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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 November 2016 (10.11.2016)

WIPO | PCT

(10) International Publication Number
WO 2016/179268 A1

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

(21) International Application Number:
PCT/US2016/030762

(22) International Filing Date:
4 May 2016 (04.05.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/157,284 5 May 2015 (05.05.2015) US

(71) Applicant: **UNIVERSITY OF FLORIDA RESEARCH FOUNDATION, INC.** [US/US]; 223 Grinter Hall, Gainesville, FL 32611 (US).

(72) Inventor: **PECK, Ammon, B.**; 346 NW 50th Boulevard, Gainesville, FL 32610 (US).

(74) Agent: **WALSH, Edmund, J.**; Wolf, Greenfield & Sacks, P.C., 600 Atlantic Avenue, Boston, MA 02210-2206 (US).

(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, 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: DIAGNOSTIC MICROARRAY, SYSTEM AND METHOD

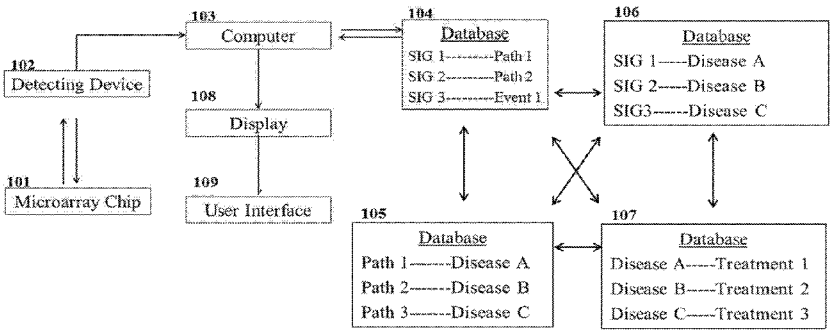

FIG. 2

(57) Abstract: In some aspects, the disclosure relates to devices, systems and methods for identifying biomarkers. In some aspects, the biomarkers are interferon-related genes (IRGs). In some embodiments, the devices are nucleic acid microarray substrates.

WO 2016/179268 A1

DIAGNOSTIC MICROARRAY, SYSTEM AND METHOD

RELATED APPLICATIONS

This Application claims the benefit under 35 U.S.C. 119(e) of U.S. provisional application USSN 62/157,284, filed May 5, 2015, entitled "DIAGNOSTIC MICROARRAY, SYSTEM AND METHOD," the entire contents of which are incorporated by reference herein.

BACKGROUND

Biomarkers are measurable indicators for the presence and/or severity of a particular disease state. Examples of biomarkers include proteins (*e.g.,* cellular antigens or autoantibodies), expression levels of genes associated with disease (*e.g.,* tumor suppressor genes) and the presence of single nucleotide polymorphisms (SNPs) in the genome of a subject. Biomarkers may be used to diagnose a patient with a particular disease or to predict the response of a patient to specific therapeutic regimens. However, currently used biomarkers are often indirect or non-specific markers of disease, which limits their usefulness. For example, a single gene or gene product used as a biomarker may be dysregulated or differentially expressed in several disease states with unrelated pathophysiologies, such that a single biomarker may not provide meaningful information about the patient's disease.

SUMMARY

Provided herein are compositions and methods related to biomarker assays that directly correlate to the etiopathological mechanisms of the disease they identify. Accordingly, in some aspects, the disclosure provides a diagnostic array, comprising a substrate with a surface having a plurality of locations, wherein each of the plurality of locations includes at least one oligonucleotide attached to the substrate at a predetermined location, each location of the plurality of locations includes at least one material targeting an interferon-regulated gene (IRG), and the material at the plurality of locations target at least 100 interferon-regulated genes.

In some embodiments, the surface comprises between 1 material and between 5 materials that target a non-IRG. In some embodiments, the surface does not comprise material targeting non-IRGs. In some embodiments, the at least one material is an oligonucleotide. In some embodiments, the surface comprises at least 200, at least 300, at least 400, at least 500, at least 750, at least 1000, or at least 1500 oligonucleotides, wherein each oligonucleotide targets a different interferon-responsive gene (IRG). In some embodiments, the IRGs are type-I

interferon responsive genes. In some embodiments, the IRGs are type-II interferon responsive genes. In some embodiments, the IRGs are type-I and type-II interferon responsive genes. In some embodiments, the at least one oligonucleotide targets a gene selected from the Interferome Database. In some embodiments, each of the at least one oligonucleotides is the same length. In some embodiments, at least two oligonucleotides are not the same length. In some embodiments, at least one oligonucleotide is a plurality of oligonucleotides, wherein each oligonucleotide is the same length.

In some embodiments, each location occupies a space of less than 1 cm^2 on the substrate. In some embodiments, each location occupies a space of less than 1 μm^2 on the substrate. In some embodiments, the substrate is glass. In some embodiments, the substrate is plastic. In some embodiments, the substrate is used for identifying interferon signature in a sample obtained from a subject having or suspected of having a disease.

In some aspects, the disclosure relates to a method for identifying an interferon signature, the method comprising: contacting a substrate as described by the disclosure with a sample having nucleic acids obtained from at least one cell; determining the gene expression profile of the sample by detecting hybridization of the nucleic acids to the substrate, wherein the gene expression profile identifies the expression level of a plurality of interferon-related genes (IRGs) relative to the expression level of a plurality of IRGs in a healthy control cell (*e.g.*, a sample having nucleic acids obtained from a healthy control cell); and, assigning an interferon signature to the sample based upon the gene expression profile of the sample.

In some embodiments, the cell is cell obtained from a subject having or suspected of having a pathological condition. In some embodiments, the pathological condition is an interferon-related disease. In some embodiments of the method, the IRGs are type-I IRGs. In some embodiments, the IRGs are type-II IRGs. In some embodiments the IRGs comprise a combination of type-I and type-II IRGs. In some embodiments, the cell is derived from blood, or a tissue biopsy. In some embodiments of the method, the nucleic acid is DNA.

In some embodiments, the plurality of IRGs comprises at least 200, at least 300, at least 400, at least 500, at least 750, at least 1000, or at least 1500 IRGs. In some embodiments, the interferon signature identifies IRGs having upregulated expression levels relative to healthy control cells. In some embodiments, the interferon signature identifies IRGs having down regulated expression levels relative to healthy control cells.

In some embodiments, the method further comprises comparing the interferon signature to a database, wherein the database comprises known interferon signatures associated with molecular events. In some embodiments, the method further comprises assigning a molecular event to the interferon signature that is identified if the interferon signature matches an interferon signature in the database.

In some aspects, the disclosure provides a method for identifying an interferon signature, the method comprising: determining the gene expression profile of a sample having nucleic acids obtained from at least one cell, wherein the gene expression profile identifies the expression level of a plurality of interferon-related genes (IRGs) in the sample relative to the expression level of a plurality of IRGs in a healthy control cell; and, assigning an interferon signature to the sample based upon the gene expression profile of the sample.

In some embodiments, the gene expression profile is determined by detecting hybridization of the nucleic acids of the sample to a plurality of IRG-detecting materials in a solution (*e.g.*, where the IRG-detecting materials are not attached to a substrate). In some embodiments, the IRG-detecting materials are nucleic acid (*e.g.*, oligonucleotide) probes. In some embodiments, each IRG-detecting material of the plurality hybridizes to a type-I IRG. In some embodiments, the gene expression profile is determined by nucleic acid sequencing (*e.g.*, DNA sequencing or RNA sequencing) or protein sequencing. In some embodiments, method further comprises analyzing the sequencing data by gene set enrichment analysis (GSEA).

In some aspects, the disclosure provides a method for determining a personalized therapeutic regimen for a subject, the method comprising obtaining a sample from a subject; determining an interferon signature of the sample; with at least one processor, comparing the interferon signature of the sample to a database of known interferon signatures, wherein each interferon signature of the database is associated with a molecular event; assigning the interferon signature of the sample to a molecular event, wherein the assignment is made by matching the interferon signature of the sample to an interferon signature of the database; and, prescribing a therapeutic regimen based upon the molecular event assigned to the interferon signature of the sample. In some embodiments, the interferon signature of the sample is determined by a method for identifying an interferon signature as described by the disclosure.

In some embodiments, the database comprises only type-I interferon signatures. In some embodiments, the database comprises only type-II interferon signatures. In some embodiments, the database comprises both type-I and type II interferon signatures. In some embodiments, the

database comprises at least 200, at least 500, at least 1000, or at least 1500 interferon signatures. In some embodiments, the database comprises at least 10, at least 20, at least 50, at least 100, at least 200, at least 500, at least 1000, at least 5000, at least 10000, at least 20000, or up to about 30000 molecular events.

In some embodiments, the therapeutic regimen comprises administering a molecule that modulates the molecular event assigned to the interferon signature of the sample to the subject. In some embodiments, the subject is a human.

In some aspects, the disclosure provides a system for identifying interferon signature in a sample, including: a substrate as described by the disclosure; a microarray scanner; a computer; and a database containing interferon signatures and associated diseases.

In some embodiments, the system further comprises a second database containing a plurality of diseases and associated therapeutic regimen for each disease. In some embodiments, the system comprises a third database containing interferon signatures and associated molecular events. In some embodiments, the database, second database and/or third database are electronically connected.

BRIEF DESCRIPTION OF DRAWINGS

FIG. 1 shows a scheme depicting the interactive roles of Interferon, Trim and Socs molecules regulating the innate response in SjS-susceptible C57BL/6.NOD-*Aec1Aec2* mice. Signal transductions following activation of the Ifnα/β receptor involves the Jak/Tyk-Stat1/Stat2 pathway. Irf9 acts as a transcription factor that is involved in the activation of Trim molecules, many of which are E3-like ubiquinating molecules known to interact at multiple points of viral infections, thus functioning as anti-microbial factors. Two major regulators of the Infαβ signaling pathway that help define the IFN-signature are Socs1 and Socs3 molecules that interfere with the activation loop of Jak kinase, thereby preventing phosphorylation of Stat molecules. Trim8, Pias1 and Pias2, in turn, function as inhibitors of Socs1, promoting continuation of Infαβ-signaling. Similarly, Trim21 stabilizes the function of Irf3 through blocking its interaction with Pin1, thereby promoting IFN-signaling. In disease states, genes encoding molecules that function to inhibit the innate response (*e.g.*, Socs1, Trim27, Trim30 and Trim40) can be shown to be down-regulated, while genes encoding factors that function to activate innate responses can be shown to be up-regulated. This scheme is consistent with the strong IFN-signature observed in rheumatic diseases, such as Systemic lupus erythematosus, Sjögren's syndrome, and Rheumatoid arthritis.

FIG. 2 shows one embodiment of a system described herein. A device as described herein, for example a microarray chip (**101**) treated with a patient sample, is scanned with a detecting device (**102**) to produce data. The data generated from the sample is transmitted to a computer (**103**), containing software used to assign an "interferon signature" to the sample. Information concerning the interferon signature of the sample is transmitted to a network of interconnected databases (**104-107**) comprising a library of interferon signatures associated with molecular events and their associated biological pathways, diseases correlated to molecular events, and treatment recommendations associated with interferon signatures, molecular events and their associated diseases. The interferon signature of the sample is matched to an interferon signature in the database defining a specific disease. The result of a match is a database-mediated (**104**) assignment of molecular events associated with the matched interferon signature to the sample. Once a molecular event has been identified, the remaining databases are contacted to identify a disease or diseases associated with said event (**105**), or associated with the interferon signature (**106**). When the specific disease is identified, a database comprising treatment options associated with each disease is searched (**107**). The resulting molecular event, disease and treatment options are transmitted back to the computer (**103**) and displayed (**108**).

FIG. 3 shows one embodiment of an example of a suitable computing system environment.

FIG. 4 shows a flowchart describing analysis of temporal differential expression of genes involved in altered biological processes and/or molecular pathways between C57BL/6J and C57BL/6.NOD-*Aec1Aec2* mice.

FIG. 5 shows biological processes altered (activated/up-regulated) at 8 weeks versus 4 weeks of age in the salivary glands (SG) of C57BL/6.NOD-*Aec1Aec2* mice compared to C57BL/6J controls.

FIG. 6 shows biological processes down-regulated at 12-16 weeks versus 8 weeks of age in the GS of C57BL/6.NOD-*Aec1Aec2* mice compared to C57BL/6J controls. Many GS up-regulated between 4-8 weeks of age are down-regulated at between 8-12 weeks of age.

FIG. 7 shows biological processes altered (activated/up-regulated) at 16 weeks of age in the SG of C57BL/6.NOD-*Aec1Aec2* mice compared to C57BL/6J controls.

FIG. 8 shows examples of gene set (GS) clusters of interest for pre-disease to early-onset SjS.

FIG. 9 shows an example of gene sets (GS) characterizing early SjS pathology.

FIG. 10 shows an example of gene sets (GS) characterizing early SjS pathology.

FIG. 11 shows an example of gene sets (GS) characterizing early SjS pathology.

FIG. 12 shows a schematic depiction of one example of the interferon signature of TLRs and Trims.

FIG. 13 shows interaction of the IFN-pathway with TAM receptor tyrosine kinase (TAM-RTK)-encoding genes.

FIG. 14 shows a schematic describing interaction of Znf512β with TGFβ signaling. Upregulation of Znf512β and TGFβ leads to autoimmunity in B6-*Aec1Aec2* mice.

FIG. 15 shows differential expression of the Zfp-Myc-TGFβ gene axis in lacrimal glands of the B6.NOD-*Aec1Aec2* mouse.

DETAILED DESCRIPTION

The inventor has recognized and appreciated approaches for identifying, from a specimen from a patient, a specific disease and/or treatment appropriate for the patent. These techniques involve novel disease-specific biomarkers and biomarker assays that directly correlate to the etiopathological mechanisms of the disease they identify. In some embodiments, a microarray containing multiple regions (*e.g.*, locations) responsive to interferon-responsive genes is used. Different ones of the regions may be responsive to different interferon-responsive genes or gene products. In accordance with some embodiments, these genes or gene products may be associated with biological pathways. By processing test results derived from locations activated by a sample based on those pathways, signatures may be developed that indicate appropriate patient-specific actions, including reporting a disease state of the patient or selecting a treatment.

The inventor has recognized and appreciated that one of the major challenges for the use of biomarkers for the identification of disease is that currently used biomarkers are often indirect or non-specific markers of the pathological state. For example, a single biomarker (*e.g.,* a gene or gene product) may be indicative of several unrelated diseases. The instant disclosure provides devices and methods to improve the diagnostic capabilities of biomarkers. In some aspects, the invention relates to processing test data in accordance with the inventor's recognition and appreciation that the presence of combinations of biomarkers with unique characteristics, *i.e.*, "signatures", represent multiple sets of direct indicators of the underlying biological processes and molecular events controlling the pathology of specific diseases.

More specifically, the disclosure relates to the signatures associated with the innate immune response of a host to certain diseases. Innate immune responses direct the nature and specificity of downstream adaptive responses in autoimmune diseases. One of the strongest markers of innate immunity is the up-regulated expression of interferon (IFN) and IFN-related genes (IRGs). For example, Sjögren's syndrome (SjS), an autoimmune rheumatic disease, is characterized by chronic immune attacks against exocrine glands leading to exocrine dysfunction, plus strong up-regulated expressions of IFN and a unique set of IRG transcripts. However, it should be appreciated that interferon signatures are not limited to autoimmune diseases, and may be present in all innate immune responses. In some aspects, therefore, the disclosure relates to the inventor's recognition and appreciation that a sufficient number of IRGs induced during host responses may be detected and, when considered in tandem, provide a high probability of uniquely identifying "interferon signatures" for each different disease. Such interferon signatures may be associated with molecular events, allowing reliable prediction of etiopathological mechanisms of disease, a treatment for a specific disease suffered by a patient, or otherwise identify information relevant to a specific patient based on a predetermined association between that information and one or more interferon signatures.

Devices for Identifying Interferon Signatures

Provided herein are devices and methods for identifying interferon signatures of a sample. In some embodiments, the devices are substrates useful for microarray assays. In some aspects, the disclosure provides a substrate with a surface comprising multiple locations that are treated to react with an interferon-responsive gene (IRG). In accordance with some embodiments, there may be at least 100 such locations. In some embodiments, the locations

may be made responsive to an IRG as a result of a material adhered to the substrate. Different material may be applied in different locations to enable the microarray to detect the presence of different IRGs in a sample applied to the microarray. These materials may include one or more oligonucleotides. As a specific example, multiple oligonucleotides may be applied to some or all of the locations of the microarray, such that each cell of the array represents a group of oligonucleotides. Oligonucleotides are used herein as an example of suitable materials.

These materials may be attached to the substrate in any suitable way. In some embodiments, the material may be covalently bonded. As a specific example, the materials may be covalently attached to the substrate at a predetermined location, and each location may target a different interferon-regulated gene (IRG). In some embodiments, more than one location targets the same IRG. In some embodiments, the substrate further comprises materials that are reactive to non-IRGs or IRG gene products. For example, a substrate may comprise a material reactive to actin or a gene encoding actin that may be used as a reference material. In some embodiments, a substrate may comprise between 1 material and 5 materials that are reactive with non-IRGs or IRG gene products.

As used herein, the term "interferon signature" refers to a unique identifier comprising expressional data for a specific combination of IRGs that are modulated in response to interferon production caused by a particular disease or pathogen. In some embodiments, an interferon signature comprises only type-I IRGs. In some embodiments, an interferon signature comprises at least 100, at least 200, at least 300, at least 400, at least 500, at least 1000, or at least 1500 IRGs.

As used herein, the term "substrate" refers to any suitable structure to which a molecule or molecules may be attached. A substrate may, but need not, be planar. In some embodiments, the substrate may serve as a scaffold for the attachment of a molecule or molecules, including for example nucleic acids- that produce a detectible reaction when coming into contact with one or more IRGs. Such a substrate may be constructed using techniques known in the art. For example, it is known to make microarray chips comprising a plurality of molecular probes (*e.g.*, oligonucleotides) covalently attached to a substrate. In some embodiments, the substrate is glass. In some embodiments, the substrate may be silicon or polymer (*e.g.*, polystyrene or polycarbonate). Substrates described herein may be square, circular, rectangular, triangular or any other shape that does not prohibit its proper function as understood by a person of ordinary skill in the art. In some embodiments, the area of the surface covalently bound to the

oligonucleotides ranges in size from 1 μm^2 to 10 cm^2. In some embodiments, the area of the surface ranges in size from 10 μm^2 to 5 cm^2. In some embodiments, the area of the surface ranges in size from 100 μm^2 to 1 cm^2.

In some aspects, the substrate comprises at least 100 locations with oligonucleotides covalently bound to its surface. As used herein, the term "oligonucleotide" refers to a single stranded, polymeric nucleic acid molecule. In some embodiments, the oligonucleotides are deoxyribonucleic acid (DNA). In some embodiments, the oligonucleotides are ribonucleic acid (RNA). Methods for covalently attaching oligonucleotides are known in the art and may be performed as described, for example in US Patent No. 5, 445,934, which is incorporated herein by reference in its entirety.

The oligonucleotides described herein may act as probes to which the nucleic acids of a sample may hybridize. Accordingly, the length of the oligonucleotides may vary depending upon the target sequence to which it is designed to hybridize. In some embodiments, the oligonucleotides range between 5 base pairs (bp) and 500 bp in length. In some embodiments, the oligonucleotides range between 10 bp and 250 bp in length. In some embodiments, the oligonucleotides range between 15 bp and 150 bp in length. In some embodiments, the oligonucleotides range between 20 bp and 100 bp in length. In some embodiments, the oligonucleotides range between 30 bp and 90 bp in length. In some embodiments, the oligonucleotides range between 40 bp and 80 bp in length. In some embodiments, the oligonucleotides range between 50 bp and 70 bp in length. In some embodiments, the oligonucleotides range between 19 bp and 25 bp in length.

In some embodiments, the length of each oligonucleotide within a group of oligonucleotides covalently attached to a substrate are each the same length. In some embodiments, at least two groups of oligonucleotides comprise oligonucleotides of the same length.

Varying levels of complementarity between oligonucleotides and target sequences are also contemplated herein. In some embodiments, the oligonucleotide and target sequences are 100% complementary. However, a response may be detected with less than 100% complementarity. In some embodiments, the oligonucleotide and target sequences are at least 99% complementary. In some embodiments, the oligonucleotide and target sequences are at least 95% complementary. In some embodiments, the oligonucleotide and target sequences are at least 90% complementary. In some embodiments, the oligonucleotide and target sequences

are at least 85% complementary. In some embodiments, the oligonucleotide and target sequences are at least 80% complementary. In some embodiments, the oligonucleotide and target sequences are at least 75% complementary. In some embodiments, the oligonucleotide and target sequences are at least 70% complementary. In some embodiments, the oligonucleotide and target sequences are at least 65% complementary. In some embodiments, the oligonucleotide and target sequences are at least 60% complementary.

In some aspects, the disclosure relates to interferon-regulated genes. Interferons are cytokines (signaling molecules) produced by cells of the immune system in response to a pathogen, tumor, and/or cell injury. There are three types of interferon protein (type-I, type-II and type III). Type-I interferons are of particular importance because in addition to functioning as part of the innate immune system, they stimulate dendritic cells and thus prime the adaptive immune response. The production of interferons results in the activation of several signaling cascades, for example the Janus kinase-signal transducer and activator of transcription (JAK/STAT) pathway, which induces the transcription of various genes that are termed "interferon-regulated genes" (IRGs). As used herein, the term interferon-regulated gene (IRG) refers to genes that are expressionally regulated by interferon proteins. Expression of IRGs may therefore be up-regulated or down regulated in response to interferon. Non-limiting examples of IRGs include SOCS1, SOCS3, TRIM8, STAT1, IRF1 and those listed in the INTERFEROME database (Samarajiwa et al., Interferome: the database of interferon-regulated genes, Nucleic Acids Res., 37: D852-D837, 2009; incorporated by reference herein in its entirety).

Without wishing to be bound by any particular theory, the inventors have recognized and appreciated that patterns of IRGs present in a sample having or suspected to have a particular disease or disorder form a unique interferon signature that is associated with the molecular events causing pathology associated with said disorder. Processing test data based on associations with these events may provide more accurate results, such as in diagnosis or treatment of a patient. Thus, in some aspects, the invention relates to a substrate comprising groups of oligonucleotides that target IRGs. It is to be appreciated that each oligonucleotide within a group targeting an IRG may be the same or may be different. For example in some embodiments, a group of oligonucleotides targeting the IRG toll-like receptor 4 (*TLR4*) may comprise oligonucleotides that each have the same oligonucleotide sequence (*e.g.,* GAGAGGA). Alternatively, in some embodiments, a group of oligonucleotides targeting the IRG toll-like receptor 4 (*TLR4*) may comprise oligonucleotides of different sequences that target discrete

locations of the *TLR4* gene. Thus, in some embodiments, each group of oligonucleotides covalently bound to the substrate targets a different IRG. In some embodiments, the substrate comprises at least 100 groups of oligonucleotides that each target a different IRG. In some embodiments, the substrate comprises at least 200, at least 300, at least 400, at least 500, at least 1000, at least 1200, or at least 1500 groups of oligonucleotides that each target a different IRG.

The disclosure also contemplates the spatial arrangement of the IRG-targeting oligonucleotides on the surface of the substrate. In some embodiments, a plurality of IRG-targeting oligonucleotides is arranged into groups, and each group is attached to a predetermined location on the surface of the substrate. Oligonucleotides may be grouped based upon any definable characteristic, for example the size of the oligonucleotide or the particular IRG targeted by the oligonucleotide. Therefore, in some embodiments, a location may comprise a group of oligonucleotides comprising a plurality of oligonucleotides where each oligonucleotide targets the same IRG. The skilled artisan recognizes that although each oligonucleotide within a group at a predetermined location may target the same IRG, the oligonucleotides need not have identical sequences. For example, a group may comprise a plurality of oligonucleotides, each oligonucleotide of the plurality targeting a different sequence fragment within the same IRG. In some embodiments a location may comprise a group comprising a plurality of oligonucleotides, wherein each oligonucleotide of the plurality targets a different IRG. For example, a location may comprise a group of oligonucleotides that targets at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 100, at least 500, at least 1000 or at least 1200 IRGs. In some embodiments, a group comprises a plurality of oligonucleotides, wherein each oligonucleotide of the plurality targets an IRG associated with a particular molecular event. For example, a group may comprise a plurality of oligonucleotides targeting IRGs involved with TLR-4-mediated signaling.

As used herein, a "predetermined location" refers to a localized area on a substrate at which a selected oligonucleotide is covalently attached. This location may therefore be predetermined at the time of that attachment. In some embodiments, the substrate may be divided into uniformly sized and distributed locations in a pattern that is the same for substrates used in sample processing. Accordingly, a machine reading results from the substrate may be programmed in advance to find every location by checking locations in accordance with the pattern. In other embodiments, information encoded on the substrate may indicate the

predetermined locations, which a machine may then determine based on the sensing the information encoded on the substrate.

As a specific example, the attachment of each oligonucleotide at a predetermined location enables each group of IRG-targeting oligonucleotides to be assigned an address (*e.g.*, x- and y-coordinates) on the substrate. During analysis, the signal at a particular address on the substrate may then provide information regarding the IRG associated with said address.

Kits comprising IRG-targeting materials (*e.g.*, IRG-targeting oligonucleotides) are also contemplated by the disclosure. For example, in some aspects the disclosure provides a kit comprising a plurality of containers, each container housing an IRG-targeting oligonucleotide. In some embodiments, each IRG-targeting oligonucleotide specifically hybridizes to a different IRG (for example, hybridizes to an IRG while in solution). In some embodiments, the IRG-targeting materials (*e.g.*, IRG-targeting oligonucleotides) provided by the kit specifically hybridize to each of the IRGs listed in the INTERFEROME database (Samarajiwa et al., Interferome: the database of interferon-regulated genes, Nucleic Acids Res., 37: D852-D837, 2009; incorporated by reference herein in its entirety). The IRG-targeting oligonucleotide can be provided in solution (*e.g.*, dissolved in a solvent) or as a solid (*e.g.*, a lyophilized powder).

Methods for identifying interferon signatures

In some aspects, the invention relates to the discovery that interferon signatures derived from the detection of the expression level of a sufficient number of IRGs can be used as biomarkers to identify the biological processes and molecular events defining a specific disease indicated by said interferon signature. In some cases, a therapeutic regimen may be prescribed based upon the disease or molecular events associated with the full interferon signature or individual parts of the interferon signature.

Therefore, in some aspects the disclosure provides a method for identifying an interferon signature, the method comprising (i) contacting a substrate as described herein with a sample comprising nucleic acids; (ii) determining the interferon gene expression profile of the sample by detecting hybridization of the nucleic acids to the substrate, wherein the gene expression profile identifies the expression level of a plurality of interferon-regulated genes (IRGs) relative to the expression level of a plurality of IRGs in a healthy control cell; and (iii) assigning an interferon signature to the sample based upon the gene expression profile.

In some embodiments of the method, the substrate has a surface comprising at least 100 oligonucleotides, wherein each oligonucleotide is covalently attached to the substrate at a predetermined location, and wherein each oligonucleotide targets a different interferon-regulated gene (IRG). In some embodiments, the IRGs are type-I IRGs. In some embodiments, the IRGs are type-II IRGs. In some embodiments, the nucleic acids are DNA or RNA.

As used herein, the term "sample" refers to a biological specimen. For example, a sample may refer to blood, sweat, urine, fecal matter, seminal fluid, a population of isolated tissue cells, isolated organ cells, a lysate of tissue cells, or a lysate of organ cells. Samples may be suitably processed to allow contact between nucleic acids of the sample and a substrate. For example, cells may be lysed and RNA extraction and purification may be performed prior to contacting the substrate with the sample. In some embodiments, the population of cells is heterogeneous. In some embodiments, the population of cells is homogenous.

In some embodiments of the method, the cell is a eukaryotic cell. In some embodiments, the cell is a human cell. Various cell types are envisioned by the disclosure. In some embodiments, the cell is a blood cell. In some embodiments, the cell is an immune cell. Examples of immune cells include but are not limited to iLC cell, T cell, B cell, Natural Killer cell, dendritic cell, macrophage, or any other immune cell. In some embodiments, the cell is a cell derived from an organ, for example kidney cell, lung cell, epithelial cell, endothelial cell, skin cell, cardiac cell, liver cell, intestinal cell, or any other cell derived from an organ. In some embodiments, the cell is a cancer cell. In some embodiments, the cell is a tumor cell.

In some embodiments of the method, the cell is derived from a subject. As used herein, the term "subject" refers to an animal having or suspected of having a disease, or an animal that is being tested for a disease. In some embodiments, the subject is a human. In some embodiments, the subject is selected from the group consisting of human, non-human primate, rodent (*e.g.,* mouse or rat), canine, feline, or equine. In some embodiments, the disease is a bacterial, viral, parasitic or autoimmune disease. In some embodiments, the disease is related to a mutation in the genome of the subject, for example cancer resulting from the mutation of a cancer suppressor gene. In some embodiments, the disease is related to a chromosomal abnormality, such as a chromosomal deletion, in the genome of the subject. In some embodiments, the disease expresses high numbers of IRGs. Examples of diseases that express high numbers of IRGs include but are not limited to rheumatoid arthritis (RA), Sjögren's

syndrome (SjS) and systemic lupus erythematosus (SLE). In some embodiments, the disease expresses low numbers of IRGs, for example kidney stone diseases.

Non-limiting examples of pathogenic bacteria include but are not limited to Salmonella, Yersinia, Shigella, Campylobacter, Helicobacter, Pseudomonas, Streptococcus, Staphylococcus, E. coli, Haemophilus, Mycobacterium, Proteus, Klebsiella, Neisseria, Branhamella, Bacteroides, Listeria, Enterococci, Vibrio, Bordetella, Clostridium, Treponema, and Mycoplasma.

Non-limiting examples of pathogenic viruses include but are not limited to human immunodeficiency virus (HIV), cowpox virus, Coxsackie virus, dengue virus, hepatitis virus (A, B, C), Ebola virus, smallpox virus, Epstein-Barr virus, human cytomegalovirus, human herpesvirus, human papillomavirus, influenza virus, measles virus, polio virus, and vaccinia virus.

Non-limiting examples of pathogenic parasites include but are not limited Plasmodium, Trypanosoma, Ascaris, Schistosoma, Cryptosporidium, Taenia, and Leishmania.

Autoimmune diseases include but are not limited to rheumatoid arthritis (RA), Sjögren's syndrome (SjS), lupus, Addison's disease, Grave's disease, celiac disease, multiple sclerosis (MS), inflammatory bowel disease, Crohn's disease and type-I diabetes.

In some embodiments, the disease is cancer. Non-limiting examples of cancer include breast cancer, lung cancer, liver cancer, bowel cancer, intestinal cancer, cervical cancer, brain cancer, kidney cancer, T cell lymphoma, B cell lymphoma, prostate cancer, testicular cancer, stomach cancer, and osteosarcomas.

In some aspects, the method comprises determining the gene expression profile of a sample. As used herein, the term "gene expression profile" refers to the measurement of activity (*i.e.,* expression) of a plurality of genes from a sample at a specific time point. The gene expression profile of a sample having or suspected of having a disease may be conveyed as relative expression of a subset of genes from the sample compared to the same subset of genes in a control cell. For example, expression of various IRGs associated with pro-inflammatory cytokines may be up-regulated in a sample suspected of having rheumatoid arthritis compared to a control cell that does not have an inflammatory disease. In some embodiments, a gene expression profile is determined relative to healthy control cell. As used herein, "healthy control cell" refers to a cell that does not have or is not suspected of having a disease or disorder. In some embodiments, the healthy control cell is the same cell type as the cell from which the sample is derived.

In some embodiments, the plurality of genes of the gene expression profile comprises only IRGs. In some embodiments, the expression profile comprises at least 100 IRGs. In some embodiments, the expression profile comprises at least 200, at least 300, at least 400, at least 500, at least 1000, at least 1200, or at least 1500 IRGs.

In some aspects, the gene expression profile of a sample is determined by detecting the presence or absence of an interaction between the sample and a material at each location on a substrate. Various methods of detecting an interaction are recognized in the art. For example, interaction between the sample and the material can be detected by measuring binding activity between the sample and the material. As used herein, the term "binding activity" refers to the chemical linkage formed between two molecules. For example, a protein ligand may become covalently bound to its cognate receptor via the chemical interaction between the amino acid residues of the ligand and the receptor. In the context of nucleic acid interactions, binding activity includes the hybridization of complementary nucleic acids. As used herein, the term "hybridization" is accorded its general meaning in the art and refers to the pairing of substantially complementary nucleotide sequences (for example, pairing of oligonucleotides and strands of nucleic acid) to form a duplex or heteroduplex through formation of hydrogen bonds between complementary base pairs in accordance with Watson-Crick base pairing. Hybridization is a specific, *i.e.*, non- random, interaction between two complementary polynucleotides. In some embodiments of the method, hybridization occurs when the nucleic acids of the sample are 100% complementary to the oligonucleotides covalently attached to the substrate. In some embodiments of the method, hybridization occurs when the nucleic acids of the sample are at least 99% complementary to the oligonucleotides covalently attached to the substrate. In some embodiments of the method, hybridization occurs when the nucleic acids of the sample are at least 95% complementary to the oligonucleotides covalently attached to the substrate. In some embodiments of the method, hybridization occurs when the nucleic acids of the sample are at least 90% complementary to the oligonucleotides covalently attached to the substrate. In some embodiments of the method, hybridization occurs when the nucleic acids of the sample are at least 85% complementary to the oligonucleotides covalently attached to the substrate. In some embodiments of the method, hybridization occurs when the nucleic acids of the sample are at least 80% complementary to the oligonucleotides covalently attached to the substrate. In some embodiments of the method, hybridization occurs when the nucleic acids of the sample are at least 75% complementary to the oligonucleotides covalently attached to the

substrate. In some embodiments of the method, hybridization occurs when the nucleic acids of the sample are at least 70% complementary to the oligonucleotides covalently attached to the substrate. In some embodiments of the method, hybridization occurs when the nucleic acids of the sample are at least 65% complementary to the oligonucleotides covalently attached to the substrate. In some embodiments of the method, hybridization occurs when the nucleic acids of the sample are at least 60% complementary to the oligonucleotides covalently attached to the substrate. Therefore, in some embodiments, the gene expression profile is determined by detecting hybridization of the nucleic acids to a substrate with a surface covalently linked to oligonucleotides.

In some embodiments, the gene expression profile of a sample is determined by nucleic acid sequencing (*e.g.*, DNA sequencing, RNA sequencing, *etc.*). Examples of sequencing methods used for gene expression profiling include but are not limited to single-molecule real-time sequencing (SMRT), ion semiconductor (Ion Torrent) sequencing, pyrosequencing, sequencing by synthesis (*e.g.*, Illumina sequencing), sequencing by ligation (SOLiD), and chain termination sequencing (Sanger sequencing), nanopore sequencing, and massively parallel sequencing (MPSS). Sequencing methods generally utilize gene specific probes (*e.g.*, oligonucleotides, primers, adaptors, *etc.*) for nucleic acid amplification. In some embodiments, IRG-targeting oligonucleotides function as gene-specific probes for a sequencing reaction.

In some aspects, the disclosure relates to assigning an interferon signature to the sample based upon its gene expression profile. Interferon signatures provide information pertaining to the relative expression levels of specific combinations of IRGs that are modulated in response to interferon production caused by a particular disease or pathogen. For example, an interferon signature of a sample having an inflammatory disease may indicate that a specific subset of IRGs related to pro-inflammatory molecules, such as TNFα, are 5-fold up-regulated relative to a healthy control. The invention is based, in part, on the discovery that detection of interferon signatures comprising combinations of a sufficient number of IRGs are useful for identifying specific molecular events that underlie pathophysiology in a subject and determining a therapeutic regimen for the subject. In some embodiments, an interferon signature comprises at least 100 IRGs. In some embodiments, an interferon signature comprises at least 200, at least 300, at least 400, at least 500, at least 1000, at least 1200, or at least 1500 IRGs.

Often, shared etiopathological mechanisms between diseases result from common molecular events or biological processes related to the host immune response. Although

different diseases have unique and specific interferon signatures, common sets of IRGs within the different signatures may overlap due to shared underlying etiopathological mechanisms between the different diseases. Thus, the instant invention relates, in part, to the surprising discovery that the detection of an interferon signature comprising a sufficient number of IRGs can be used to identify the etiopathophysiology associated with a specific disease, even in the presence of overlapping molecular events common to multiple diseases. Without wishing to be bound by any particular theory, identification of the etiopathophysiology associated with a specific disease allows for the selection of an appropriate therapeutic regimen to target the underlying cause of the disease.

Accordingly, in some aspects, the method further comprises comparing the interferon signature of the sample to a database, wherein the database comprises known interferon signatures associated with molecular events. In some embodiments, if the interferon signature of the sample matches an interferon signature in the database, the molecular event associated with the signature of the database may be assigned to the interferon signature of the sample and compared to a database comprising unique molecular events associated with specific diseases. The interferon signature of the sample may also be compared to a database comprising known interferon signatures associated with specific diseases. The diseases may be associated directly or indirectly with interferon signatures. For example, in some embodiments, signatures may be associated with molecular events (such as pathways activated), and the combination of one or more molecular events may be uniquely associated with specific diseases. Alternatively or additionally, the interferon signature or molecular event assigned to the sample can be compared to a database comprising known effective therapeutic regimens for a subject exhibiting the interferon signature. Such association also may be direct or indirect, with a database linking signatures to treatments directly, or with databases linking signatures to diseases and linking diseases to treatments.

The skilled artisan recognizes that the information contained in the foregoing databases may be combined into a single database or data tables within a database. Moreover, it should be appreciated that "database" is used generically and does not imply any dependency on a format in which such information is stored. It should be appreciated that databases as described herein may also be electronically networked together. Therefore, in some embodiments, the method further comprises identifying the molecular event(s) associated with an interferon signature, diagnosing the subject with a disease based upon the presence of the interferon signature and the

molecular event(s) and selecting a therapeutic regimen based upon the disease diagnosis. These actions may be performed by a processor or processors accessing one or multiple databases over a computer network.

As used herein, the term "molecular event" refers to a biological process that directly or indirectly results from the induction or suppression of IRGs by a pathological condition. Examples of molecular events include but are not limited to signal transduction pathways (*e.g.*, TLR4 signal transduction pathway) and disease associated events (*e.g.*, inflammatory response induced by bacterial infection). In some embodiments, the database comprises at least 100, at least 200, at least 300, at least 400, at least 500, at least 600, at least 700, at least 800, at least 900, at least 1000, at least 2000, at least 3000, at least 5000, at least 10,000, at least 20,000, or at least 30,000 molecular events associated with interferon signatures. In some embodiments the database comprises no more than about 30,000 molecular events associated with interferon signatures. Matching the interferon signature of the sample to an interferon signature in the database indicates the presence of the molecular event or biological pathway associated with the database interferon signature in the sample.

The determination of the molecular events underlying a pathological condition may be useful in the context of personalized medicine and/or pharmacogenomics. As used herein, "personalized medicine" refers to the customization of medical decisions and/or treatments based upon the genetic makeup of a patient. For example, breast cancers that overexpress the HER2 gene are sensitive to treatment with the monoclonal antibody trastuzumab. Therefore, breast cancer patients may be genotyped to identify whether their cancer is HER2-positive and if so, treated with trastuzumab. However unlike cancer, the underlying etiological causes of many other diseases remain unclear. Without wishing to be bound by any particular theory, identification of molecular events by identification of and association with specific interferon signatures may provide new therapeutic strategies and targets by recognizing previously unknown pathological mechanisms of a disease.

Aspects of the disclosure relate to the inventor's recognition that the interferon signature of a particular disease may change depending upon the biological pathways and molecular events that are regulated as the disease develops or progresses. For example, the early (*e.g.*, covert) stages of a disease may exhibit an interferon signature characterized by upregulation of innate immune response IRGs (*e.g.*, type-I IRGs), whereas later (*e.g.*, overt) stages of the same disease may exhibit an interferon signature characterized by upregulation of adaptive immune

response IRGs (*e.g.*, type-II IRGs). Accordingly, in some aspects, the disclosure provides a method for monitoring the progression of a disease, the method comprising determining a gene expression profile from a sample obtained from a subject having or suspected of having a disease on a first occasion; assigning an interferon signature to the sample; determining a gene expression profile from a sample obtained from the subject on a second occasion; assigning a second interferon signature to the sample; and comparing the interferon signature of the first occasion sample to the interferon signature of the second occasion sample; and, determining that the subject has the disease or the disease has progressed.

Alternatively, in some embodiments, development of a disease in an otherwise healthy subject is identified by detecting a change in interferon signatures of the subject between two time points. For example, in some embodiments, certain diseases can be identified by detecting IRG-mediated changes in biological pathways and molecular events as the disease develops over time. In some embodiments, the length of time between which an interferon signature is measured ranges from about 1 week to about 1 year. In some embodiments the length of time is about one week, four weeks (one month), three months, six months, or one year. In some embodiments, the length of time is more than one year. In some embodiments, interferon signatures change over time of disease onset, development and clinical appearance as the immune response evolves from an innate to an adaptive response. In some embodiments, the interferon signature can thus predict the stage of disease as well, whether the disease is progressing or regressing, or whether the host defense has successfully combated the disease or not. Accordingly, in some embodiments, temporally monitoring an interferon signature can therefore help determine potential treatments.

In some aspects, the disclosure provides a method for determining a personalized therapeutic regimen for a subject in need thereof, the method comprising (i) obtaining a sample from a subject; (ii) determining interferon signature of the sample; and, (iii) comparing the interferon signature of the sample to a database of known interferon signatures. In some embodiments, the interferon signature is based on a sufficient number of measurements of IRGs that interferon signatures of the database may be associated with a molecular event. In some embodiments, the interferon signature is based on measurement of the expression levels of at least 100 IRGs. The method may further comprise: (iv) assigning the interferon signature of the sample to a molecular event or biological pathway, wherein the assignment is made by matching the interferon signature of the sample to an interferon signature of the database; and, (v)

prescribing a therapeutic regimen based upon the molecular event or biological pathway assigned to the interferon signature of the sample.

In some embodiments, methods for determining a personalized therapeutic regimen described herein are useful for identifying subpopulations of responsive (or non-responsive) patients in the context of clinical trials. For example, certain diseases comprise subpopulations of patients having different etio-pathological (*e.g.*, biological pathways and molecular events) causes of disease but sharing the same symptoms (*e.g.*, the same overall biological response). Differences in the biological pathways and molecular events characterizing each subpopulation may lead to variations in drug response. Without wishing to be bound by any particular theory, methods for identifying the interferon signature of a subject described herein allow for biological pathways and molecular events characterizing subpopulations of such patients to be identified and an appropriate therapeutic regimen prescribed.

In some embodiments, the disclosure provides a method of treating a certain disease (*e.g.*, an interferon-related disease), the method comprising: administering a therapeutic regimen (*e.g.*, a drug or combination of drugs) based upon the identification of an interferon signature of the subject, wherein the interferon signature of the subject indicates the appropriate target for the drug being administered. For example, in some embodiments, a subject is administered an anti-viral therapy (*e.g.*, an anti-retroviral drug) based upon identification of an interferon signature associated with a viral disease in the subject.

As used herein, the term "therapeutic regimen" refers to the use of at least one therapeutic compound for the treatment of a disease or disorder. The term "a subject in need thereof" refers to a subject having or suspected of having a disease or disorder, or a subject having clinical signs or symptoms associated with a disease or disorder. In some embodiments, the disease or disorder is an interferon-related disease or disorder. In some embodiments the disease or disorder is a bacterial, viral, parasitic or autoimmune disease. In some embodiments, the disease or disorder is cancer.

In some aspects, the method comprises the step of prescribing a therapeutic regimen based upon the molecular event or biological pathway assigned to the interferon signature of the sample. In some embodiments, the therapeutic regimen prescribed comprises at least one therapeutic compound that enhances, inhibits or otherwise modifies the molecular event assigned to the interferon signature of the sample. For example, in some embodiments, identification of an interferon signature associated with STAT-mediated inflammation may

result in prescription of a therapeutic regimen comprising inhibitors of $NF_{\kappa\beta}$ or JAK/STAT pathways, an antiviral if the interferon signature indicates infection with a dsRNA virus, or development of new drugs targeting disease-specific biological processes utilizing IRGs.

Therapeutic compounds may be any compound that is clinically acceptable for the treatment of a particular disease. For example, therapeutic compounds may be small molecules, peptides or proteins, antibodies, nucleic acids (*e.g.,* siRNA or shRNA), transgenes, or any combination of the foregoing. In some embodiments, the therapeutic compound is a pharmaceutical composition comprising a therapeutic compound. In some embodiments, pharmaceutical compositions further comprise a pharmaceutically acceptable carrier or excipient, for example those described in Remington's Pharmaceutical Sciences, 18th Ed. (1990).

Systems for identifying interferon signatures

The disclosure also contemplates systems for identifying interferon signatures. One embodiment of a system for identifying interferon signatures is depicted in FIG. 2. In some aspects, provided herein is a system for identifying interferon signature in a sample, comprising (i) a substrate as described herein; (ii) a microarray scanner; (iii) a computer; (iv) at least one database containing known interferon signatures associated to molecular events and/or molecular events associated to specific diseases and/or specific diseases associated to therapeutic regimens.

In certain embodiments, systems comprise a substrate as described herein, for example a microarray chip (**101**) that has been treated with a sample. The chip may contain, for example, a sufficient number of locations treated with material reactive with IRG's such that results read from the chip contain enough information to differentiate molecular events within the subject from which the sample was collected. Some locations on the array may react with the same IRG.

The chip (**101**) is then scanned with a detecting device (**102**) to produce data. Scanning of such chips is known in the art, and any suitable scanning technique may be used. In some embodiments, for example, the scanning is performed with a visible imaging system, detecting changes in appearance of locations in response to a reaction between a material on the chip and the sample. The data generated from the sample is transmitted to a computer (**103**), containing software used to assign an "interferon signature" to the sample. Information concerning the

interferon signature of the sample may be compared to information in a database (**104**), comprising a library of interferon signatures associated with biological processes and molecular events. The interferon signature of the sample may be matched to an interferon signature in the database. The result of a match is database-mediated (**104**) assignment of the molecular event associated with the matched interferon signature to the sample. Once a molecular event has been identified, the remaining databases are contacted to identify a disease or diseases associated with said event (**105**), or associated with the interferon signature (**106**). Once a disease has been identified, a database comprising treatment options associated with each disease is searched (**107**). The resulting molecular event, disease and treatment options are transmitted back to the computer (**103**) and displayed (**108**).The resulting biological process or molecular event, is transmitted back to the computer (**103**) and displayed (**105**).

Representative Computing Environment

Techniques as described herein may yield more accurate diagnosis and treatment recommendations for specific subjects. Such techniques involve collecting and processing data on a sufficient number of IRGs to produce data sets including adequate information to distinguish between biological pathways triggered within a subject, even though many biological paths involve regulation of overlapping IRGs. The collection and/or processing of such data may be controlled by execution of a computing device.

FIG. 3 illustrates an example of a suitable computing system environment **400** on which such processing may be implemented. The computing system environment **400** is only one example of a suitable computing environment and is not intended to suggest any limitation as to the scope of use or functionality of the invention. Neither should the computing environment **400** be interpreted as having any dependency or requirement relating to any one or combination of components illustrated in the exemplary operating environment **400**.

The invention is operational with numerous other general purpose or special purpose computing system environments or configurations. Examples of well-known computing systems, environments, and/or configurations that may be suitable for use with the invention include, but are not limited to, personal computers, server computers, smartphones, tablets, hand-held or laptop devices , multiprocessor systems, microprocessor-based systems, set top boxes, programmable consumer electronics, network PCs, minicomputers, mainframe computers, distributed computing environments that include any of the above systems or

devices, and the like. Some of the elements illustrated in FIG. 3 may not be present, depending on the specific type of computing device. Alternatively, additional elements may be present in some implementations.

The computing environment may execute computer-executable instructions, such as program modules. Generally, program modules include routines, programs, objects, components, data structures, *etc.* that perform particular tasks or implement particular abstract data types. Some embodiments may also be practiced in distributed computing environments where tasks are performed by remote processing devices that are linked through a communications network. These distributed systems may be what are known as enterprise computing systems or, in some embodiments, may be "cloud" computing systems. In a distributed computing environment, program modules may be located in both local and/or remote computer storage media including memory storage devices.

With reference to FIG. 3, an exemplary system for implementing the invention includes a general purpose computing device in the form of a computer **410**. Components of computer **410** may include, but are not limited to, a processing unit **420**, a system memory **430**, and a system bus **421** that couples various system components including the system memory to the processing unit **420**. The system bus **421** may be any of several types of bus structures including a memory bus or memory controller, a peripheral bus, and a local bus using any of a variety of bus architectures. By way of example, and not limitation, such architectures include Industry Standard Architecture (ISA) bus, Micro Channel Architecture (MCA) bus, Enhanced ISA (EISA) bus, Video Electronics Standards Association (VESA) local bus, and Peripheral Component Interconnect (PCI) bus, also known as Mezzanine bus.

Computer **410** typically includes a variety of computer readable media. Computer readable media can be any available media that can be accessed by computer **410** and includes both volatile and nonvolatile media, removable and non-removable media. By way of example, and not limitation, computer readable media may comprise computer storage media and communication media. Computer storage media includes both volatile and nonvolatile, removable and non-removable media implemented in any method or technology for storage of information such as computer readable instructions, data structures, program modules or other data. Computer storage media includes, but is not limited to, RAM, ROM, EEPROM, flash memory or other memory technology, CD-ROM, digital versatile disks (DVD) or other optical disk storage, magnetic cassettes, magnetic tape, magnetic disk storage or other magnetic storage

devices, or any other medium which can be used to store the desired information and which can accessed by computer **410**.

Communication media typically embodies computer readable instructions, data structures, program modules or other data in a modulated data signal such as a carrier wave or other transport mechanism and includes any information delivery media. The term "modulated data signal" means a signal that has one or more of its characteristics set or changed in such a manner as to encode information in the signal. By way of example, and not limitation, communication media includes wired media such as a wired network or direct-wired connection, and wireless media such as acoustic, RF, infrared and other wireless media. Combinations of the any of the above should also be included within the scope of computer readable media.

The system memory **430** includes computer storage media in the form of volatile and/or nonvolatile memory such as read only memory (ROM) **431** and random access memory (RAM) **432**. A basic input/output system **433** (BIOS), containing the basic routines that help to transfer information between elements within computer **410**, such as during start-up, may be stored in ROM **431**. RAM **432** may contain data and/or program modules that are immediately accessible to and/or presently being operated on by processing unit **420**. By way of example, and not limitation, FIG. 3 illustrates operating system **434**, application programs **435**, other program modules **436**, and program data **437**.

The computer **410** may also include other removable/non-removable, volatile/nonvolatile computer storage media. By way of example only, FIG. 3 illustrates a hard disk drive **441** that reads from or writes to non-removable, nonvolatile magnetic media. Such a hard disk drive may be implemented by a rotating disk drive or as a solid state drive, such as is implemented with FLASH memory.

FIG. 3 also illustrates a magnetic disk drive **451** that reads from or writes to a removable, nonvolatile magnetic disk **452**, and an optical disk drive **455** that reads from or writes to a removable, nonvolatile optical disk **456** such as a CD ROM or other optical media. Other removable/non-removable, volatile/nonvolatile computer storage media that can be used in the exemplary operating environment include, but are not limited to, magnetic tape cassettes, flash memory cards, digital versatile disks, digital video tape, solid state RAM, solid state ROM, and the like. The hard disk drive **441** may be connected to the system bus **421** through a non-removable memory interface such as interface **440**, and magnetic disk drive **451** and optical disk drive **455** may be connected to the system bus **421** by a removable memory interface, such as

interface **450**. However, it should be appreciated that, in some embodiments, some or all of the computer readable media available to a device may be accessed over a communication network.

The drives and their associated computer storage media discussed above and illustrated in FIG. 3, provide storage of computer readable instructions, data structures, program modules and other data for the computer **410**. In FIG. 3, for example, hard disk drive **441** is illustrated as storing operating system **444**, application programs **445**, other program modules **446**, and program data **447**. Note that these components can either be the same as or different from operating system **434**, application programs **435**, other program modules **436**, and program data **437**. Operating system **444**, application programs **445**, other program modules **446**, and program data **447** are given different numbers here to illustrate that, at a minimum, they are different copies.

A computing environment may include one or more input/output devices. Some such input/out devices may provide a user interface. A user may enter commands and information into the computer **410** through input devices such as a keyboard **462** and pointing device **461**, depicted as a mouse. However, other forms of pointing devices may be used, including a trackball, touch pad or touch screen. Other input devices (not shown) may include a microphone, joystick, game pad, satellite dish, scanner, or the like. The microphone, for example, may support voice input, which may be recorded as an audio file or may be translated, such as using speech recognition, to a text format for further processing. These and other input devices are often connected to the processing unit **420** through a user input interface **460** that is coupled to the system bus, but may be connected by other interface and bus structures, such as a parallel port, game port or a universal serial bus (USB).

The computing device may include one or more output devices, including an output device that may form a portion of a user interface. A monitor **491** or other type of display device may also be connected to the system bus **421** via an interface, such as a video interface **490**, to form a visual output device. In addition to the monitor, computers may also include other peripheral output devices such as speakers **497** and printer **496**, which may be connected through an output peripheral interface **495**. The speaker, for example, may enable output via synthesized voice or in any other suitable way.

The computer **410** may operate in a networked environment using logical connections to one or more remote computers, such as a remote computer **480**. The remote computer **480** may be a personal computer, a server, a router, a network PC, a peer device or other common

network node, and typically includes many or all of the elements described above relative to the computer **410**, although only a memory storage device **481** has been illustrated in FIG. 3. The logical connections depicted in FIG. 3 include a local area network (LAN) **471** and a wide area network (WAN) **473**, but may also include other networks. Such networking environments are commonplace in offices, enterprise-wide computer networks, intranets and the Internet. Alternatively or additionally, the WAN may include a cellular network.

When used in a LAN networking environment, the computer **410** is connected to the LAN **471** through a network interface or adapter **470**. When used in a WAN networking environment, the computer **410** typically includes a modem **472** or other means for establishing communications over the WAN **473**, such as the Internet. The modem **472**, which may be internal or external, may be connected to the system bus **421** via the user input interface **460**, or other appropriate mechanism.

In a networked environment, program modules depicted relative to the computer **410**, or portions thereof, may be stored in the remote memory storage device. By way of example, and not limitation, FIG. 3 illustrates remote application programs **485** as residing on memory device **481**. It will be appreciated that the network connections shown are exemplary and other means of establishing a communications link between the computers may be used.

Depending on the nature of the computing device, one or more additional elements may be present. For example, a smart phone or other portable electronic device may include a camera, capable of capturing still or video images. In some embodiments, a computing device may include sensors such as a global positioning system (GPS) to sense location and inertial sensors such as a compass, an inclinometer and/o ran accelerometer. The operating system may include utilities to control these devices to capture data from them and make it available to applications executing on the computing device.

As another example, in some embodiments, a computing device may include a network interface to implement a personal area network. Such an interface may operate in accordance with any suitable technology, including a Bluetooth, Zigbee or an 802.11 ad hoc mode, for example.

Having thus described several aspects of at least one embodiment of this invention, it is to be appreciated that various alterations, modifications, and improvements will readily occur to those skilled in the art.

Such alterations, modifications, and improvements are intended to be part of this disclosure, and are intended to be within the spirit and scope of the invention. Further, though advantages of the present invention are indicated, it should be appreciated that not every embodiment of the invention will include every described advantage. Some embodiments may not implement any features described as advantageous herein. Accordingly, the foregoing description and drawings are by way of example only.

The above-described embodiments of the present invention can be implemented in any of numerous ways. For example, the embodiments may be implemented using hardware, software or a combination thereof. When implemented in software, the software code can be executed on any suitable processor or collection of processors, whether provided in a single computer or distributed among multiple computers. Such processors may be implemented as integrated circuits, with one or more processors in an integrated circuit component, including commercially available integrated circuit components known in the art by names such as CPU chips, GPU chips, microprocessor, microcontroller, or co-processor. Alternatively, a processor may be implemented in custom circuitry, such as an ASIC, or semicustom circuitry resulting from configuring a programmable logic device. As yet a further alternative, a processor may be a portion of a larger circuit or semiconductor device, whether commercially available, semi-custom or custom. As a specific example, some commercially available microprocessors have multiple cores such that one or a subset of those cores may constitute a processor. Though, a processor may be implemented using circuitry in any suitable format.

Further, it should be appreciated that a computer may be embodied in any of a number of forms, such as a rack-mounted computer, a desktop computer, a laptop computer, or a tablet computer. Additionally, a computer may be embedded in a device not generally regarded as a computer but with suitable processing capabilities, including a Personal Digital Assistant (PDA), a smart phone or any other suitable portable or fixed electronic device.

Also, a computer may have one or more input and output devices. These devices can be used, among other things, to present a user interface. Examples of output devices that can be used to provide a user interface include printers or display screens for visual presentation of output and speakers or other sound generating devices for audible presentation of output. Examples of input devices that can be used for a user interface include keyboards, and pointing devices, such as mice, touch pads, and digitizing tablets. As another example, a computer may receive input information through speech recognition or in other audible format. In the

embodiment illustrated, the input/output devices are illustrated as physically separate from the computing device. In some embodiments, however, the input and/or output devices may be physically integrated into the same unit as the processor or other elements of the computing device. For example, a keyboard might be implemented as a soft keyboard on a touch screen. Alternatively, the input/output devices may be entirely disconnected from the computing device, and functionally integrated through a wireless connection.

Such computers may be interconnected by one or more networks in any suitable form, including as a local area network or a wide area network, such as an enterprise network or the Internet. Such networks may be based on any suitable technology and may operate according to any suitable protocol and may include wireless networks, wired networks or fiber optic networks.

Also, the various methods or processes outlined herein may be coded as software that is executable on one or more processors that employ any one of a variety of operating systems or platforms. Additionally, such software may be written using any of a number of suitable programming languages and/or programming or scripting tools, and also may be compiled as executable machine language code or intermediate code that is executed on a framework or virtual machine.

In this respect, the invention may be embodied as a computer readable storage medium (or multiple computer readable media) (*e.g.*, a computer memory, one or more floppy discs, compact discs (CD), optical discs, digital video disks (DVD), magnetic tapes, flash memories, circuit configurations in Field Programmable Gate Arrays or other semiconductor devices, or other tangible computer storage medium) encoded with one or more programs that, when executed on one or more computers or other processors, perform methods that implement the various embodiments of the invention discussed above. As is apparent from the foregoing examples, a computer readable storage medium may retain information for a sufficient time to provide computer-executable instructions in a non-transitory form. Such a computer readable storage medium or media can be transportable, such that the program or programs stored thereon can be loaded onto one or more different computers or other processors to implement various aspects of the present invention as discussed above. As used herein, the term "computer-readable storage medium" encompasses only a computer-readable medium that can be considered to be a manufacture (*i.e.*, article of manufacture) or a machine. Alternatively or

additionally, the invention may be embodied as a computer readable medium other than a computer-readable storage medium, such as a propagating signal.

The terms "program" or "software" are used herein in a generic sense to refer to any type of computer code or set of computer-executable instructions that can be employed to program a computer or other processor to implement various aspects of the present invention as discussed above. Additionally, it should be appreciated that according to one aspect of this embodiment, one or more computer programs that when executed perform methods of the present invention need not reside on a single computer or processor, but may be distributed in a modular fashion amongst a number of different computers or processors to implement various aspects of the present invention.

Computer-executable instructions may be in many forms, such as program modules, executed by one or more computers or other devices. Generally, program modules include routines, programs, objects, components, data structures, *etc.* that perform particular tasks or implement particular abstract data types. Typically the functionality of the program modules may be combined or distributed as desired in various embodiments.

Also, data structures may be stored in computer-readable media in any suitable form. For simplicity of illustration, data structures may be shown to have fields that are related through location in the data structure. Such relationships may likewise be achieved by assigning storage for the fields with locations in a computer-readable medium that conveys relationship between the fields. However, any suitable mechanism may be used to establish a relationship between information in fields of a data structure, including through the use of pointers, tags or other mechanisms that establish relationship between data elements.

Various aspects of the present invention may be used alone, in combination, or in a variety of arrangements not specifically discussed in the embodiments described in the foregoing and is therefore not limited in its application to the details and arrangement of components set forth in the foregoing description or illustrated in the drawings. For example, aspects described in one embodiment may be combined in any manner with aspects described in other embodiments.

Also, the invention may be embodied as a method, of which an example has been provided. The acts performed as part of the method may be ordered in any suitable way. Accordingly, embodiments may be constructed in which acts are performed in an order different

than illustrated, which may include performing some acts simultaneously, even though shown as sequential acts in illustrative embodiments.

Use of ordinal terms such as “first,” “second,” “third,” *etc.*, in the claims to modify a claim element does not by itself connote any priority, precedence, or order of one claim element over another or the temporal order in which acts of a method are performed, but are used merely as labels to distinguish one claim element having a certain name from another element having a same name (but for use of the ordinal term) to distinguish the claim elements.

Also, the phraseology and terminology used herein is for the purpose of description and should not be regarded as limiting. The use of "including," "comprising," or "having," “containing,” “involving,” and variations thereof herein, is meant to encompass the items listed thereafter and equivalents thereof as well as additional items.

The present invention is further illustrated by the following Example, which in no way should be construed as further limiting.

EXAMPLES

Example 1

Introduction

Sjögren's syndrome (SjS) is a chronic systemic human autoimmune disease, yet one characterized primarily by an immune-mediated reduction and destruction of lacrimal, meibomian and salivary gland function resulting, respectively, in dry eye (keratoconjunctivitis sicca / xerophthalmia) and/or dry mouth (stomatitis sicca / xerostomia) diseases (Fox et al., 2008; Fox, 2005; Hansen et al., 2003; Jonsson et al., 2000; Manthorpe et al., 2006). However, in addition to the apparent primary sites of autoimmunity in SjS, multiple tissues can develop pathologies including the lungs, kidneys, GI tract, skin, vasculature, bladder and vagina. Interestingly, as many as 20% of SjS patients exhibit various neuropathies, including sensory, peripheral, cranial and myelopathic complications (Delalande et al., 2004), plus various cognitive impairments such as dementia, lack of concentration, memory loss and various psychiatric disorders (ranging from depression to anxiety). Depression, loss of energy and memory impairment, often noted in patients during clinic visits (Belin et al., 1999; Malinow et al., 1985; Valtysdottir et al., 2000), is referred to as “mental fogginess”, while involvement of

the musculature can lead to fibromyalgia-like symptoms and chronic fatigue (Fox, 2005; Manthorpe et al., 2006). Fatigue is considered the most prevalent complaint and believed to be due to high levels of Interferon (IFN). Increased IFN levels, in turn, activate multiple IFN-responsive/stimulated genes (IRGs/ISGs) involved in innate and adaptive immune activities, defining a specific SjS-associated "IFN signature". Non-limiting examples of IFN-responsive genes and their gene families are provided in Table 1.

Table 1:

Gene Family	Examples of the family genes
IRF	*Irf1, Irf2, Irf3, Irf4, Irf5, Irf6, Irf7, Irf8, Irf9*
ISG	*Isg15, Isg20, Isg20l1*
IFI/IFHI	*Ifi27, Ifi30, Ifi44, Ifi203, Ifi204, Ifi35, Ifi47, Ifi202b, Ifi205, Ifih1*
IFIT	*Ifit1, Ifit3*
IFITM	*Ifitm1, Ifitm5, Ifitm6, Ifitm7, Ifitm2, Ifitm3*
IFN-induced GTPase	*Igtp, Iigp1*
IFR	*Ifrd1, Ifrd2, Ifrg15*
TRIM	*Trim2, Trim6, Trim9, Trim10, Trim11, Trim12, Trim8, Trim14, Trim16, Trim21(Ro), Trim14, Trim15, Trim17, Trim30, Trim31, Trim32, Trim25, Trim28, Trim29, Trim 36, Trim37, Trim42, Trim45, Trim46, Trim54, Trim39, Trim41, Trim44, Trim60, Trim61, Trim63, Trim65, Trim68, Trim69, Trim47, Trim62, Trim7, Trim14, Trim23, Trim26, Trim30, Trim33, Trim34, Trim62, Trim71, Trim75*
SOCS	*Socs1, Socs2, Socs3, Socs4, Socs5, Socs6, Socs7, Cish*
TLR	*Tlr1, Tlr2, Tlr3, Tlr4, Tlr5, Tlr6, Tlr7, Tlr8, Tlr9, Tlr10, Tlr11, Tlr12, Tlr13*

Misc.	*Inf1, Inf2, Il12p35, Il12p40, Cd40, Cd40L*

One fascinating feature of SjS autoimmunity in both humans and animal models of SjS is the reported high levels of interferon (IFN), both IFN-α/β and IFN-γ (Gottenberg et al., 2006; Hjelmervik et al., 2005; Kawakami et al., 2007; Kimoto et al., 2011; Peck et al., 2011; Perez et al., 2009; Spachidou et al., 2007). Although elevated levels of the interferons are often associated with viral infections, there remained little proof to date that SjS is a viral-based disease, despite recent observations that genes encoding TLR3, TLR7, TLR9 and factors in both the TLR- and IFN-signaling pathways are markedly up-regulated prior to disease onset, *i.e.*, the innate immune phase, and apparently independent of detectable adaptive autoimmunity (Devauchelle-Pensec et al., 2010; Obermoser and Pascual, 2010; Peck et al., 2011; Wakamatsu et al., 2007). Furthermore, SjS-susceptible mice expressing non-functional *Ifng* or *IfngR* genes fail to develop any signs of a SjS-like disease (Kimoto et al., 2011; Perez et al., 2009), while mice expressing a non-functional *IfnaR1*gene fail to develop the clinical disease (Cha et al., 2004).

Despite extensive efforts to define the genetic, environmental and/or immunological basis for human SjS, prior to the instant disclosure, the underlying etiology was poorly-defined, due, in part, to the fact that patients are currently diagnosed only after onset of overt clinical disease, sometimes as many as 10 years post-onset.

Cell-autonomous biological processes defined by the SjS IFN-signature

The number of genes being routinely added to the *Interferome* database (Rusinova et al., 2013), together with their arrays of functions, underscores the fact that IRGs/ISGs are not merely activated or suppressed during development and onset of SjS-like disease, but also act as both positive and negative feedback regulatory molecules to ensure maximum host defenses against microbial infections while preventing hyperreactivity leading to unwanted host injury. Furthermore, the IFNs can no longer be viewed as purely anti-viral molecules as IFNs are a central player in innate immunity that is part of the general inflammatory response to injury. Prolonged activation of interferon signaling is critical in dealing with chronic infections, not only for activating an adaptive response, but also for orchestrating cooperative anti-microbial processes between IRGs/ISGs and autophagic factors that opsonize cytosolic pathogens or disrupt compartmentalized pathogens to facilitate efficient killing in autophagolysosomes

(Macmicking, 2012). To this end, unique inducible molecular mechanisms have evolved in mammalian hosts to counter the many schemes used by microorganisms to gain entry into host cells and organs. Thus, by studying the multitude of functions displayed by IRG/ISG family proteins and global transcriptome data, the different IFN-induced cell-autonomous effector biological processes used to kill and/or clear specific pathogens were identified. The first consideration in such an analysis is whether the pathogen is compartmentalized, *e.g.*, in phagocytic vacuoles or pathogen-containing inclusion bodies, or residing freely as a cytosolic pathogen. The second consideration is whether the make-up of an interferon-signature profile at the transcription level can identify, first and foremost, a specific molecular mechanism, then a specific pathogen. As described in the following examples, the transcriptomic analyses of SjS support the concept that the exocrine tissues are mounting an anti-viral host response and not a defensive response against bacteria or parasites.

Identification of a candidate etiological agent for SjS defined by the IFN-signature

Analyses of global temporal transcriptome data collected during development of SjS-like disease in the C57BL/6.NOD-*Aec1Aec2* model of primary SjS defined an unique IFN-signature that could be used to model molecular events and their biological processes underlying SjS. Briefly, differential expression of genes involved in altered biological processes and/or molecular pathways between C57BL/6J and C57BL/6.NOD-*Aec1Aec2* mice at specific time points correlating with pathological processes in developmental and onset stages of disease were investigated (FIG. 4). Genes that were differentially expressed in exocrine glands (*e.g.*, lacrimal and salivary gland tissue) of C57BL/6J and C57BL/6.NOD-*Aec1Aec2* mice between the ages of 4 and 20 weeks (4, 8, 12, 16, and 20 weeks) were identified using Gene Set Enrichment Analysis (GSEA); a total of 21,673 genes was analyzed after data collapsing. Leading Edge (LE) analysis and Network Building were performed to connect Gene Sets (GS) through leading edges.

FIG. 5 depicts data showing biological processes altered (activated/up-regulated) at 8 weeks versus 4 weeks of age in the salivary glands (SG) of C57BL/6.NOD-*Aec1Aec2* mice compared to C57BL/6J controls.

FIG. 6 depicts data showing biological processes down-regulated at 12-16 weeks versus 8 weeks of age in the GS of C57BL/6.NOD-*Aec1Aec2* mice compared to C57BL/6J controls. Many GS up-regulated between 4-8 weeks of age are down-regulated at between 8-12 weeks of age.

FIG. 7 depicts data showing biological processes altered (activated/up-regulated) at 16 weeks of age in the SG of C57BL/6.NOD-*Aec1Aec2* mice compared to C57BL/6J controls.

Gene sets associated with pre-disease to SjS early onset were examined (FIG. 8). Analyses identified two GS clusters of interest: IFN-associated processes with structural changes and protein ubiquitination with cell apoptosis. Specifically, 79 up-regulated GS and 29 down-regulated GS were identified. The up-regulated GS are characterized by four biological processes: innate immunity, Insr>mTor signaling, focal adhesion and motility, and cell-cell junction formation (FIGs. 9-11).

Restricted differential expression of Toll-like receptor genes (*Tlr*) and pathogen recognition receptor (PRR) genes (Tables 2 and 3). With the exception of *Ifih1* (encoding Mda-5), no common PRR genes (*e.g.*, NLRs, RAGE, RHLs (Rig1), LGRs, DAI, CLECs) are up-regulated. Table 4 shows differential expression of IRG in the salivary glands of C57BL/6.NOD-*Aec1Aec2* mice.

Table 2: Relative gene expressions of the TLRs in C57BL/6J and C57BL/6.NOD-Aec1Aec2 mice.

Strain	TLR1	TLR2	TL3	TLR4	TLR5	TLR6	TLR7	TLR8	TLR9	TLR12	TLR13
C57BL/6J	0.65*	0.98	0.91	**1.69**‡	1.02	1.10	**1.46**‡	**1.39**‡	1.10	1.00	**1.22**‡
				p<0.001			p=0.004	p=0.004			p=0.04
C57BL/6.NOD-Aec1Aec2	1.35	0.94	**1.47**	**1.65**	0.98	0.97	0.78	0.98	0.89	0.99	0.96
	p=0.19		p=0.04	p=0.02							

*Differential gene expression in the salivary glands of mice at 8 weeks of age, time of maximal expressions.

‡Bolded numbers are statistically significantly up-regulated by LIMMA pair-wise comparisons.

Table 3: Example of the interferon gene expression profile in the salivary glands of C57BL/6.NOD-Aec1Aec2 mice.

Mouse Line	Ifna1	Ifna2	Ifna4	Ifna5	Ifna6	Ifna7	Ifna9	Ifna11	Ifnab	Ifnb1	Ifng	Ifnz
Aec1Aec2	1.03[a]	0.93	0.99	1.28	0.95	1.01	0.98	0.99	0.82	1.06	1.00	0.99
p value =	(0.64)	(0.11)	(0.62)	**(0.01)**	(0.66)	(0.76)	(0.71)	(0.60)	(0.11)	(0.41)	(0.92)	(0.98)

C57BL/6J	1.03	1.05	1.01	1.01	0.89	**2.39**	1.08	1.50	1.00	1.11	*0.88*	1.38
p value =	(0.86)	(0.29)	(0.74)	(0.42)	(0.22)	**(0.008)**	(0.53)	(0.27)	(0.96)	(0.27)	**(0.08)**	(0.07)

Values represent the mean microarray data of n=5 mice at 16 weeks of age: 16 weeks of age was the only time point with any significant differentially-expressed gene value

Table 4: Example of IFN-responsive genes and their temporal expressions in salivary glands of C57BL/6.NOD-Aec1Aec2 mice

Gene Family	No temporal change	Up-regulated temporal change
IRF	*Irf2, Irf4, Irf5*	*Irf1, Irf3, Irf6, Irf7, Irf8, Irf9*
ISG	Isg15, Isg20	*Isg20l1*
IFI/IFIH	*Ifi27, Ifi30, Ifi44, Ifi203, Ifi204*	*Ifi35, Ifi47, Ifi202b, Ifi205, Ifih1*
IFIT	---	*Ifit1, Ifit3*
IFITM	*Ifitm1, Ifitm5, Ifitm6, Ifitm7*	*Ifitm2, Ifitm3*
IFN-induced GTPase	---	*Igtp, Iigp1*
IFR	*Ifrd1, Ifrd2*	*Ifrg15*
TRIM [a]	*Trim6, Trim9, Trim10, Trim11, Trim12, Trim14,* ***Trim15****, Trim17,* ***Trim19(Pml)****,* ***Trim20(Mefv)****,* ***Trim30(5a)****, Trim31, Trim32, Trim 36, Trim37, Trim42, Trim45, Trim46, Trim54, Trim60, Trim61, Trim63, Trim65,* ***Trim68****, Trim69, Trim71, Trim75*	***Trim8****, Trim14, Trim16,* ***Trim21(Ro)****,* ***Trim25****, Trim28, Trim29, Trim 36, Trim37, Trim39, Trim41, Trim44, Trim47,* ***Trim56****, Trim62, Trim66*
	[a] *Trim13, Trim24,* ***Trim27****, Trim35,* ***Trim40****, Trim59, Trim62*	[b] *Trim2 Trim3, Trim7, Trim14,* ***Trim23****, Trim26, Trim33, Trim34*
SOCS	***Socs1****,* ***Socs3****, Socs6, Socs7*	***Socs2****, Socs4, Socs5,* ***Cish***

[a] Trim genes that are down-regulated

[b] Trim genes that are differentially and temporally up-regulated, but never reaching a $p<0.05$ statistical value

Although there was previously little proof to date that human SjS is a viral-based disease, multiple lines of evidence clearly point to the role of a dsRNA viral etiology in the mouse models described here: (a) an up-regulated expression of *Tlr3* and *Tlr4*, two genes encoding pathogen recognition receptors (PPRs) that signal through Traf3 via Trif and/or through Traf6 via a Trif-Trim23 complex to activate NF-κβ and Irf3/Irf7 transcription of pro-inflammatory cytokines including IFN, (b) the up-regulation of *Ifih1*, encoding Mda-5, with a concomitant down-regulation of *Ddx58*, encoding Rig-1, (c) the up-regulation of the interferon-responsive factors *Irf3*, *Irf7*, *Irf8* and *Irf9* critical for transcription of a vast variety of genes, and (d) the down-regulation of *Trim27*, *Trim30* and *Trim40* with concomitant up-regulation of *Trim8*, *Trim21* (encoding Ro52), *Trim25* and *Trim56*, whose proteins impact viral replication and regulate aspects of innate immunity. While additional genes exist within each of these gene families that also exhibit differential expressions (Jefferies et al., 2011), the genes mentioned point directly to two important concepts: the first questions whether SjS might be a viral-induced autoimmunity, while the second suggests that the cytokine storm exhibited in this disease is under the direction of regulatory Trim molecules.

With respect to the first point, the three activated pathogen-recognition receptors (PRRs) in this model (Tlr3, Tlr4 and Mda-5) are receptors involved in the recognition of dsRNA viruses. No other PRR (or class of PRRs) have been found to be activated in this model, including Nod, Nalp, Ipaf, Naip, Rage, Rxfp1, and Dai receptors. Of particular interest, however, is the fact that *Mda5* (*Ifih1*), but not *Rig1* (*Ddx58*), is up-regulated coordinately with *Tlr3*. Rig-1 tends to recognize viruses of the *Paramyxoviridae* family (*e.g.*, mumps, measles, respiratory syncytial and parainfluenza viruses), while Mda-5 tends to recognize viruses of the *Picornaviridae* family (*e.g.*, coxsackie, encephalomyocarditis and rhinoviruses) or *Reoviridae* family (*e.g.*, rotovirus).

The second point, that Trim molecules may be directing both the molecular mechanisms underlying the cytokine storm observed in SjS patients and the transition from an enhanced innate response to an adaptive autoimmune response, is supported by the *Trim* gene expression profile present in the exocrine glands (Jefferies et al., 2011). In essence, the three Trim molecules (Trim27, Trim30 and Trim40), whose gene expressions are down-regulated, function to suppress the signal transductions of the Tlr4, Tlr3 and Mda5 signaling pathways at various

signaling points. In contrast, the genes encoding Trim21, Trim23, Trim25 and Trim56, four molecules whose functions are to up-regulate the Tlr3, Tlr4 and Mda5 pathways at different signaling steps, are each up-regulated. In addition, the gene encoding *Trim8*, whose function is to suppress the action of the Socs (Suppressor of cytokine synthesis) molecules (Toniato et al., 2002) is strongly up-regulated. A graphic depiction of Socs vs. Trim regulation of innate and adaptive immunity is shown in FIG. 1. Additionally, FIG. 12 provides a schematic depiction of the interferon signature of TLRs and Trims in SjS.

Taken as a whole, this interferon signature indicates up-regulation of pathways leading to strong transcription of pro-inflammatory cytokines, interferons and molecules known to activate adaptive responses (*e.g.*, IL6, IL12p40, Rantes, CD40, CD80 and CD56). For example, FIG. 13 shows interaction of the IFN-pathway with TAM receptor tyrosine kinase (TAM-RTK)-encoding genes, which down regulate inflammation (*e.g.*, IFN cytokine activity) by blocking activation of Traf molecules by TLR complexes. Specifically, expression of TAM-RTK genes (Gas6, Mertk, and ADAM10) is upregulated in C57BL/6.NOD-*Aec1Aec2* mice compared to C57BL/6J controls, confirming the lack of a down-regulated innate immune response in SjS.

In another example, the TGFβ response was investigated. Znf512β (also referred to as ZnF-36 and ZFP-512B) regulates microRNA cluster 17-92 and is known to down-regulate the TGFβ canonical pathway and increase apoptosis (FIG. 14). Differential expression of the zinc finger protein 512β (Znf512β)-Myc-TGFβ gene axis in lacrimal glands of B6.NOD-*Aec1Aec2* mice was examined. Data indicates that both Znf512β and TGFβ, as well as p21, Mapk14, Ctgr, Tsp1, Smad4, Pten, Bcl2l11 and Stat3, are upregulated in B6.NOD-*Aec1Aec2* mice (FIG. 15). Therefore, interferon signature analysis identifies pathways that lead down to the transcriptional level of a disease.

Comparison between mouse and human SjS-associated IFN-signatures

Genes used to establish the SjS interferon signatures in both mouse and humans overlap, even though the specific underlying etiologic agents are now suspected to be species-specific dsRNAs. IRGs/ISGs that have thus far been reported as differentially-expressed in human SjS patients by several groups (Devauchelle-Pensec et al., 2010; Emamian et al., 2009; Gottenberg et al., 2006; Hjelmervik et al., 2005; Kimoto et al., 2011; Perez et al., 2009; Wakamatsu et al., 2007) include *IRF7, MX1, GIP2, GIP3, OAS1, OAS2, PKR, IFI16, IFI27, IFI30, IFI35, IFI44, ISG20, ISG56K, IFIT1, IFIT2, IFIT4, IFITM1, IFITM3, IP10/CXCL10, APOBEC3,*

SAMHD1,TETHERIN, VIPERIN and *STAT1a*, the majority of which are also represented in the differentially-expressed genes up-regulated in the exocrine glands of the C57BL/6.NOD-*Aec1Aec2* mice (Peck and Nguyen, 2012). This overlapping set of differentially-expressed genes must be considered an important subset of responsive genes that point to specific and shared etiopathological processes.

While it is natural to focus on the many similarities in the IFN-associated gene sets differentially-expressed in human SjS patients and SjS-susceptible C57BL/6.NOD-*Aec1Aec2* mice, there are also important differences. A few that stand out include expression profiles for *Irf8*, *Ifi202b - Ifi205* encoding the p200 family molecules, and the three interferon-inducible genes, *Ifi27*, *Ifi30* and *Ifi44*. For example, the *Irf8* gene, which encodes a factor that is involved in myeloid differentiation and Fas-mediated apoptosis, as well as B cell development and transcriptional regulation of germinal center formation (Wang and Morse, 2009), deserves attention due its highly up-regulated expression in the C57BL/6.NOD-*Aec1Aec2* mice. Earlier studies postulated that myeloid cells enter the exocrine glands during the early innate response (8-12 weeks of age) in response to Fas-FasL-mediated apoptosis of acinar tissue, while B cells enter the salivary glands transiently during the adaptive immune phase (post-16 weeks of age). Interestingly, the temporal expression profile of *Irf8* showed a bimodal profile, in line with this hypothesis. The inability to detect an up-regulated expression of *IRF8* in SjS patients is an interesting aspect to examine further, as binding of the transcriptional factor PU.1 to Irf8 leads to up-regulation of OAS1 and/or OAS2, two molecules that can bind and degrade dsRNA viral RNA (Rogozin et al., 2003), and are highly up-regulated in SjS patients. In contrast, the p200 molecules, encoded by the *Ifi200* family of genes, are known to sense cytoplasmic DNA, leading to the formation and activation of inflammasomes with subsequent production of anti-nuclear antibodies (Choubey et al., 2010). Although there was an *Ifi202b* up-regulated gene expression in the exocrine glands of C57BL/6.NOD-*Aec1Aec2* mice, there is no evidence for activation of inflammasomes in these mice, in contrast to their comparative SjS-non-susceptible C57BL/6J partners. Lastly, whereas *IFI27*, *IFI30* and *IFI44* have been consistently found to be up-regulated in SjS patients (Devauchelle-Pensec et al., 2010; Emamian et al., 2009; Hjelmervik et al., 2005; Kimoto et al., 2011), these three *Ifi* genes with distinct functions were not found to be differentially-expressed in the exocrine glands of C57BL/6.NOD-*Aec1Aec2* mice. Considering IFI44 is associated with HCV and RSV infections, this difference between humans and mice lies in the fact that the underlying etiological agent(s) of SjS in these two species is different and

invokes different environmental triggers. Interestingly, IFI44L was identified as a marker gene in RA (Raterman et al., 2012). These data support the concept that an individual gene or gene product is insufficient to accurately identify the underlying molecular mechanisms that permit identification of a specific disease. Thus, differentially-expressed gene sets common to both species indicate activation of similar pathways of immunopathological processes, highlighting the fact that both similarities and differences in the IFN-signatures are critical to understanding the small nuances of molecular events associated with disease.

REFERENCES

Ansell, P., Simpson, J., Lightfoot, T., Smith, A., Kane, E., Howell, D., Newton, R., McGonagle, D., Jack, A., and Roman, E. (2011). Non-Hodgkin lymphoma and autoimmunity: does gender matter? Int J Cancer 129, 460-466.

Belin, C., Moroni, C., Caillat-Vigneron, N., Debray, M., Baudin, M., Dumas, J.L., Moretti, J.L., Delaporte, P., and Guillevin, L. (1999). Central nervous system involvement in Sjogren's syndrome: evidence from neuropsychological testing and HMPAO-SPECT. Ann Med Interne (Paris) 150, 598-604.

Bemark, M., Holmqvist, J., Abrahamsson, J., and Mellgren, K. (2012). Translational Mini-Review Series on B cell subsets in disease. Reconstitution after haematopoietic stem cell transplantation - revelation of B cell developmental pathways and lineage phenotypes. Clin Exp Immunol 167, 15-25.

Cha, S., Brayer, J., Gao, J., Brown, V., Killedar, S., Yasunari, U., and Peck, A.B. (2004). A dual role for interferon-gamma in the pathogenesis of Sjogren's syndrome-like autoimmune exocrinopathy in the nonobese diabetic mouse. Scand. J. Immunol. 60, 552-565.

Cha, S., Nagashima, H., Brown, V.B., Peck, A.B., and Humphreys-Beher, M.G. (2002a). Two NOD Idd-associated intervals contribute synergistically to the development of autoimmune exocrinopathy (Sjogren's syndrome) on a healthy murine background. Arthritis. Rheum. 46, 1390-1398.

Cha, S., Peck, A.B., and Humphreys-Beher, M.G. (2002b). Progress in understanding autoimmune exocrinopathy using the non-obese diabetic mouse: an update. Crit. Rev. Oral. Biol. Med. 13, 5-16.

Cha, S., van Blockland, S.C., Versnel, M.A., Homo-Delarche, F., Nagashima, H., Brayer, J., Peck, A.B., and Humphreys-Beher, M.G. (2001). Abnormal organogenesis in salivary gland

development may initiate adult onset of autoimmune exocrinopathy. Exp Clin Immunogenet 18, 143-160.

Choubey, D., Duan, X., Dickerson, E., Ponomareva, L., Panchanathan, R., Shen, H., and Srivastava, R. (2010). Interferon-inducible p200-family proteins as novel sensors of cytoplasmic DNA: role in inflammation and autoimmunity. Journal of interferon & cytokine research : the official journal of the International Society for Interferon and Cytokine Research 30, 371-380.

De Vita, S., Boiocchi, M., Sorrentino, D., Carbone, A., Avellini, C., Dolcetti, R., Marzotto, A., Gloghini, A., Bartoli, E., Beltrami, C.A., and Ferraccioli, G. (1997). Characterization of prelymphomatous stages of B cell lymphoproliferation in Sjogren's syndrome. Arthritis. Rheum. 40, 318-331.

Delalande, S., de Seze, J., Fauchais, A.L., Hachulla, E., Stojkovic, T., Ferriby, D., Dubucquoi, S., Pruvo, J.P., Vermersch, P., and Hatron, P.Y. (2004). Neurologic manifestations in primary Sjogren syndrome: a study of 82 patients. Medicine (Baltimore) 83, 280-291.

Delaleu, N., Nguyen, C.Q., Peck, A.B., and Jonsson, R. (2011). Sjogren's syndrome: studying the disease in mice. Arthritis research & therapy 13, 217.

Devauchelle-Pensec, V., Cagnard, N., Pers, J.O., Youinou, P., Saraux, A., and Chiocchia, G. (2010). Gene expression profile in the salivary glands of primary Sjogren's syndrome patients before and after treatment with rituximab. Arthritis and rheumatism 62, 2262-2271.

Dimitriou, I.D., Kapsogeorgou, E.K., Moutsopoulos, H.M., and Manoussakis, M.N. (2002). CD40 on salivary gland epithelial cells: high constitutive expression by cultured cells from Sjogren's syndrome patients indicating their intrinsic activation. Clin Exp Immunol 127, 386-392.

Emamian, E.S., Leon, J.M., Lessard, C.J., Grandits, M., Baechler, E.C., Gaffney, P.M., Segal, B., Rhodus, N.L., and Moser, K.L. (2009). Peripheral blood gene expression profiling in Sjogren's syndrome. Genes Immun 10, 285-296.

Fox, P.C., Bowman, S.J., Segal, B., Vivino, F.B., Murukutla, N., Choueiri, K., Ogale, S., and McLean, L. (2008). Oral involvement in primary Sjogren syndrome. J Am Dent Assoc 139, 1592-1601.

Fox, R.I. (2005). Sjogren's syndrome. Lancet 366, 321-331.

Gottenberg, J.E., Cagnard, N., Lucchesi, C., Letourneur, F., Mistou, S., Lazure, T., Jacques, S., Ba, N., Ittah, M., Lepajolec, C., et al. (2006). Activation of IFN pathways and plasmacytoid

dendritic cell recruitment in target organs of primary Sjogren's syndrome. Proceedings of the National Academy of Sciences of the United States of America 103, 2770-2775.

Hansen, A., Lipsky, P.E., and Dorner, T. (2003). New concepts in the pathogenesis of Sjogren syndrome: many questions, fewer answers. Curr. Opin. Rheumatol. 15, 563-570.

Hjelmervik, T.O., Petersen, K., Jonassen, I., Jonsson, R., and Bolstad, A.I. (2005). Gene expression profiling of minor salivary glands clearly distinguishes primary Sjogren's syndrome patients from healthy control subjects. Arthritis. Rheum. 52, 1534-1544.

Humphreys-Beher, M.G., Hu, Y., Nakagawa, Y., Wang, P.L., and Purushotham, K.R. (1994). Utilization of the non-obese diabetic (NOD) mouse as an animal model for the study of secondary Sjogren's syndrome. Adv. Exp. Med. Biol. 350, 631-636.

Isaacson, P.G., and Du, M.Q. (2004). MALT lymphoma: from morphology to molecules. Nat. Rev. Cancer. 4, 644-653.

Jefferies, C., Wynne, C., and Higgs, R. (2011). Antiviral TRIMs: friend or foe in autoimmune and autoinflammatory disease? Nature reviews. Immunology 11, 617-625.

Jonsson, R., Haga, H.J., and Gordon, T.P. (2000). Current concepts on diagnosis, autoantibodies and therapy in Sjogren's syndrome. Scand. J. Rheumatol. 29, 341-348.

Kawakami, A., Nakashima, K., Tamai, M., Nakamura, H., Iwanaga, N., Fujikawa, K., Aramaki, T., Arima, K., Iwamoto, N., Ichinose, K., et al. (2007). Toll-like receptor in salivary glands from patients with Sjogren's syndrome: functional analysis by human salivary gland cell line. The Journal of rheumatology 34, 1019-1026.

Killedar, S.Y., Eckenrode, S.E., McIndoe, R.A., She, J.X., Nguyen, C.Q., Peck, A.B., and Cha, S.R. (2006). Early pathogenic events associated with Sjogren's syndrome (SjS)-like disease of the nod mouse using microarray analysis. Lab Invest 86, 1243-1260.

Kimoto, O., Sawada, J., Shimoyama, K., Suzuki, D., Nakamura, S., Hayashi, H., and Ogawa, N. (2011). Activation of the interferon pathway in peripheral blood of patients with Sjogren's syndrome. The Journal of rheumatology 38, 310-316.

Macmicking, J.D. (2012). Interferon-inducible effector mechanisms in cell-autonomous immunity. Nature reviews. Immunology 12, 367-382.

Malinow, K.L., Molina, R., Gordon, B., Selnes, O.A., Provost, T.T., and Alexander, E.L. (1985). Neuropsychiatric dysfunction in primary Sjogren's syndrome. Ann Intern Med 103, 344-350.

Manthorpe, R., Bredberg, A., Henriksson, G., and Larsson, A. (2006). Progress and regression within primary Sjogren's syndrome. Scand. J. Rheumatol. 35, 1-6.

Masaki, Y., and Sugai, S. (2004). Lymphoproliferative disorders in Sjogren's syndrome. Autoimmun. Rev. 3, 175-182.

Nguyen, C.Q., Cha, S.R., and Peck, A.B. (2007). Sjögren's syndrome (SjS)-like disease of mice: the importance of B lymphocytes and autoantibodies. Frontiers in Bioscience 12, 1767-1789.

Nguyen, C.Q., and Peck, A.B. (2009). Unraveling the pathophysiology of Sjogren syndrome-associated dry eye disease. Ocul Surf 7, 11-27.

Niiro, H., and Clark, E.A. (2002). Regulation of B-cell fate by antigen-receptor signals. Nat. Rev. Immunol. 2, 945-956.

Nordmark, G., Eloranta, M.L., and Ronnblom, L. (2012). Primary Sjogren's syndrome and the type I interferon system. Curr Pharm Biotechnol.

Obermoser, G., and Pascual, V. (2010). The interferon-alpha signature of systemic lupus erythematosus. Lupus 19, 1012-1019.

Ohlsson, M., Szodoray, P., Loro, L.L., Johannessen, A.C., and Jonsson, R. (2002). CD40, CD154, Bax and Bcl-2 expression in Sjogren's syndrome salivary glands: a putative anti-apoptotic role during its effector phases. Scand. J. Immunol. 56, 561-571.

Peck, A.B., and Nguyen, C.Q. (2012). Transcriptome analysis of the interferon-signature defining the autoimmune process of Sjogren's syndrome. Scand J Immunol 76, 237-245.

Peck, A.B., Nguyen, C.Q., Sharma, A., McIndoe, R.A., and She, J.X. (2011). The interferon-signature of Sjögren's syndrome: What does it say about the etiopathology of autoimmunity. J Clinical Rheum & Musculoskeletal Med 1, 1-17.

Perez, P., Anaya, J.M., Aguilera, S., Urzua, U., Munroe, D., Molina, C., Hermoso, M.A., Cherry, J.M., Alliende, C., Olea, N., et al. (2009). Gene expression and chromosomal location for susceptibility to Sjogren's syndrome. Journal of autoimmunity 33, 99-108.

Poe, J.C., Hasegawa, M., and Tedder, T.F. (2001). CD19, CD21, and CD22: multifaceted response regulators of B lymphocyte signal transduction. Int. Rev. Immunol. 20, 739-762.

Pourmand, N., Wahren-Herlenius, M., Gunnarsson, I., Svenungsson, E., Lofstrom, B., Ioannou, Y., Isenberg, D.A., and Magnusson, C.G. (1999). Ro/SSA and La/SSB specific IgA autoantibodies in serum of patients with Sjogren's syndrome and systemic lupus erythematosus. Ann. Rheum. Dis. 58, 623-629.

Raterman, H.G., Vosslamber, S., de Ridder, S., Nurmohamed, M.T., Lems, W.F., Boers, M., van de Wiel, M., Dijkmans, B.A., Verweij, C.L., and Voskuyl, A.E. (2012). Interferon type I

signature may predict non response upon rituximab in rheumatoid arthritis patients. Arthritis research & therapy 14, R95.

Rogozin, I.B., Aravind, L., and Koonin, E.V. (2003). Differential action of natural selection on the N and C-terminal domains of 2'-5' oligoadenylate synthetases and the potential nuclease function of the C-terminal domain. Journal of molecular biology 326, 1449-1461.

Rusinova, I., Forster, S., Yu, S., Kannan, A., Masse, M., Cumming, H., Chapman, R., and Hertzog, P.J. (2013). Interferome v2.0: an updated database of annotated interferon-regulated genes. Nucleic Acids Res 41, D1040-1046.

Sawalha, A.H., and Harley, J.B. (2004). Antinuclear autoantibodies in systemic lupus erythematosus. Curr Opin Rheumatol 16, 534-540.

Shiboski, S.C., Shiboski, C.H., Criswell, L., Baer, A., Challacombe, S., Lanfranchi, H., Schiodt, M., Umehara, H., Vivino, F., Zhao, Y., et al. (2012). American College of Rheumatology classification criteria for Sjogren's syndrome: a data-driven, expert consensus approach in the Sjogren's International Collaborative Clinical Alliance cohort. Arthritis Care Res (Hoboken) 64, 475-487.

Spachidou, M.P., Bourazopoulou, E., Maratheftis, C.I., Kapsogeorgou, E.K., Moutsopoulos, H.M., Tzioufas, A.G., and Manoussakis, M.N. (2007). Expression of functional Toll-like receptors by salivary gland epithelial cells: increased mRNA expression in cells derived from patients with primary Sjogren's syndrome. Clin Exp Immunol 147, 497-503.

Staeheli, P., and Sutcliffe, J.G. (1988). Identification of a second interferon-regulated murine Mx gene. Mol Cell Biol 8, 4524-4528.

Toniato, E., Chen, X.P., Losman, J., Flati, V., Donahue, L., and Rothman, P. (2002). TRIM8/GERP RING finger protein interacts with SOCS-1. The Journal of biological chemistry 277, 37315-37322.

Valtysdottir, S.T., Gudbjornsson, B., Lindqvist, U., Hallgren, R., and Hetta, J. (2000). Anxiety and depression in patients with primary Sjogren's syndrome. J Rheumatol 27, 165-169.

Vitali, C., Bombardieri, S., Jonsson, R., Moutsopoulos, H.M., Alexander, E.L., Carsons, S.E., Daniels, T.E., Fox, P.C., Fox, R.I., Kassan, S.S., et al. (2002). Classification criteria for Sjogren's syndrome: a revised version of the European criteria proposed by the American-European Consensus Group. Ann. Rheum. Dis. 61, 554-558.

Voulgarelis, M., and Moutsopoulos, H.M. (2003). Lymphoproliferation in autoimmunity and Sjogren's syndrome. Curr. Rheumatol. Rep. 5, 317-323.

Wakamatsu, E., Nakamura, Y., Matsumoto, I., Goto, D., Ito, S., Tsutsumi, A., and Sumida, T. (2007). DNA microarray analysis of labial salivary glands of patients with Sjogren's syndrome. Annals of the rheumatic diseases 66, 844-845.

Wang, H., Lee, C.H., Qi, C., Tailor, P., Feng, J., Abbasi, S., Atsumi, T., and Morse, H.C., 3rd (2008). IRF8 regulates B-cell lineage specification, commitment, and differentiation. Blood 112, 4028-4038.

Wang, H., and Morse, H.C., 3rd (2009). IRF8 regulates myeloid and B lymphoid lineage diversification. Immunol Res 43, 109-117.

Yang, K., Shi, H.X., Liu, X.Y., Shan, Y.F., Wei, B., Chen, S., and Wang, C. (2009). TRIM21 is essential to sustain IFN regulatory factor 3 activation during antiviral response. Journal of immunology 182, 3782-3792.

Yao, Y., Liu, Z., Jallal, B., Shen, N., and Ronnblom, L. (2012). Type I interferons in Sjogren's syndrome. Autoimmun Rev.

CLAIMS

What is claimed is:

1. A diagnostic array, comprising

a substrate with a surface comprising a plurality of locations, wherein:

each of the plurality of locations comprises at least one oligonucleotide attached to the substrate at a predetermined location,

each location of the plurality of locations comprises at least one material targeting an interferon-regulated gene (IRG), and

the material at the plurality of locations target at least 100 interferon-regulated genes.

2. The substrate of claim 1, wherein the surface comprises between 1 material and between 5 materials that target a non-IRG.

3. The substrate of claim 1, wherein the surface does not comprise material targeting non-IRGs.

4. The substrate of any one of claims 1 to 3, wherein the at least one material is an oligonucleotide.

5. The substrate of any one of claims 1 to 4, wherein the surface comprises at least 200, at least 300, at least 400, at least 500, at least 750, at least 1000, or at least 1200 oligonucleotides, wherein each oligonucleotide targets a different interferon-responsive gene (IRG).

6. The substrate of any one of claims 1 to 5, wherein the IRGs are type-I interferon responsive genes.

7. The substrate of any one of claims 1 to 6, wherein at least one oligonucleotide targets a gene selected from the Interferome Database.

8. The substrate of any one of claims 1 to 7, wherein each of the at least one oligonucleotides is the same length.

9. The substrate of any one of claims 1 to 8, wherein at least two oligonucleotides are not the same length.

10. The substrate of any one of claims 1 to 9, wherein the at least one oligonucleotide is a plurality of oligonucleotides, wherein each oligonucleotide is the same length.

11. The substrate of any one of claims 1 to 10, wherein each location occupies a space of less than 1 cm^2 on the substrate.

12. The substrate of claim 11, wherein each location occupies a space of less than 1 μm^2 on the substrate.

13. The substrate of any one of claims 1 to 12, wherein the substrate is glass.

14. The substrate of any one of claims 1 to 12, wherein the substrate is plastic.

15. The substrate of any one of claims 1 to 14, for use in identifying interferon signature in a sample obtained from a subject having or suspected of having a disease.

16. A method for identifying an interferon signature, the method comprising:

i. contacting the substrate of any one of claims 1 to 14 with a sample comprising nucleic acids obtained from at least one cell (*e.g.*, sample);

ii. determining the gene expression profile of the sample by detecting hybridization of the nucleic acids to the substrate, wherein the gene expression profile identifies the expression level of a plurality of interferon-related genes (IRGs) relative to the expression level of a plurality of IRGs in a healthy control cell; and,

iii. assigning an interferon signature to the sample based upon the gene expression profile of the sample.

17. The method of claim 16, wherein the cell (*e.g.*, sample) is obtained from a subject having or suspected of having a pathological condition.

18. The method of claim 17, wherein the pathological condition is an interferon-related disease.

19. The method of any one of claims 16 to 18, wherein the IRGs are type-I IRGs.

20. The method of any one of claims 16 to 19, wherein the cell is derived from blood, or a tissue biopsy.

21. The method of any one of claims 16 to 20, wherein the nucleic acid is DNA.

22. The method of any one of claims 16 to 21, wherein the plurality of IRGs comprises at least 200, at least 300, at least 400, at least 500, at least 750, at least 1000, or at least 1500 IRGs.

23. The method of any one of claims 16 to 22, wherein the interferon signature identifies IRGs having upregulated expression levels relative to healthy control cells.

24. The method of any one of claims 16 to 23, wherein the interferon signature identifies IRGs having down regulated expression levels relative to healthy control cells.

25. The method of any one of claims 16 to 24, wherein the method further comprises comparing the interferon signature to a database, wherein the database comprises known interferon signatures associated with molecular events.

26. The method of claim 25, wherein the method further comprises assigning a molecular event to the interferon signature of (iii) if the interferon signature of (iii) matches an interferon signature in the database.

27. A method for determining a personalized therapeutic regimen for a subject, the method comprising:

i. obtaining a sample from a subject;

ii. determining an interferon signature of the sample;

iii. with at least one processor, comparing the interferon signature of the sample to a database of known interferon signatures, wherein each interferon signature of the database is associated with a molecular event;

iv. assigning the interferon signature of the sample to a molecular event, wherein the assignment is made by matching the interferon signature of the sample to an interferon signature of the database ; and,

v. prescribing a therapeutic regimen based upon the molecular event assigned to the interferon signature of the sample.

28. The method of claim 27, wherein the interferon signature of the sample is determined by the method of any one of claims 16 to 24.

29. The method of claim 27, wherein the interferon signature of the sample is determined by identifying a gene expression profile, wherein the gene expression profile is identified by detecting hybridization of nucleic acids in the sample to a plurality of IRG-detecting materials in a solution.

30. The method of claim 29, wherein the IRG-detecting materials are oligonucleotide probes.

31. The method of claim 29 or 30, wherein each IRG-detecting material specifically hybridizes to a type-I IRG.

32. The method of claim 27, wherein detecting the interferon signature is performed by detecting a gene expression profile by nucleic acid sequencing or protein sequencing.

33. The method of claim 32, wherein the nucleic acid sequencing is DNA sequencing or RNA sequencing.

34. The method of claim 32 or 33, wherein detecting the gene expression profile further comprises analyzing the sequencing data by gene set enrichment analysis (GSEA).

35. The method of any one of claims 27 to 34, wherein the database comprises only type-I interferon signatures.

36. The method of any one of claims 27 to 35, wherein the database comprises at least 200, at least 500, at least 1000, or at least 1500 interferon signatures.

37. The method of any one of claims 27 to 36, wherein the database comprises at least 10, at least 20, at least 50, at least 100, at least 200, at least 500, at least 1000, at least 5000, at least 10000, at least 20000, or up to about 30000 molecular events.

38. The method of any one of claims 27 to 37, wherein the therapeutic regimen comprises administering a molecule that modulates the molecular event assigned to the interferon signature of the sample to the subject.

39. The method of any one of claims 27 to 38, wherein the subject is a human.

40. A system for identifying an interferon signature in a sample, comprising:

(i) a substrate from any one of claims 1 to 14;

(ii) a microarray scanner;

(iii) a computer; and

(iv) a database containing interferon signatures and associated diseases.

41. The system of claim 40, further comprising a second database containing a plurality of diseases and associated therapeutic regimens for each disease.

42. The system of claim 40 or 41, comprising a third database containing interferon signatures and associated molecular events.

43. The system of claim 41 or 42, wherein the database, second database and/or third database are electronically connected.

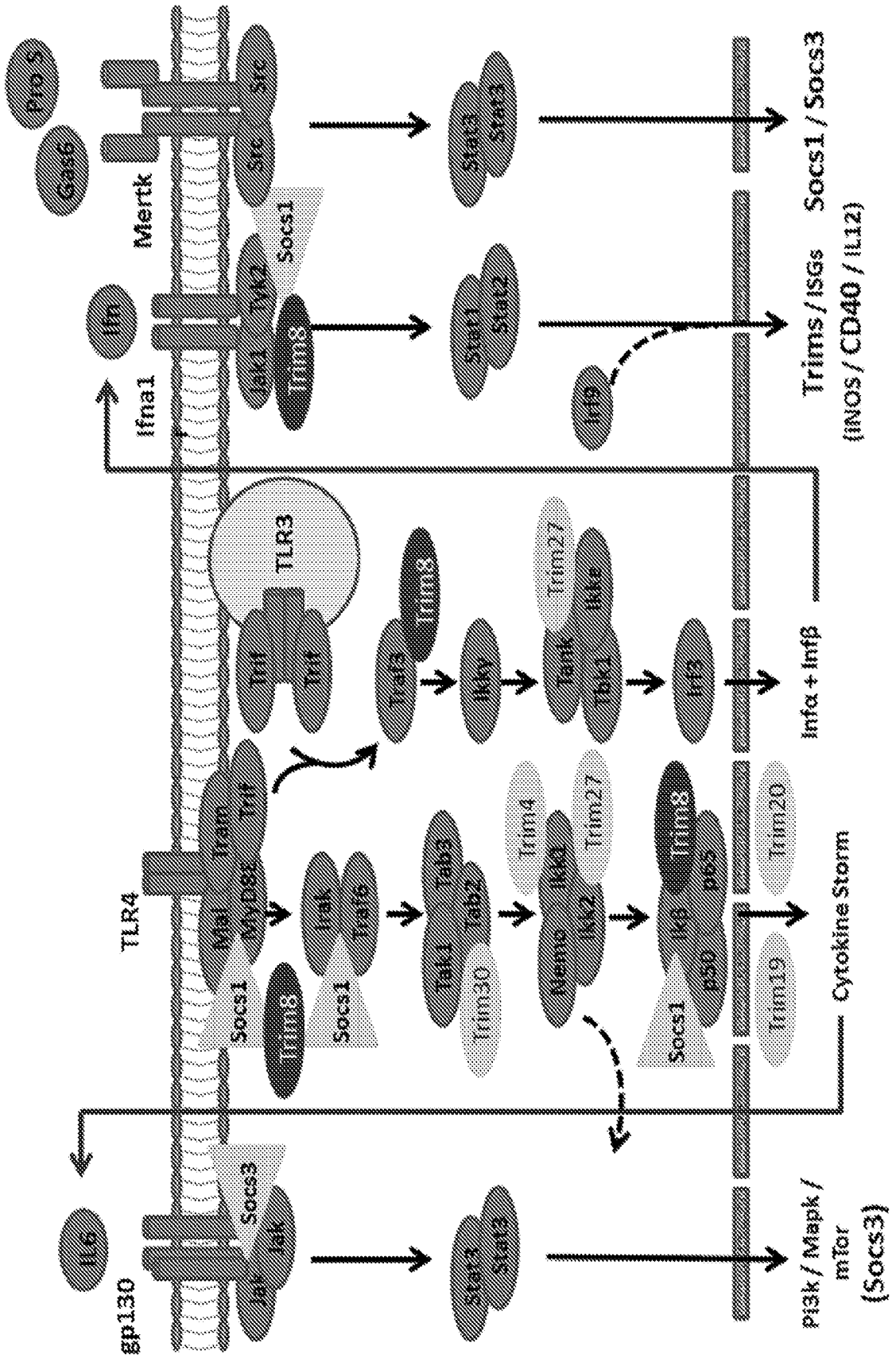
Pro S
Gas6
Mertk
Src
Src
Ifn
Ifna1
Jak1
Tyk2
Socs1
Trim8
Stat3
Stat3
Stat1
Stat2
Irf9
Trims / ISGs
(iNOS / CD40 / IL12)
Socs1 / Socs3
TLR3
Trif
Trif
Traf3
Trim8
Ikky
Trim27
Tank
Ikke
Tbk1
Irf3
Infα + Infβ
TLR4
Tram
Trif
Mal
MyD88
Irak
Tarfe
Socs1
Trim8
Socs1
Tak1
Tab3
Tab2
Trim30
Trim4
Nemo
Ikk1
Ikk2
Trim27
Trim8
Ikb
p65
p50
Socs1
Trim20
Trim19
Cytokine Storm
Il6
gp130
Socs3
Jak
Jak
Stat3
Stat3
Pi3k / Mapk /
mTor
(Socs3)

FIG. 1

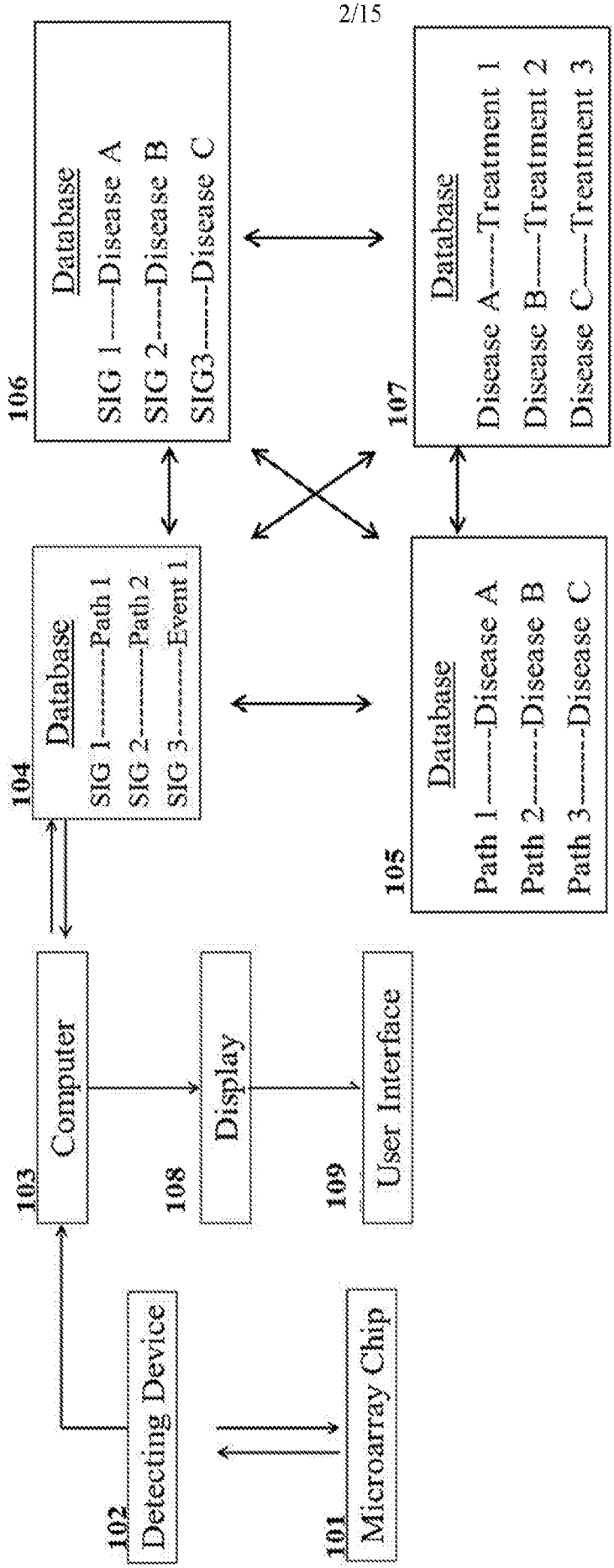
Database
SIG 1----Disease A
SIG 2----Disease B
SIG3----Disease C
106
Database
Disease A----Treatment 1
Disease B----Treatment 2
Disease C----Treatment 3
107
Database
SIG 1--------Path 1
SIG 2--------Path 2
SIG 3--------Event 1
104
Database
Path 1--------Disease A
Path 2--------Disease B
Path 3--------Disease C
105
Computer
103
Display
108
User Interface
109
Detecting Device
102
Microarray Chip
101

FIG. 2

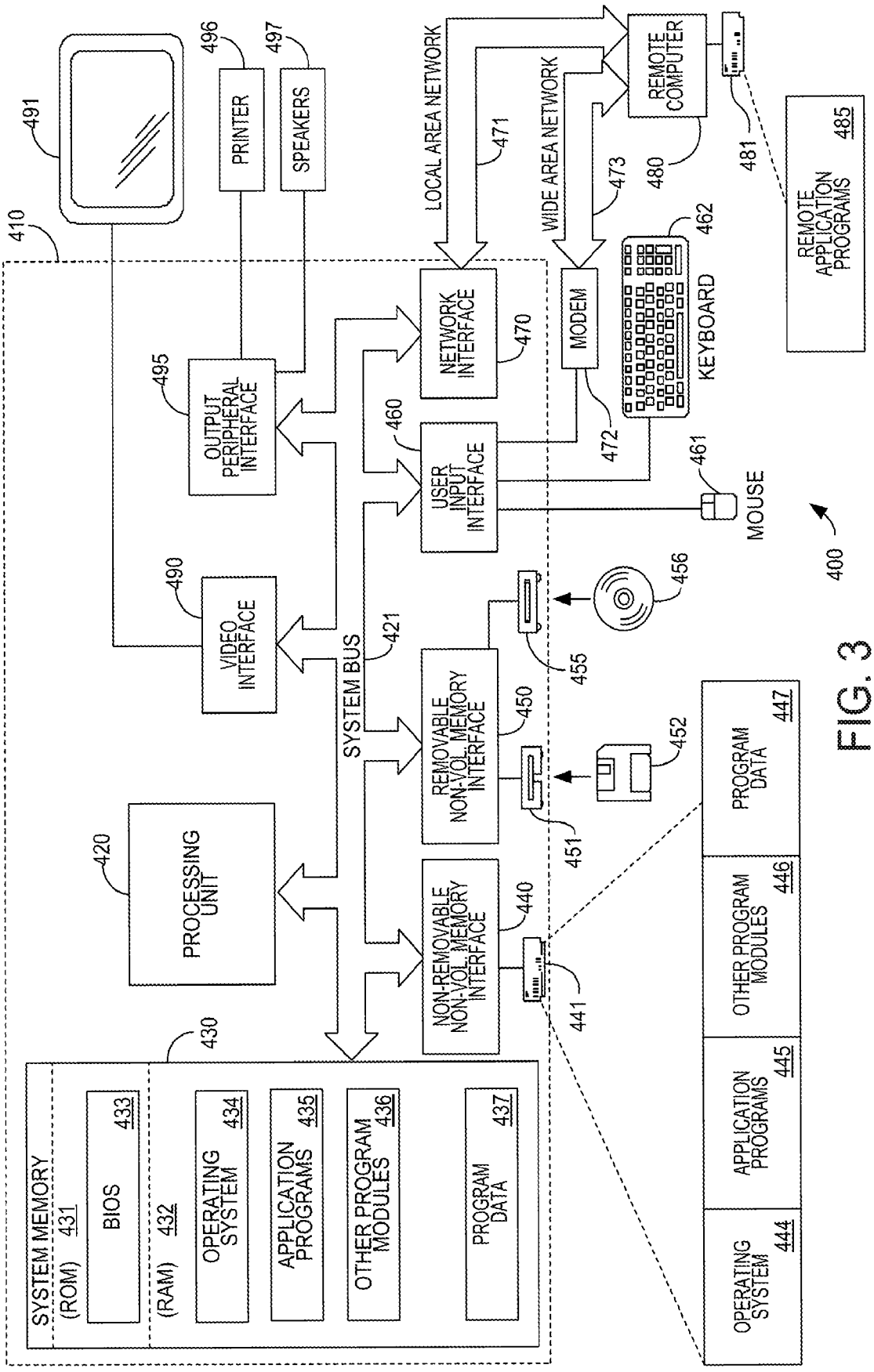
400
410
420
421
430
431
432
433
434
435
436
437
440
441
444
445
446
447
450
451
452
455
456
460
461
462
470
471
472
473
480
481
485
490
491
495
496
497
BIOS
OPERATING SYSTEM
APPLICATION PROGRAMS
OTHER PROGRAM MODULES
PROGRAM DATA
SYSTEM MEMORY
(ROM)
(RAM)
PROCESSING UNIT
SYSTEM BUS
VIDEO INTERFACE
OUTPUT PERIPHERAL INTERFACE
PRINTER
SPEAKERS
NETWORK INTERFACE
LOCAL AREA NETWORK
MODEM
WIDE AREA NETWORK
REMOTE COMPUTER
REMOTE APPLICATION PROGRAMS
USER INPUT INTERFACE
KEYBOARD
MOUSE
REMOVABLE NON-VOL. MEMORY INTERFACE
NON-REMOVABLE NON-VOL. MEMORY INTERFACE
OPERATING SYSTEM
APPLICATION PROGRAMS
OTHER PROGRAM MODULES
PROGRAM DATA

FIG. 3

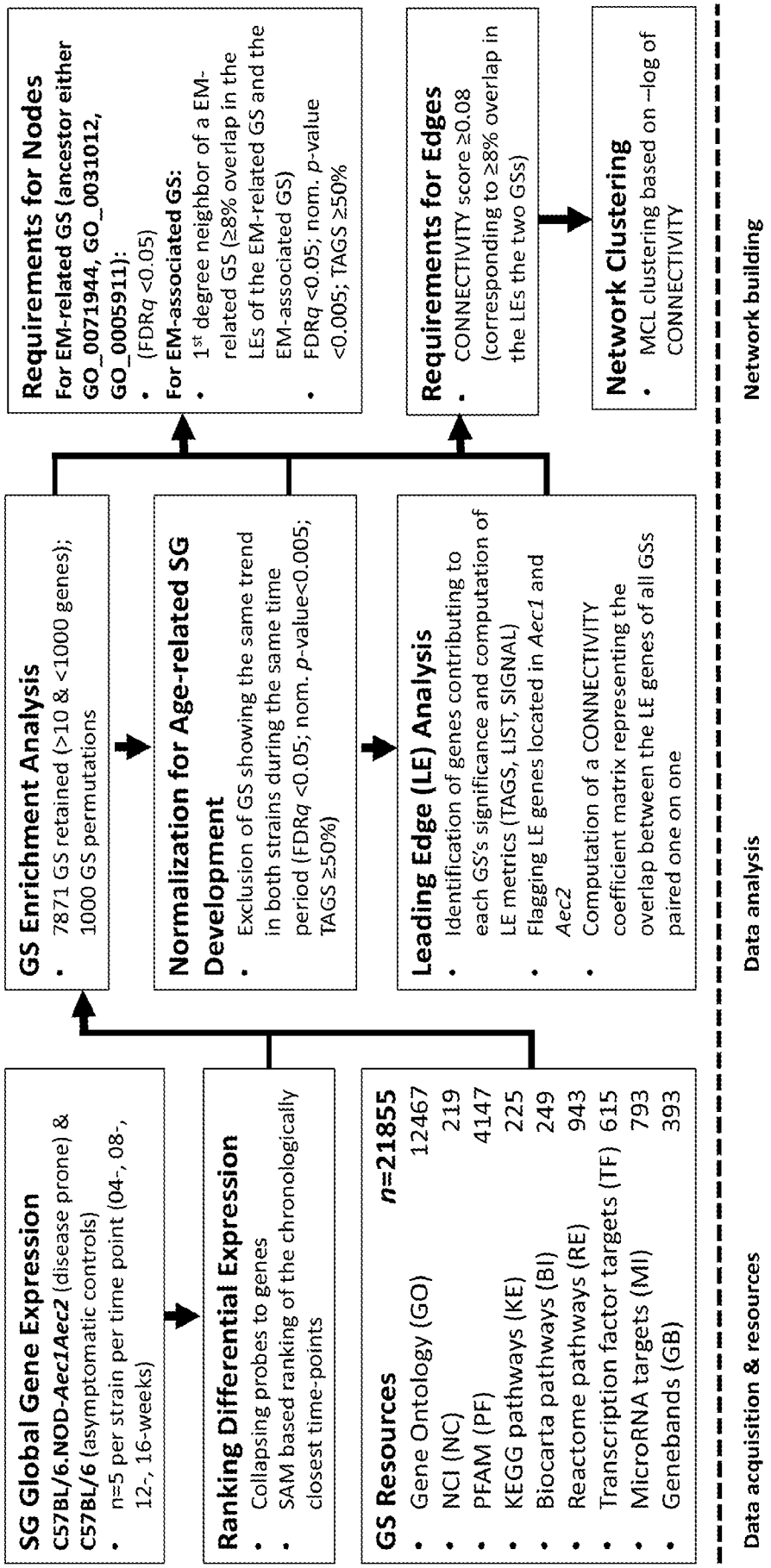
SG Global Gene Expression
C57BL/6.NOD-Aec1Aec2 (disease prone) & C57BL/6 (asymptomatic controls)
n=5 per strain per time point (04-, 08-, 12-, 16-weeks)
Ranking Differential Expression
Collapsing probes to genes
SAM based ranking of the chronologically closest time-points
GS Resources n=21855
Gene Ontology (GO) 12467
NCI (NC) 219
PFAM (PF) 4147
KEGG pathways (KE) 225
Biocarta pathways (BI) 249
Reactome pathways (RE) 943
Transcription factor targets (TF) 615
MicroRNA targets (MI) 793
Genebands (GB) 393
GS Enrichment Analysis
7871 GS retained (>10 & <1000 genes); 1000 GS permutations
Normalization for Age-related SG Development
Exclusion of GS showing the same trend in both strains during the same time period (FDRq <0.05; nom. p-value<0.005; TAGS ≥50%)
Leading Edge (LE) Analysis
Identification of genes contributing to each GS's significance and computation of LE metrics (TAGS, LIST, SIGNAL)
Flagging LE genes located in Aec1 and Aec2
Computation of a CONNECTIVITY coefficient matrix representing the overlap between the LE genes of all GSs paired one on one
Requirements for Nodes
For EM-related GS (ancestor either GO_0071944, GO_0031012, GO_0005911):
(FDRq <0.05)
For EM-associated GS:
1st degree neighbor of a EM-related GS (≥8% overlap in the LEs of the EM-related GS and the EM-associated GS)
FDRq <0.05; nom. p-value <0.005; TAGS ≥50%
Requirements for Edges
CONNECTIVITY score ≥0.08 (corresponding to ≥8% overlap in the LEs the two GSs)
Network Clustering
MCL clustering based on –log of CONNECTIVITY
Data acquisition & resources
Data analysis
Network building

FIG. 4

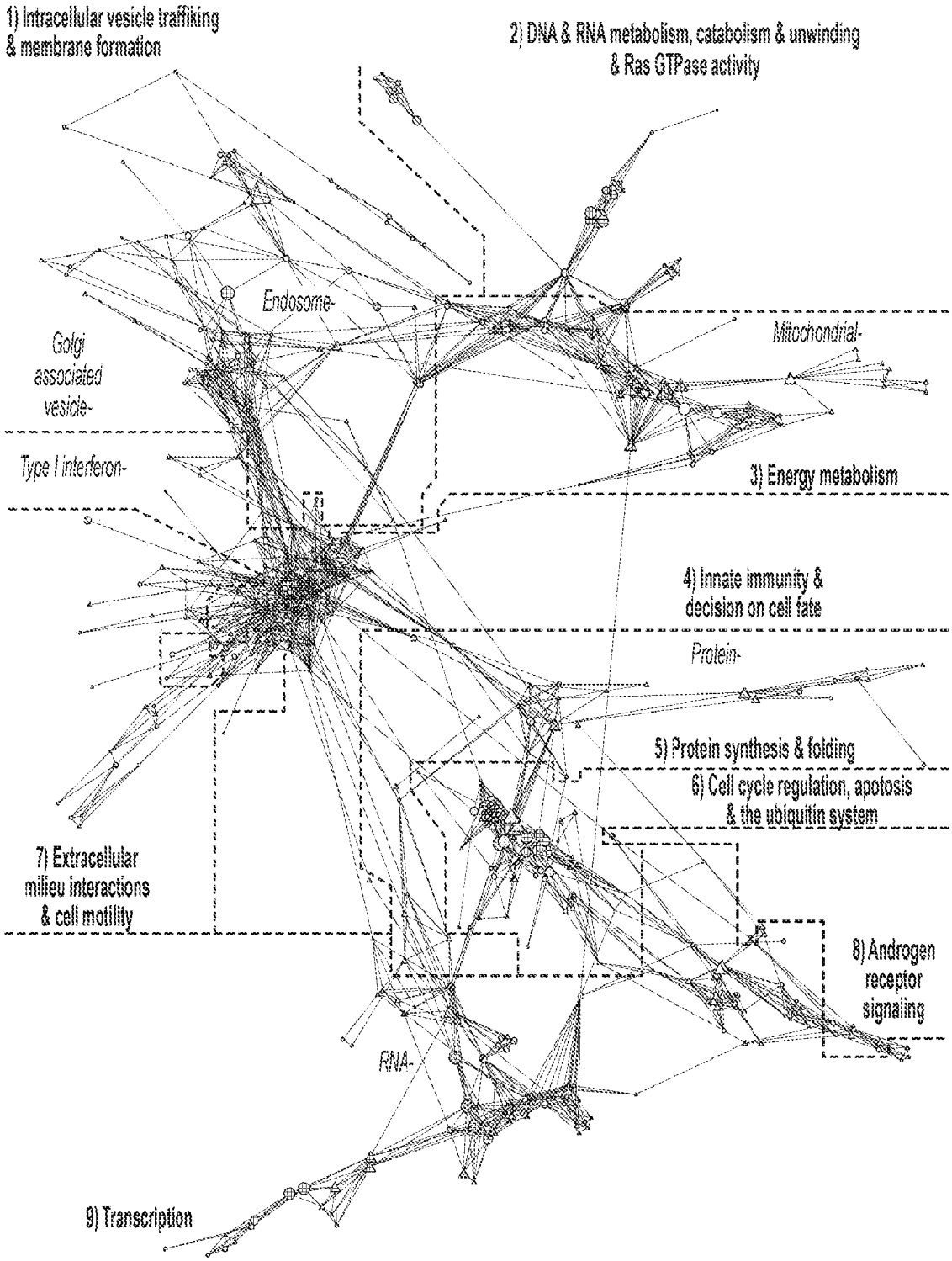
1) Intracellular vesicle trafficking & membrane formation
2) DNA & RNA metabolism, catabolism & unwinding & Ras GTPase activity
Endosome-
Golgi associated vesicle-
Mitochondrial-
Type I interferon-
3) Energy metabolism
4) Innate immunity & decision on cell fate
Protein-
5) Protein synthesis & folding
6) Cell cycle regulation, apotosis & the ubiquitin system
7) Extracellular milieu interactions & cell motility
8) Androgen receptor signaling
RNA-
9) Transcription

FIG. 5

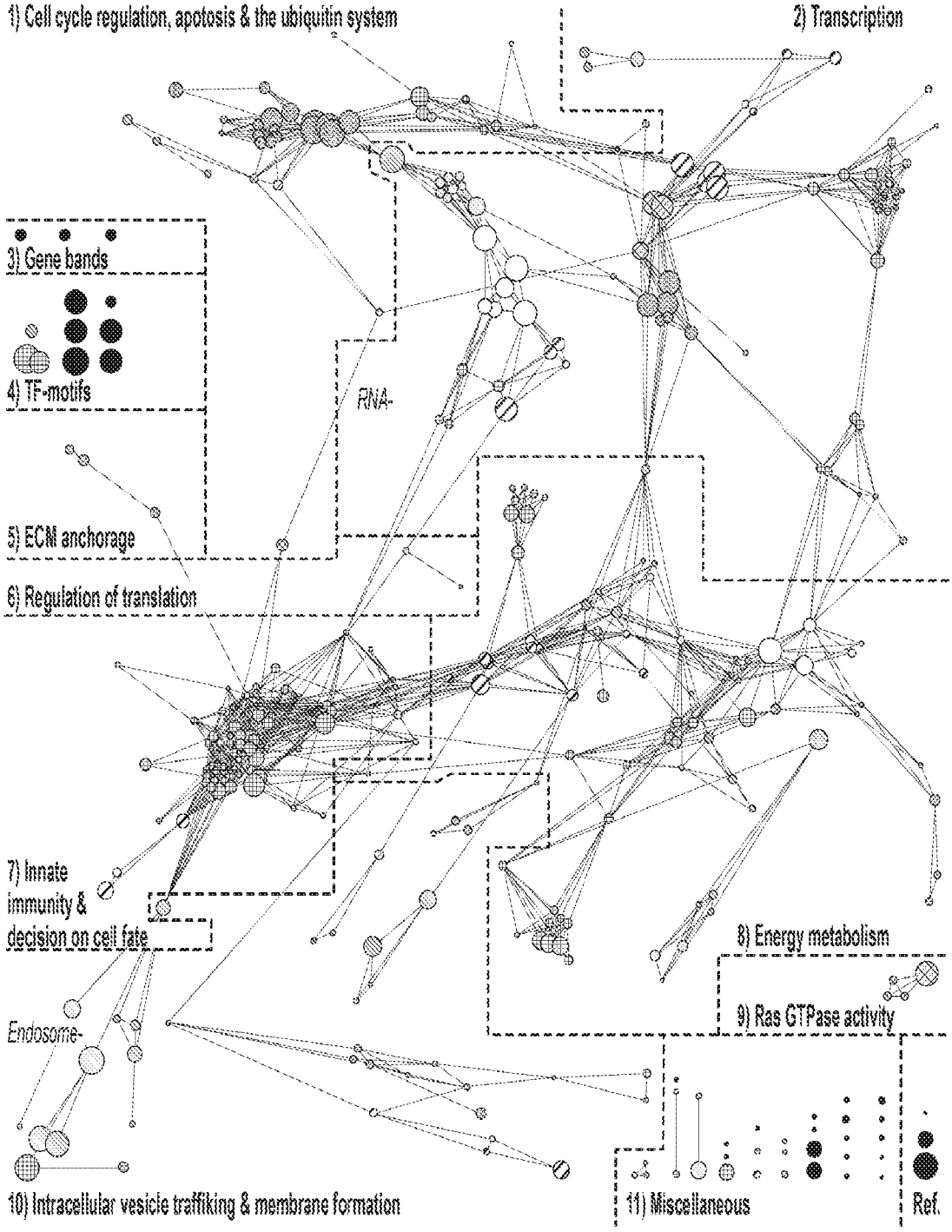
1) Cell cycle regulation, apotosis & the ubiquitin system
2) Transcription
3) Gene bands
4) TF-motifs
RNA-
5) ECM anchorage
6) Regulation of translation
7) Innate immunity & decision on cell fate
8) Energy metabolism
9) Ras GTPase activity
Endosome-
10) Intracellular vesicle trafficking & membrane formation
11) Miscellaneous
Ref.

FIG. 6

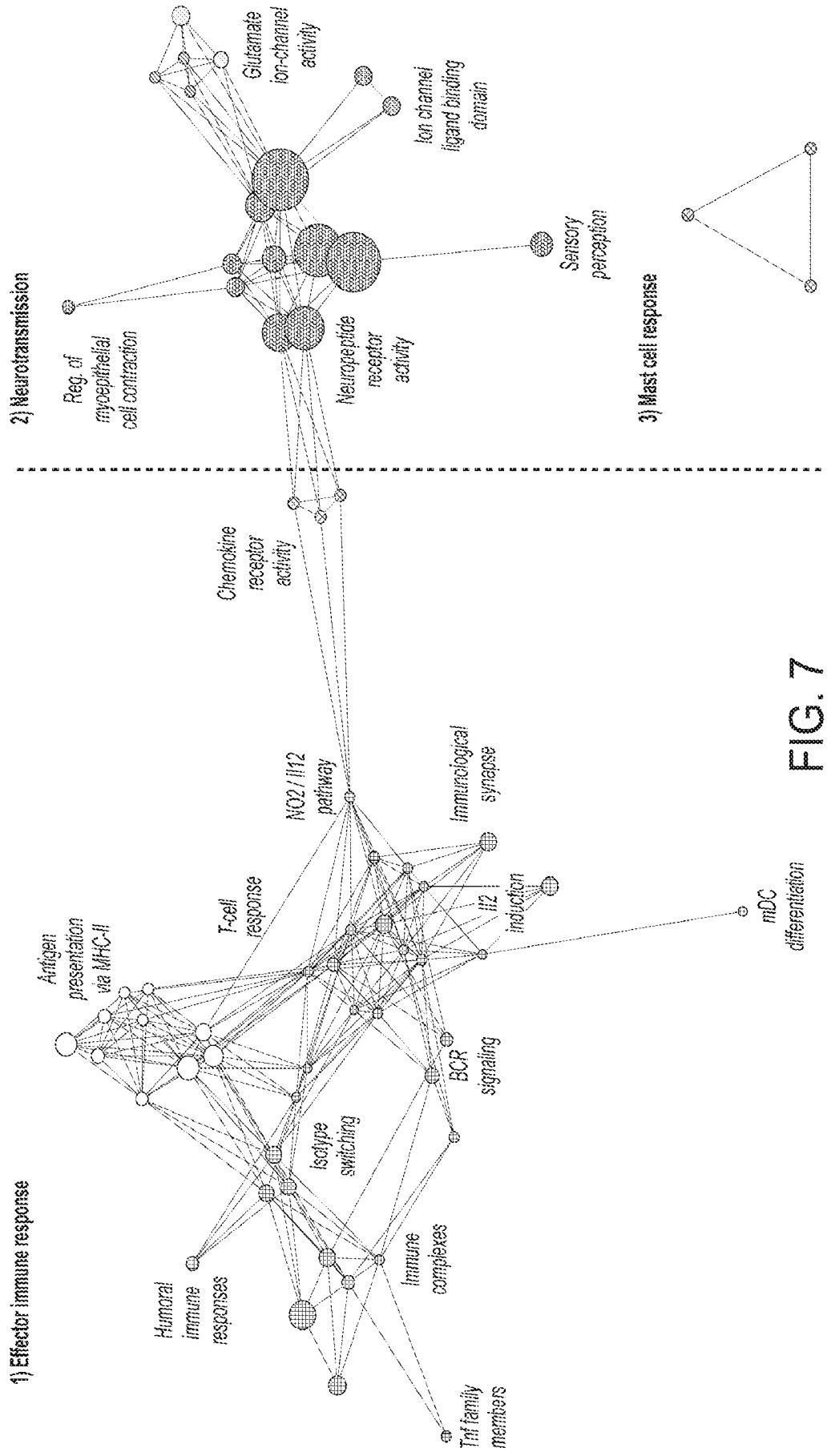
1) Effector immune response
2) Neurotransmission
3) Mast cell response
Antigen presentation via MHC-II
T-cell response
NO2 / Il12 pathway
Immunological synapse
Il2 induction
mDC differentiation
BCR signaling
Isotype switching
Humoral immune responses
Immune complexes
Tnf family members
Chemokine receptor activity
Reg. of myoepithelial cell contraction
Neuropeptide receptor activity
Glutamate ion-channel activity
Ion channel ligand binding domain
Sensory perception

FIG. 7

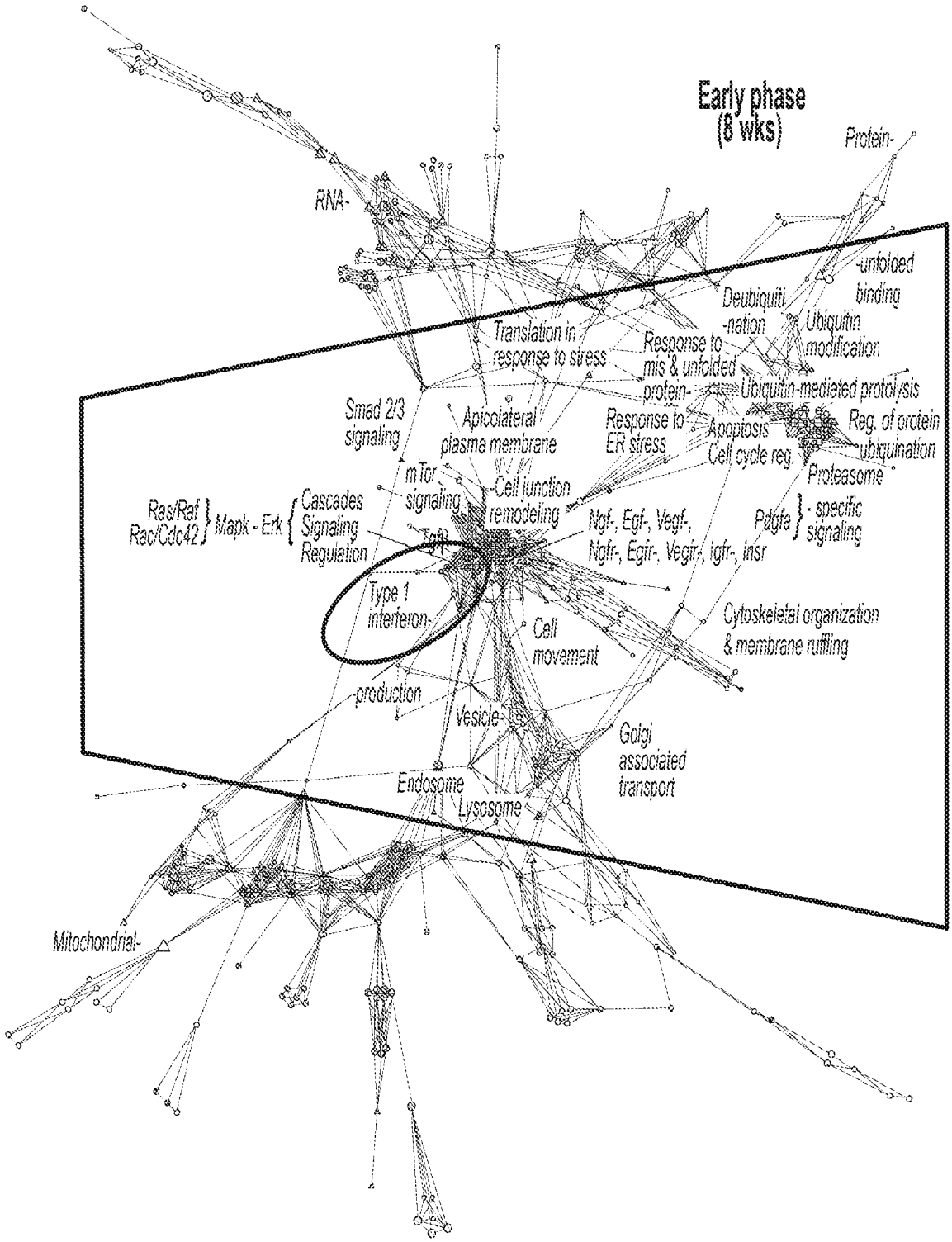
Early phase
(8 wks)
RNA-
Protein-
-unfolded
binding
Deubiquiti
-nation
Ubiquitin
modification
Translation in
response to stress
Response to
mis & unfolded
protein-
Ubiquitin-mediated proteolysis
Smad 2/3
signaling
Apicolateral
plasma membrane
Response to
ER stress
Apoptosis
Cell cycle reg.
Reg. of protein
ubiquination
Proteasome
mTor
signaling
Cell junction
remodeling
Ras/Raf
Rac/Cdc42
Mapk - Erk
Cascades
Signaling
Regulation
Tgfb
Ngf-, Egf-, Vegf-,
Ngfr-, Egfr-, Vegfr-, Igfr-, Insr
Pdgfa
- specific
signaling
Type 1
interferon-
Cell
movement
Cytoskeletal organization
& membrane ruffling
-production
Vesicle-
Golgi
associated
transport
Endosome
Lysosome
Mitochondrial-

FIG. 8

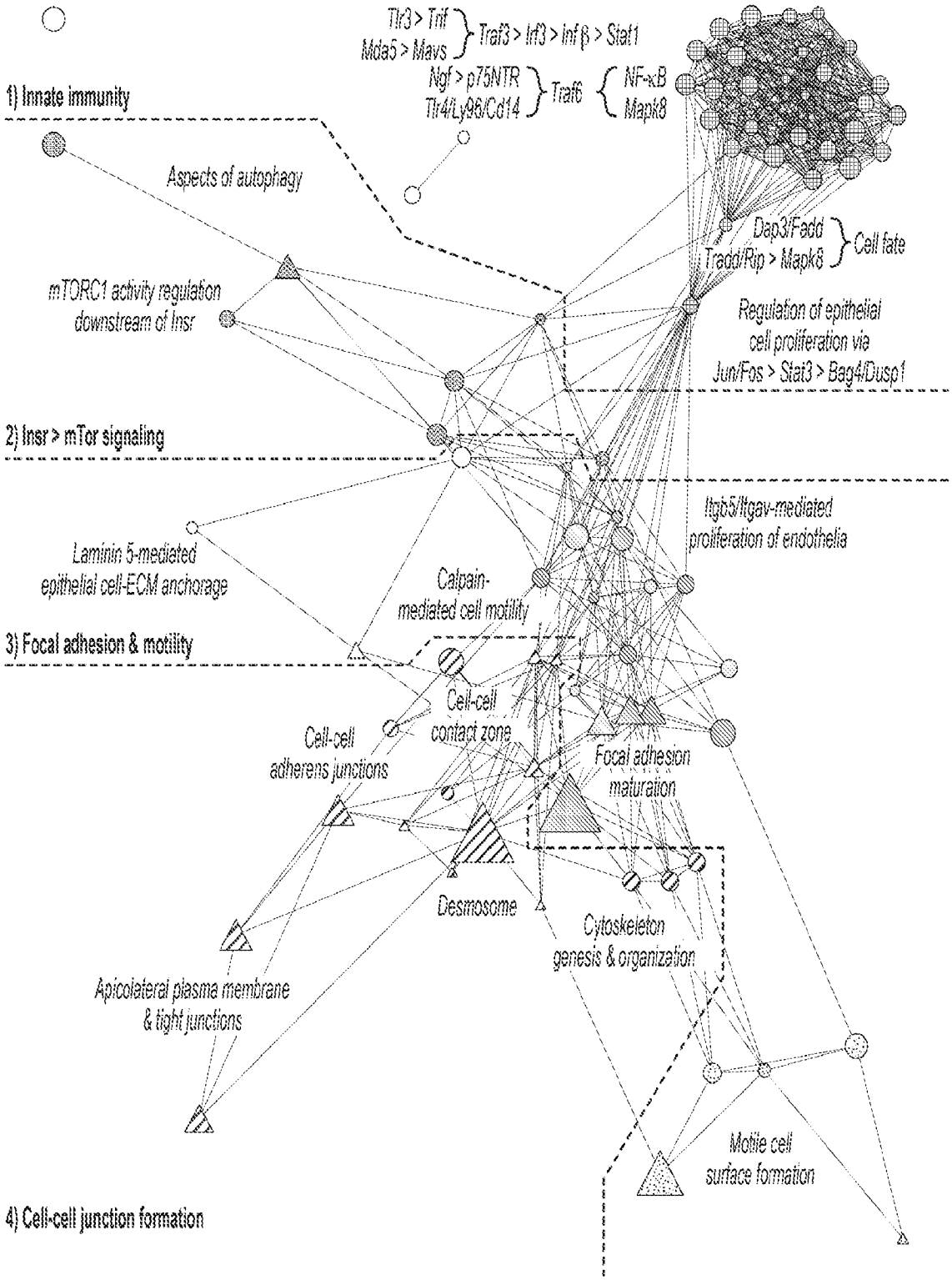
Tlr3 > Trif
Mda5 > Mavs
Traf3 > Irf3 > Infβ > Stat1
Ngf > p75NTR
Tlr4/Ly96/Cd14
Traf6
NF-κB
Mapk8
1) Innate immunity
Aspects of autophagy
Dap3/Fadd
Tradd/Rip > Mapk8
Cell fate
mTORC1 activity regulation downstream of Insr
Regulation of epithelial cell proliferation via Jun/Fos > Stat3 > Bag4/Dusp1
2) Insr > mTor signaling
Itgb5/Itgav-mediated proliferation of endothelia
Laminin 5-mediated epithelial cell-ECM anchorage
Calpain-mediated cell motility
3) Focal adhesion & motility
Cell-cell contact zone
Cell-cell adherens junctions
Focal adhesion maturation
Desmosome
Cytoskeleton genesis & organization
Apicolateral plasma membrane & tight junctions
Motile cell surface formation
4) Cell-cell junction formation

FIG. 9

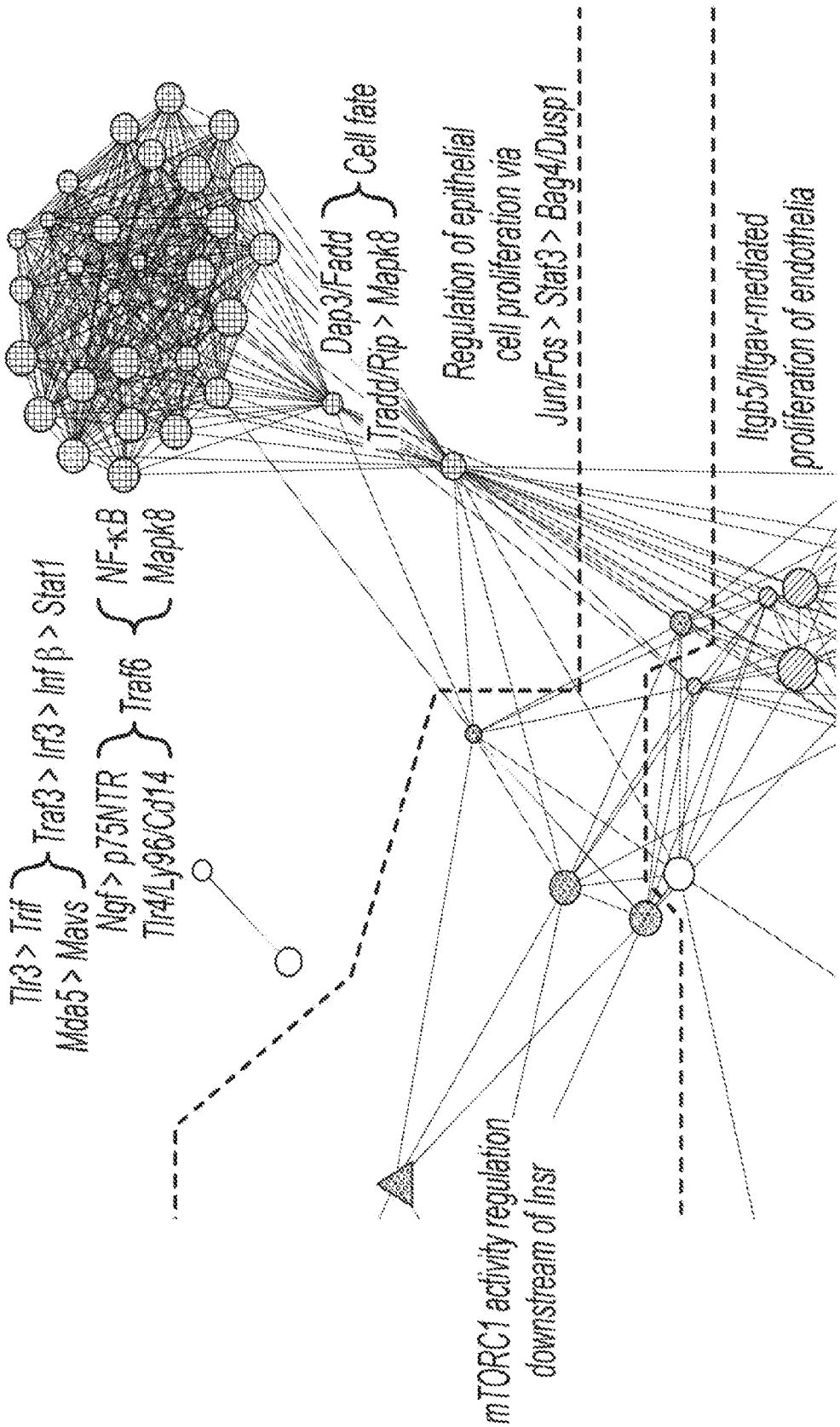
Tlr3 > Trif
Mda5 > Mavs
Ngf > p75NTR
Tlr4/Ly96/Cd14
Traf3 > Irf3 > Inf β > Stat1
Traf6
NF-κB
Mapk8
Dap3/Fadd
Tradd/Rip > Mapk8
Cell fate
Regulation of epithelial cell proliferation via Jun/Fos > Stat3 > Bag4/Dusp1
Itgb5/Itgav-mediated proliferation of endothelia
mTORC1 activity regulation downstream of Insr

FIG. 10

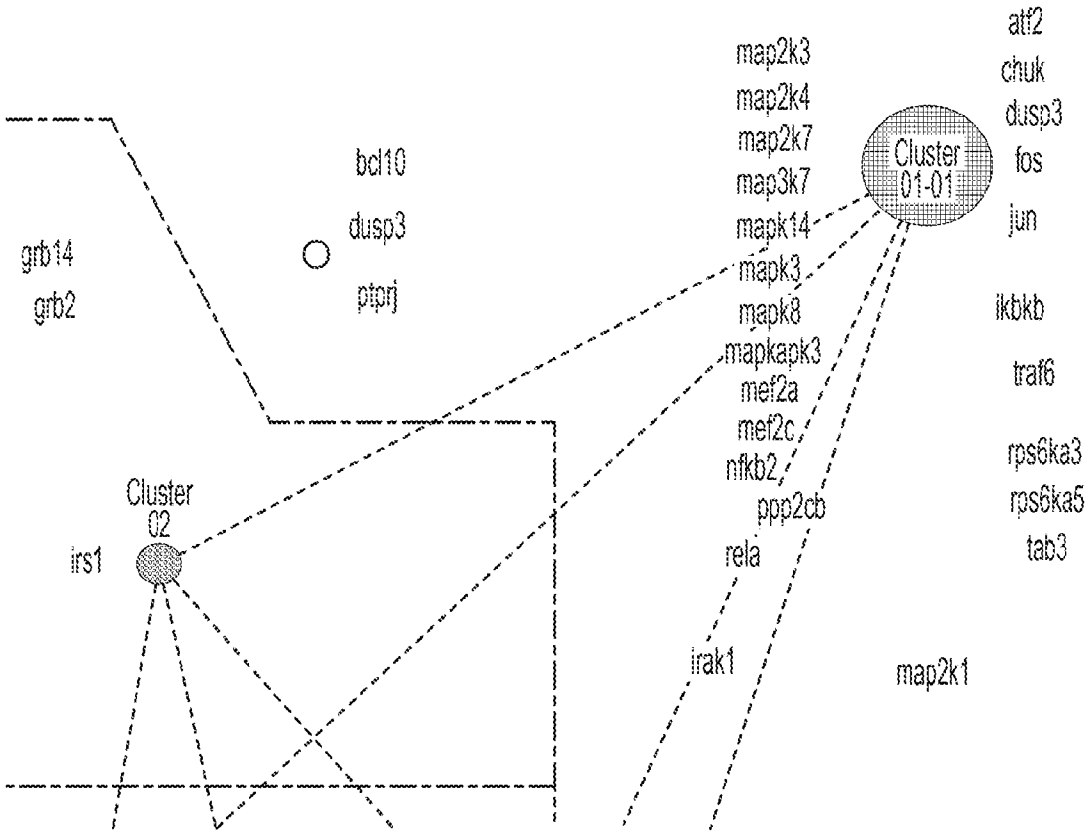
atf2
map2k3
chuk
map2k4
dusp3
map2k7
Cluster
01-01
fos
bcl10
map3k7
jun
dusp3
mapk14
grb14
mapk3
ptprj
ikbkb
grb2
mapk8
mapkapk3
traf6
mef2a
mef2c
rps6ka3
Cluster
02
nfkb2
rps6ka5
ppp2cb
irs1
rela
tab3
irak1
map2k1

FIG. 11

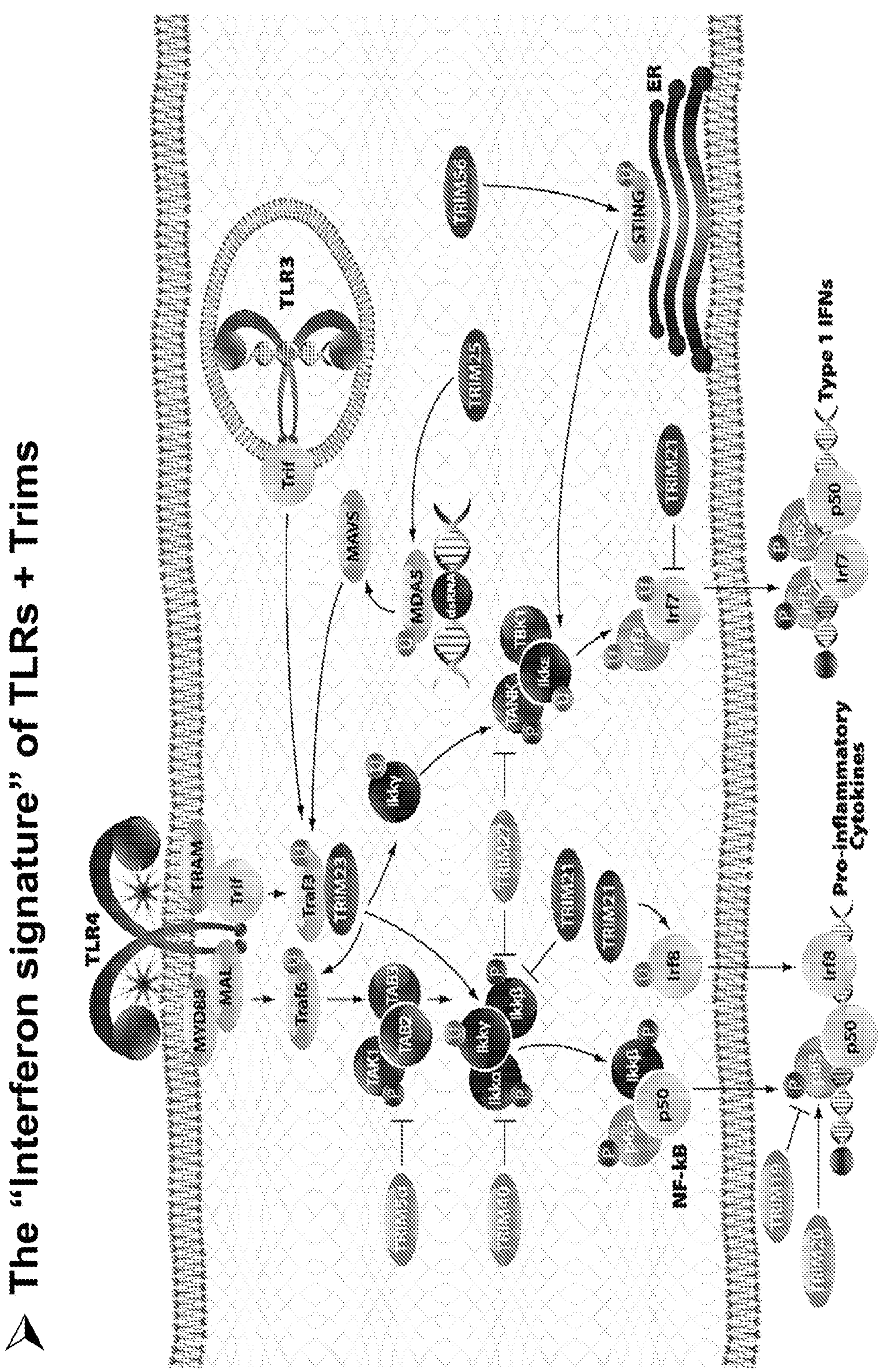
The "Interferon signature" of TLRs + Trims
TLR4
TLR3
MYD88
MAL
TRAM
Trif
Traf6
Traf3
TRIM25
MAVS
MDA5
TRIM56
STING
ER
Irf7
Irf8
p50
NF-kB
Pro-inflammatory Cytokines
Type 1 IFNs

FIG. 12

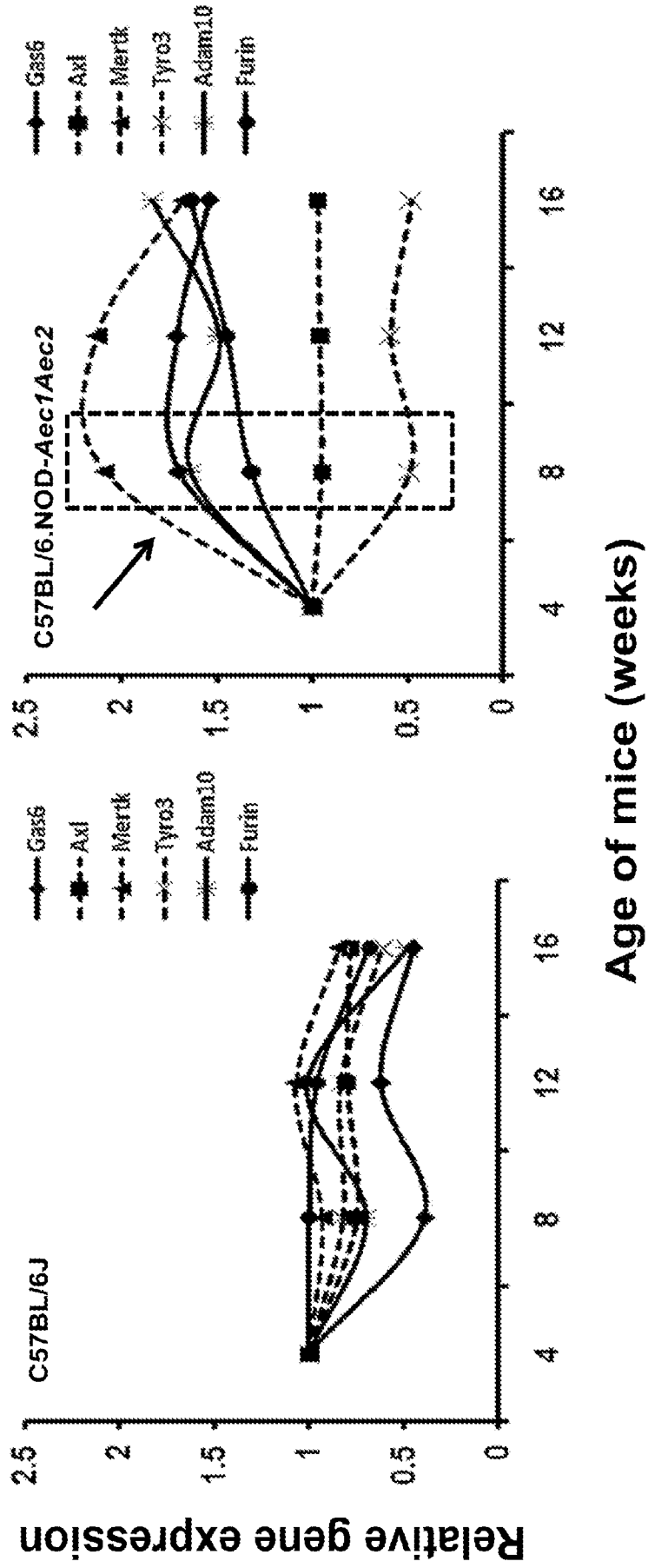
Gas6
Axl
Mertk
Tyro3
Adam10
Furin
C57BL/6.NOD-Aec1Aec2
C57BL/6J
2.5
2
1.5
1
0.5
0
4
8
12
16
Relative gene expression
Age of mice (weeks)

FIG. 13

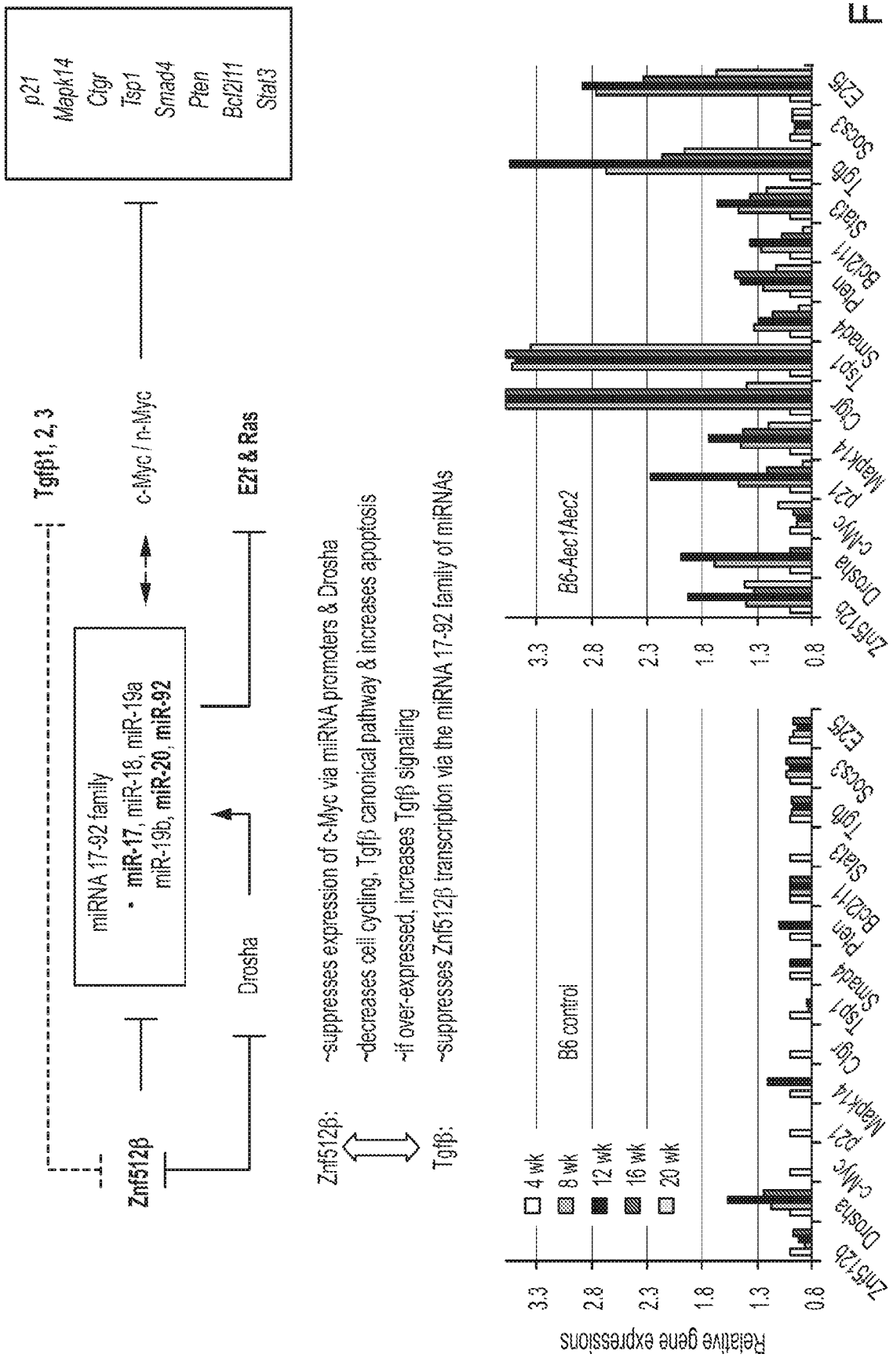
Tgfβ1, 2, 3
Znf512β
miRNA 17-92 family
• miR-17, miR-18, miR-19a
miR-19b, miR-20, miR-92
Drosha
c-Myc / n-Myc
E2f & Ras
p21
Mapk14
Ctgf
Tsp1
Smad4
Pten
Bcl2l11
Stat3
Znf512β:
~suppresses expression of c-Myc via miRNA promoters & Drosha
~decreases cell cycling, Tgfβ canonical pathway & increases apoptosis
~if over-expressed, increases Tgfβ signaling
Tgfβ:
~suppresses Znf512β transcription via the miRNA 17-92 family of miRNAs
4 wk
8 wk
12 wk
16 wk
20 wk
B6 control
B6-Aec1Aec2
Relative gene expressions
3.3
2.8
2.3
1.8
1.3
0.8
Znf512b
Drosha
c-Myc
p21
Mapk14
Ctgf
Tsp1
Smad4
Pten
Bcl2l11
Stat3
Tgfb
Socs3
E2f3

FIG. 14

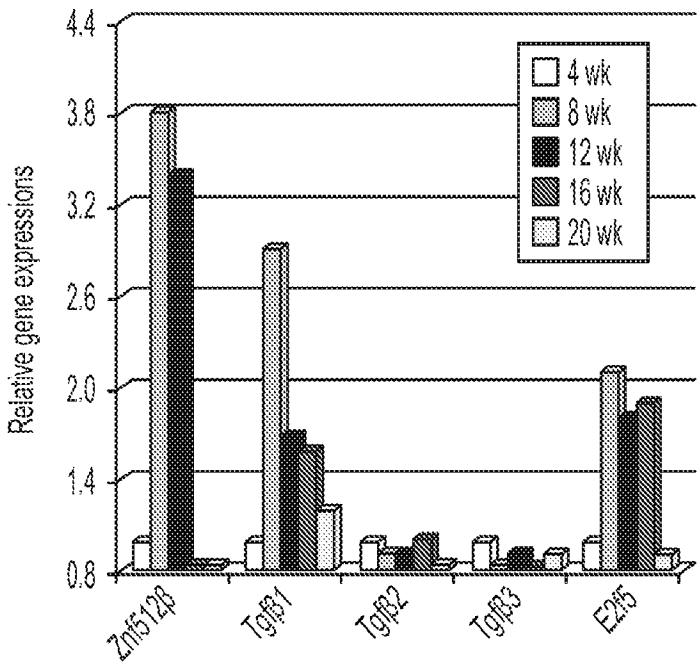
4.4
3.8
3.2
2.6
2.0
1.4
0.8
Relative gene expressions
4 wk
8 wk
12 wk
16 wk
20 wk
Znf512B
Tgfβ1
Tgfβ2
Tgfβ3
E2f5

FIG. 15

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 16/30762

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - C12Q 1/68 (2016.01)
CPC - C12Q 1/6883, C07H 21/00, C12Q 1/6837
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)
IPC (8):C12Q 1/68 (2016.01)
CPC: C12Q 1/6883, C07H 21/00, C12Q 1/6837

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC: 435/6.11, 435/91.1, 435/287.2

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Google patents, Google scholar, Google web, PatBase, Proquest Dialog
diagnosis; array/microarray; expression; level; interferon; regulated/responsive/stimulate; gene; signature/profile; oligonucleotide/probe

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	US 6,331,396 B1 (SILVERMAN et al.) 18 December 2001 (18.12.2001) Abstract; claim 1; claim 2; Tables 2-8; Col. 2, ln 35-48	1, 4/1 --- 2-3, 4/2-3
X --- Y	US 2014/0135225 A1 (NEW YORK SOC FOR THE RUPTURED AND CRIPPLED MAINTAINING THE HOSPITAL FOR SPE) 15 May 2014 (15.05.2014) Abstract; claim 1; para [0023]; para [0024]; Table 2; para [0124]; para [0301]; para [0304]; para [0170]; para [0274]; para [0306]; para [0031]; para [0065]; para [0170]; para [0286]; para [0317]; para [0109]; para [0141]; para [0177]	27, 29-33 --- 34
Y	US 2010/0279298 A1 (ABBAS) 04 November 2010 (04.11.2010) claim 8; para [0091]; para [0185]; para [0020]	2-3, 4/2-3
Y	US 2014/0030261 A1 (VERENIGING VOOR CHRISTELIJK HOGER ONDERWIJS et al.) 30 January 2014 (30.01.2014) para [0175]; para [0184]; para [0046]; claim 1; para [0003]; Abstract	34

☐ Further documents are listed in the continuation of Box C. ☐

* 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
05 July 2016 (05.07.2016)	12 AUG 2016
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-8300	Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 16/30762

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 5-26, 28, 35-43
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.