(54) **Title:** AUTOMATED EVALUATION OF EMBRYO IMPLANTATION POTENTIAL

FIG. 1

(57) **Abstract:** A system comprising: at least one hardware processor; and a non-transitory computer-readable storage medium having stored thereon program code executable by the at least one hardware processor to: receiving a plurality of video segments, each depicting prenatal embryogenesis of a corresponding embryo; dividing each of said video segments into a plurality of consecutive packets each comprising a specified number of frames; training a first machine learning model on a training set comprising (i) said packets, and (ii) labels indicating a developmental parameter associated with each of said corresponding embryos; and training a second machine learning model on a training set comprising: (iii) sets of outputs of said first machine learning model, wherein each of said sets is associated with said packets comprising one of said video segments, and (iv) labels indicating said developmental parameter associated with said corresponding embryo depicted in said one of said video segments.

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))*

AUTOMATED EVALUATION OF EMBRYO IMPLANTATION POTENTIAL

FIELD OF THE INVENTION

[0001]   The invention relates generally to the field of machine learning.

CROSS REFERENCE TO RELATED APPLICATIONS

[0002]   This application claims the benefit of priority from U.S. Provisional Patent Application No. 62799384 filed January 31, 2019, the contents of which are incorporated herein by reference in their entirety.

BACKGROUND

[0003]   In IVF treatments, early identification of embryos with high implantation potential is essential for avoiding clinical complications to the newborn and/or to the mother, and for shortening the time until achieving a successful pregnancy. Known embryo automated classification tools are used to evaluate embryo developmental competence, for example, based on manual scoring of multiple morphological properties of an embryo at a single time point just before transfer into the uterus. The incorporation of time-lapse incubators in IVF clinics provides continuous visual monitoring of the embryos, while maintaining them in optimal culture conditions. Based on such video recordings, embryos can be represented by the time sequence of developmental events that are used by morphokinetic algorithms for predicting embryo blastulation and implantation. Both morphological and morphokinetic classifiers require manual annotation by expert personnel and are limited to a discrete representation of embryo preimplantation development while ignoring dynamic features associated with embryo quality.

[0004]   The foregoing examples of the related art and limitations related therewith are intended to be illustrative and not exclusive. Other limitations of the related art will become apparent to those of skill in the art upon a reading of the specification and a study of the figures.

SUMMARY

[0005]   The following embodiments and aspects thereof are described and illustrated in conjunction with systems, tools and methods which are meant to be exemplary and illustrative, not limiting in scope.
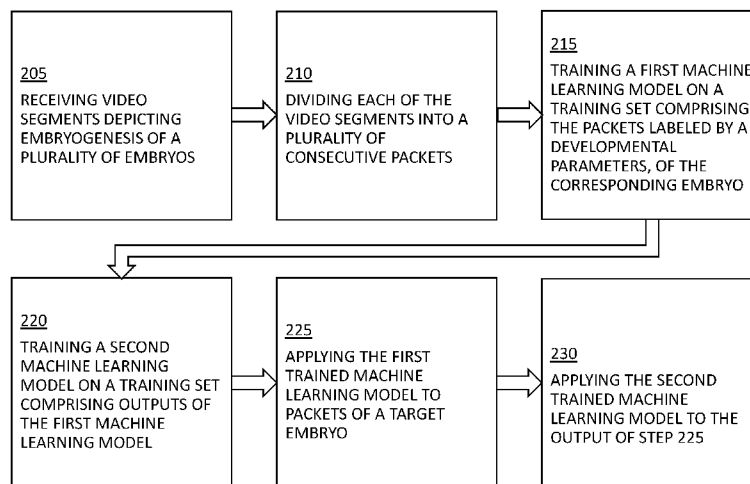
[0006]   There is provided in an embodiment, a system comprising: at least one hardware processor; and a non-transitory computer-readable storage medium having stored thereon program code, the program code executable by the at least one hardware processor to: receive a plurality of video segments, each depicting prenatal embryogenesis of a corresponding embryo; divide each of said video segments into a plurality of consecutive packets, wherein each of said plurality of packets comprises a specified number of frames; train a first machine learning model on a training set comprising (i) said packets, and (ii) labels indicating a developmental parameter associated with each of said corresponding embryos; and train a second machine learning model on a training set comprising: (iii) sets of outputs of said first machine learning model, wherein each of said sets is associated with said packets comprising one of said video segments, and (iv) labels indicating said developmental parameter associated with said corresponding embryo depicted in said one of said video segments.

[0007]   There is also provided in an embodiment a method comprising: receiving a plurality of video segments, each depicting prenatal embryogenesis of a corresponding embryo; dividing each of said video segments into a plurality of consecutive packets, wherein each of said plurality of packets comprises a specified number of frames; training a first machine learning model on a training set comprising (i) said packets, and (ii) labels indicating a developmental parameter associated with each of said corresponding embryos; and training a second machine learning model on a training set comprising: (iii) sets of outputs of said first machine learning model, wherein each of said sets is associated with said packets comprising one of said video segments, and (iv) labels indicating said developmental parameter associated with said corresponding embryo depicted in said one of said video segments

[0008]   A computer program product comprising a non-transitory computer-readable storage medium having program code embodied therewith, the program code executable by

at least one hardware processor to: receive a plurality of video segments, each depicting prenatal embryogenesis of a corresponding embryo; divide each of said video segments into a plurality of consecutive packets, wherein each of said plurality of packets comprises a specified number of frames; train a first machine learning model on a training set comprising (i) said packets, and (ii) labels indicating a developmental parameter associated with each of said corresponding embryos; and train a second machine learning model on a training set comprising: (iii) sets of outputs of said first machine learning model, wherein each of said sets is associated with said packets comprising one of said video segments, and (iv) labels indicating said developmental parameter associated with said corresponding embryo depicted in said one of said video segments.

[0009] In some embodiments, with respect to each of said packets, the output of said trained first machine learning model is a numerical representation indicating a probability associated with said developmental parameter.

[0010] In some embodiments, the numerical representation is one of: a scalar representation, a vector representation, and a matrix representation.

[0011] In some embodiments, the numerical representation reflects a dimensionality reduction.

[0012] In some embodiments, the trained second machine learning model predicts a developmental potential associated with each of said corresponding embryos.

[0013] In some embodiments, the program instructions are further executable to apply, and the method further comprises applying, at an inference stage: (i) said trained first machine learning model to target packets associated with a target video segment depicting prenatal embryogenesis of a target embryo, to obtain said numerical representations for each of said target packets; and said trained second machine learning model to said obtained numerical representations, to predict a developmental potential of said target embryo.

[0014] In some embodiments, the first machine learning model comprises at least two machine learning models, wherein: (i) with respect to a first of said machine learning models, said developmental parameter indicated by said labels is a blastulation state; and (ii) with

respect to a second of said machine learning models, said developmental parameter indicated by said labels is an implantation state.

[0015]   In some embodiments, the developmental parameter comprises at least one of: morphological stage, cleavage stage, number of cells, cell fragmentation, cell symmetry, inner cell mass, trophectoderm, pronuclei symmetry, pronuclei movement, pronuclei location, cell location, and cell movement. In some embodiments, the trained second machine learning model predicts at least one of: morphological stage, cleavage stage, number of cells, cell fragmentation, cell symmetry, inner cell mass, trophectoderm, pronuclei symmetry, pronuclei movement, pronuclei location, cell location, and cell movement.

[0016]   In some embodiments, the first machine learning model comprises a self-supervised algorithm. In addition to the exemplary aspects and embodiments described above, further aspects and embodiments will become apparent by reference to the figures and by study of the following detailed description.

BRIEF DESCRIPTION OF THE FIGURES

[0017]   Exemplary embodiments are illustrated in referenced figures. Dimensions of components and features shown in the figures are generally chosen for convenience and clarity of presentation and are not necessarily shown to scale. The figures are listed below.

[0018]   Fig. 1 is a flowchart of functional steps in a process for training a machine learning model to predict embryo implantation, in accordance with some embodiments of the present invention;

[0019]   Figs. 2A and 2B show tables of exemplary databases, in accordance with some embodiments of the present disclosure;

[0020]   Fig. 3 is a schematic illustration of exemplary data preprocessing, in accordance with some embodiments of the present disclosure;

[0021]   Figs. 4A and 4B show exemplary automated screenings of empty well images and cropping embryo region of interest (ROI), in accordance with some embodiments of the present disclosure;

[0022]   Figs. 5A and 5B show exemplary embodiments of image preprocessing comprising automated segmentation, down sampling of embryo region of interest (ROI), and discarding empty well images, in accordance with some embodiments of the present disclosure;

[0023]   Fig. 6 shows a table of exemplary morphokinetic time windows, in accordance with some embodiments of the present disclosure;

[0024]   Fig. 7 shows a table of exemplary morphokinetic intervals, in accordance with some embodiments of the present disclosure;

[0025]   Figs. 8A, 8B, 8C, 8D, 8E, 8F, and 8G show exemplary embodiments of identification of developmentally arrested embryos that fail to reach blastulation (BLAST_n), in accordance with some embodiments of the present disclosure;

[0026]   Figs. 9A, 9B and 9C show an exemplary database comprising of embryo video files and associated clinical metadata, in accordance with some embodiments of the present disclosure;

[0027]   Figs. 10A, 10B, 10C, and 10D are exemplary embodiments of automated predictions of embryo blastulation, in accordance with some embodiments of the present disclosure;

[0028]   Figs. 11A, 11B, 11C, 11D, 11E, and 11F show exemplary embodiments of high resolution morphokinetic analysis of embryo preimplantation development, in accordance with some embodiments of the present disclosure;

[0029]   Figs. 12A, 12B, 12C, and 12D show exemplary embodiments of statistical characteristics of blastulation prediction, in accordance with some embodiments of the present disclosure;

[0030]   Figs. 13A and 13B show exemplary embodiments of morphokinetic overlaps between different maternal age embryos, in accordance with some embodiments of the present disclosure;

[0031]   Figs. 14A, 14B, and 14C are exemplary embodiments of automated predictions of embryo implantation, in accordance with some embodiments of the present disclosure;

[0032]   Figs. 15A and 15B are exemplary embodiments of machine learning model optimization for prediction of embryo blastulation and implantation, in accordance with some embodiments of the present disclosure;

[0033]   Figs. 16A 16B and 16C are exemplary embodiments of machine learning model optimization for prediction of embryo blastulation and implantation, in accordance with some embodiments of the present disclosure;

[0034]   Fig. 17 is a schematic illustration of exemplary learning process, in accordance with some embodiments of the present disclosure;

[0035]   Figs. 18A and 18B show exemplary embodiments of top positive versus negative SHAP-scored embryo frames directing blastulation prediction, in accordance with some embodiments of the present disclosure;

[0036]   Figs. 19A and 19B show exemplary embodiments of top positive versus negative SHAP-scored embryo frames directing implantation prediction, in accordance with some embodiments of the present disclosure;

[0037]   Figs. 20A and 20B are exemplary embodiments of automated predictions of embryo blastulation, in accordance with some embodiments of the present disclosure;

[0038]   Figs. 21A and 21B shows exemplary embodiments of first direct unequal cleavage (DUC1) embryos that are selected by SHIFRA$_K$, in accordance with some embodiments of the present disclosure;

[0039]   Figs. 22A, 22B, and 22C show statistical characteristics of blastulation prediction by SHIFRA$_B$, in accordance with some embodiments of the present disclosure;

[0040]   Figs. 23A, 23B, and 23C are exemplary demonstrations of fully automated predictions of embryo implantation, in accordance with some embodiments of the present disclosure;

[0041]   Figs. 24A, 24B, 24C, 24D, and 24E show the SHIFRA$_B$ and SHIFRA$_K$ databases are optimized to predict blastulation and implantation with partial dependence on embryo state, in accordance with some embodiments of the present disclosure;

[0042]   Figs. 25A and 25B are exemplary embryonic state distributions statistical analysis, in accordance with some embodiments of the present disclosure;

[0043]   Figs. 26A and 26B are exemplary statistical characteristics of implantation prediction by SHIFRA$_K$, in accordance with some embodiments of the present disclosure;

[0044]   Fig. 27 is a table of the KIDScore-D3, in accordance with some embodiments of the present disclosure;

[0045]   Figs. 28A and 28B are exemplary embryo transfer and implantation statistics, in accordance with some embodiments of the present disclosure;

[0046]   Figs. 29A, 29B, 29C, and 29D shows that classification of embryo blastulation and implantation is robust to differences in maternal age, in accordance with some embodiments of the present disclosure; and

[0047]   Figs. 30A, 30B, 30C, and 30D show exemplary embodiments of statistical characteristics of implantation prediction by SHIFRA$_K$, in accordance with some embodiments of the present disclosure.

DETAILED DESCRIPTION

[0048]   Disclosed herein are a system, method, and computer program product for automated evaluation of embryonic developmental potential and/or competence. In some embodiments, the present disclosure provides for an accurate prediction of embryo viability and/or implantation potential.

[0049]   The present disclosure provides for one or more machine learning models trained to predict embryo viability, developmental, and/or implantation competence, which were developed by training deep neural networks using video files depicting a plurality of blastulation-labelled and implantation-labelled embryos. In some embodiments, the present machine learning models provide for greater prediction accuracy compared to known classification techniques.

[0050]   In some embodiments, the present disclosure employs deep learning techniques to generate automated, accurate and standardized machine learning models for early prediction of embryo viability, developmental, and/or implantation potential. In some embodiments, a

prediction accuracy of the present machine learning models remains high irrespective of maternal age, without maternal age input. Deep learning methods employed by the present disclosure offer an automated, standardized and accurate substitute to human-based evaluation of embryonic developmental competence.

[0051] In some embodiments, the present machine learning models, which provide early evaluation of blastulation and implantation potential, may be incorporated into a systemic decision-making tool that may provide a personalized, multi-step embryo transfer strategy. Accordingly, given a finite number of embryos obtained from a patient and their assessed quality, this tool will specify the multistep order and timing of embryo transfers (including transfers of multiple embryos), as well as which embryos are to be cryopreserved for subsequent transfers. The general framework of the present disclosure opens the door for the implementation of such personalized clinical tools that will optimize conception rates while shortening time to pregnancy in IVF treatments.

[0052] In some embodiments, the present disclosure provides for training one or more machine learning models, based, at least in part, on training data comprising image data depicting at least a portion of a prenatal embryogenesis process of a plurality of embryos.

[0053] In some embodiments, the image data comprises a series of images and/or a video segment. In some embodiments, the image data depicts, with respect to each embryo, at least a portion of prenatal embryonic embryogenesis or embryo development process. In some embodiments, the image data with respect to each embryo comprises one or more time-lapse video segments, wherein an image is captured with a specified frequency, e.g., every several minutes, e.g., every 18-20 minutes.

[0054] In some embodiments, the image data comprises multiple video segments with respect to at least some of the embryos. In some embodiments, at least some video segments with respect to each embryo are acquired using multifocal plane microscopy. In some embodiments, multifocal plane microscopy or multiplane microscopy allows the tracking of the 3D dynamics in embryonic cell-level development at high temporal and spatial resolution, by simultaneously imaging different focal planes within the specimen.

[0055] In some embodiments, the image data is obtained using one or more of RGB imaging techniques, monochrome imaging, near infrared (NIR), sort-wave infrared (SWIR),

8

infrared (IR), ultraviolet (UV), multi spectral, hyperspectral, and/or any other and/or similar imaging techniques. In some embodiments, the image data may be taken using different imaging techniques, imaging equipment, from different distances and angles, using varying backgrounds and settings, and/or under different illumination and ambient conditions. In some embodiments, the image data is obtained using various magnification levels, which can be optical and/or digital magnification.

[0056]   In some embodiments, the obtained images undergo image processing analyses comprising at least some of: data cleaning, data normalization, data standardization, and/or similar and/or additional image preprocessing steps. In some embodiments, image data preprocessing may comprise any one or more of object detection, object identification, image   transformations,   and/or   object   segmentation,   similar and/or   additional image preprocessing steps.

[0057]   In some embodiments, the training data comprises image data and/or additional and/or other clinical metadata associated with embryo development covering, e.g., at least a portion of the period between fertilization and implantation.

[0058]   In some embodiments, image data may comprise a plurality of video segments each depicting prenatal embryogenesis of a corresponding embryo. In some embodiments, each video segment may be associated with an indication of a developmental parameter of the depicted embryo.

[0059]   In some embodiments, the developmental parameter may be a blastulation and/or implantation state. In some embodiments, the developmental parameter may be any one of morphological stage, cleavage stage, number of cells, cell fragmentation, cell symmetry, inner cell mass, trophectoderm, pronuclei symmetry, pronuclei movement, pronuclei location, cell location, and cell movement.

[0060]   In some embodiments, each video segment may be divided into a plurality of consecutive packets, wherein each of said plurality of packets comprises a specified number of frames, e.g., between 3 and 7, e.g., 5 image frames. In some embodiments, each packet may be associated with an indication and/or annotation and/or labels reflecting one or more developmental parameters of the corresponding embryo.

[0061] In some embodiments, one or more machine learning models may be trained to produce a reduced dimension representation of each packer, e.g., a scalar representation, a vector representation, and/or a matrix representation.

[0062] In some embodiments, the reduced dimension representation of the packets may be used as part of a training set, to train a second machine learning model on sets of packets associated with each complete video segment (and hence, a corresponding embryo) wherein the sets may be associated with an indication and/or annotation and/or labels reflecting one or more developmental parameters of the corresponding embryo.

[0063] In some embodiments, the second machine learning model may output a prediction associated with embryo viability, developmental, and/or implantation competence and/or potential.

[0064] In some embodiments, one or more training sets may be generated based on the training data. In some embodiments, a first training set may be labeled with an indication of blastocyst formation in each of the embryos depicted in the image data. In some embodiments, a second training set may be labeled with an indication of a successful eventual implantation associated with each embryo depicted in the image data.

[0065] In some embodiments, one or more blastulation machine learning models may be trained on the first dataset, to predict an embryo blastulation outcome. In some embodiments, one or more implantation machine learning models may be trained on the second dataset to predict embryo implantation success.

[0066] In some embodiments, at an inference stage, one or more of the trained machine learning models may be applied to image data associated with an embryo, to predict implantation success of the embryo. In some embodiments, implantation success may be determined based on a combination of prediction scores of the different machine learning models.

[0067] In some embodiments, the method comprises training a first implantation machine learning model on a training set comprising a plurality of packets associated with an implantation indication. In some embodiments, the first implantation machine learning model is trained using a database of packets associated with an implantation indication. In

some embodiments, the first implantation machine learning model is trained to score at least one of the frames of each packet a plurality of times, based on, at least in part, temporal features of the embryo depicted by the frame. In some embodiments, the first implantation machine learning model is configured to output a sum of the scores of each frame. In some embodiments, first implantation machine learning model is configured to output a vector comprising the sum of the scores of each packet.

[0068]   In some embodiments, the second implantation machine learning model is trained using a database of packets, wherein each packet comprises a label associated with a label of the corresponding embryo depicted by the packet, the scores and/or vector outputted by the first implantation machine learning model of each packet.

[0069]   In some embodiments, the second implantation machine learning model is trained using at least one vector outputted by the first blastulation machine learning model for each packet. In some embodiments, for each frame and/or packet, the scores outputted by the first implantation machine learning model and the first blastulation machine learning model were summed into a single number and/or vector.

[0070]   In some embodiments, the labels inputted into the second implantation machine learning model for each frame depicting a specific embryo indicate if the specific embryo has resulted in successful implantation. In some embodiments, the labels of an embryo include positive implantation for an embryo that has resulted in successful implantation, negative implantation for an embryo that has not resulted in successful implantation, and unknown implantation for embryos for which the implantation success was undetermined.

[0071]   In some embodiments, the second implantation machine learning model is configured to evaluate the implantation potential of an embryo depicted by a packet based on, at least in part, the scores and/or vectors of the frames and/or packets outputted by one or more of the first implantation machine learning model and the first blastulation machine learning model.

[0072]   In some embodiments, the method comprises obtaining a target video segment depicting a development period of a target embryo pre-implantation. In some embodiments, the method comprises applying the first and second blastulation machine learning models to obtain a blastulation evaluation associated with each of said target packets. In some

embodiments, the method comprises applying the first and second implantation machine learning models to obtain an implantation evaluation associated with each of said target packets.

[0073] Reference is made to Fig. 1, which is a flowchart of functional steps in a process for training a machine learning model to predict embryo implantation, in accordance with some embodiments of the present invention.

[0074] According to some embodiments of the present disclosure, there is provided a method for evaluating the developmental competence of an embryo. In some embodiments, at step 205, the method comprises receiving video segments each depicting a development period of a corresponding embryo during pre-implantation stages. In some embodiments, each of the video segments comprises an indication of a developmental parameters of a corresponding embryo.

[0075] In some embodiments, at step 210, the method comprises dividing each of the video segments into a plurality of consecutive packets. In some embodiments, each embryo is associated with a plurality of packets which depict a sequence of developmental events. In some embodiments, each of the plurality of packets comprises a specified number of frames. In some embodiments, the frames are consecutive or separated by specific time increments. In some embodiments, the frames depict one or more focal points of the video segments.

[0076] In some embodiments, at step 215, the method comprises training a first blastulation machine learning model on a training set comprising a plurality of packets labeled and/or annotated with class labels indicating the developmental parameter. associated with a blastulation indication. In some embodiments, the first machine learning model outputs, with respect to each of the packets, a numerical representations indicating a probability associated with the developmental parameter.

[0077] In some embodiments, at step 220, a second machine learning model is trained using the output of the first machine learning model, wherein the numerical representations of packets are grouped together into sets associated with each individual video segment. In some embodiments, each set is labeled and/or annotated with class labels indicating the developmental parameter.

[0078]   In some embodiments, at an inference step 225, the trained first machine learning model is applied to target packets associated with a target video segment depicting prenatal embryogenesis of a target embryo, to obtain numerical representations for each target packet.

[0079]   In some embodiments, at an inference step 230, the trained second machine learning model is applied to the output of step 225, to predict a developmental potential of said target embryo.

EMBRYONIC DEVELOPMENTAL IMAGE DATASET

[0080]   In some embodiments, machine learning models (e.g., comprising deep neural networks, DNNs) were trained to generate automated and standardized classification algorithms of embryo blastulation and implantation. In some embodiments of the present disclosure, machine learning models were trained directly on the raw video files. In some embodiments, a dataset was assembled and compiled by collecting video depicting at least portion of embryogenesis of files of over 20,000 embryos, cultured in time-lapse-monitored incubators in several medical centers.

[0081]   Reference is made to Figs. 2A and 2B, which show tables of exemplary datasets, in accordance with some embodiments of the present disclosure. In some embodiments, such as depicted by Fig. 2A, the dataset includes thousands of embryos with clinical characteristics and statistical information, wherein TL: Time-lapse incubator, H: Hospital, and KID: Known implantation data.

[0082]   A potential advantage of using retrospective over prospective embryo transfer datasets for machine learning model training is their ethical feasibility and far greater size. However, prediction accuracy is limited due to lacking critical information about endometrial receptivity and using a homogenous dataset of embryos that were retrospectively preselected for transfer according to established morphological and/or morphokinetic criteria. In some embodiments, in order to overcome these limitations, one or more machine learning models may first be trained on blastulation outcome using a heterogeneous set of labeled embryos, and then integrated with one or more machine learning models trained separately on implantation outcome. Accordingly, accuracy of day-

implantation prediction improved over implantation prediction based on implantation outcome alone.

[0083]   In some embodiments, the dataset is arranged in a database with a front-end website that supports display, query and data annotation.

[0084]   In some embodiments, the image dataset comprises anonymized time-lapse video files and the corresponding metadata. In some embodiments, only embryos that were fertilized via intracytoplasmic sperm injection (ICSI) and/or show two-pronuclei appearance (2PNa) inside the incubator are included in the dataset. In this manner, the time of fertilization may be accurately defined, non-fertilized embryos discarded, and a full morphokinetic profile, starting from tPNa, may be obtained. In some embodiments, PGD/PGS tested embryos may be discarded from the dataset as well.

[0085]   In some embodiments, morphokinetic annotations may be performed by qualified and experienced embryologists, adhering to established protocols. In some embodiments, annotation quality assurance (QA) may be carried out by comparing the morphokinetic annotations of 253 randomly selected embryos with the annotations of an expert embryologist in a blinded manner.

IMAGE DATA PREPROCESSING

[0086]   In some embodiments, dataset video segments are preprocessed and/or decreased in size by 16 fold, while the dynamic nature of pre-implantation embryo development is captured and retained in the data. In some embodiments, the preprocessing of the input video files decreases their size from ~100 MB each to less than 1 MB each, while retaining the dynamic nature of pre-implantation embryo development. In some embodiments, the preprocessing comprises 4X-resizing of embryo images (i.e., 2X biaxially) and/or segmentation of the embryo region of interest (ROI).

[0087]   Fig. 3 is an overview of the preprocessing of the input embryo video files, including automated screening of empty well images, cropping and/or segmentation of embryo regions of interest (ROI), and 2X down sampling along each axis. In some embodiments, five consecutive ROIs of the same embryo and/or the same focal plane are grouped into packets.

[0088] Reference is made to Figs. 4A and 4B, which show exemplary automated screenings of empty well images and cropping embryo region of interest (ROI), in accordance with some embodiments of the present disclosure.

[0089] Fig. 4A shows (i) Empty well images detected for a 3-cells embryo input image, (ii) prewitt operators are applied bi-axially to generate a gradient map, wherein the sharp boundaries of the culture wells are then segmented and pixels outside the well are set to zero, (iii) low gradient pixels are set to zero and salt-and-pepper noise and/or impulse noise is removed ($II2$), and (iv) the embryo is identified and cropped as the highest energy object. In some embodiments, the pixels of other objects are set to zero. In some embodiments, images with integrated intensity smaller than threshold are identified as empty wells.

[0090] Fig. 4B shows (i) embryo cropping algorithm is applied only to images that were not identified as empty wells, as demonstrated here for a 2-cells embryo image, (ii) a gradient map is generated using an SSIM descriptor, (iii) next, a 256x256 pixels convolution mask is applied, which highlights high-gradient regions, (iv) based on the texture differences of the cytoplasm and the surrounding zona pellucida compared with the surrounding well, the coordinate of the ROI mask with maximal intensity is identified as the ROI in which the embryo is bounded (cropped ROI).

[0091] In some embodiments, preprocessing comprises at least one of identifying and/or discarding of empty well images, cropping of square regions of interests' (ROIs) that contain the embryos, and down-sampling of cropped embryo ROIs.

[0092] In some embodiments, the image data in which empty wells are depicted is screened manually and/or automatically. In some embodiments, an algorithm determines whether an image contains an embryo or the culture well is empty. In some embodiments, horizontal and vertical 3x3 Prewitt operators are applied on the input image (Fig. 4A-i). In each pixel, the L^2-norm of the absolute values of both channels is calculated to generate a gradient map which is then normalized by the median gradient value. To find the boundaries of the circular culture well, each side of the image is treated separately (left, top, right, bottom) as illustrated below for the top side of the image. In some embodiments, the central fifty columns (columns 226 to 275) are scanned from top to bottom and the first pixel with value greater than 0.4 is marked in each column. In some embodiments, the average y-axis coordinate of

all fifty marked pixels is calculated. By applying these steps also to the bottom, left and right sides, the rectangle in which the culture well is bounded is define. In some embodiments, the boundaries of the well are obtained by fitting a circle inside.

[0093]   In some embodiments, all the pixels that are located outside the well boundaries are set to zero (I_1, Fig. 4A-ii). In some embodiments, low gradient levels are removed by setting low-intensity pixels $I_1 < 0.2$ to zero followed by removing 'salt and pepper' noise using 10x10 convolution mask (I_2, Fig. 4A-iii). In some embodiments, the remaining objects in I_2 are typically larger than the size of the convolution mask. In some embodiments, and under the assumption that the embryo is the largest object with the sharpest edges, all pixels are set to zero, except for the pixels that belong to the highest-energy cluster, namely the object with the maximal integrated intensity in I_2 (I_3, Fig. 4A-iv). In some embodiments, an image is labelled as empty if the sum of I_3 pixels is lower than a threshold value, e.g., 8000.

[0094]   Fig. 4B illustrates an exemplary automated segmentation of an embryo's ROI. In some embodiments, a self-similarity descriptor (SSIM) is applied on the input image (Fig. 4B-i) to generate a gradient map, I_1. In some embodiments, gradients are calculated at each pixel along eight equally-rotated directions (45° apart) at a distance of three pixels away. For each angle, noise is reduced by averaging the gradients over 3x3 regions centered at distal pixels. Next, the L^2-norm of the all eight directional gradients is calculated and the value of each pixel in I_1 is obtained by normalizing by the median SSIM value of the entire image. In some embodiments, the SSIM depicts and highlights the edges of the well and of the embryo (Fig. 4B-ii). In some embodiments, a convolution between I_1 and a 256x256 mask of ones is performed. Since I_1 is a 500x500 matrix, the product of this convolution, I_2, is 245x245 (Fig. 4B-iii). In some embodiments, the location of the region of interest (ROI) in which the embryo is contained is obtained by the argument of maxima (ArgMax) of I_2 (Fig. 4B-iv). In some embodiments, the cropped ROI is down-sampled two-fold biaxially (128x128 pixels).

[0095]   Reference is made to Figs. 5A and 5B, which show exemplary embodiments of image preprocessing comprising automated segmentation, down sampling of embryo region of interest (ROI), and discarding empty well images, in accordance with some embodiments of the present disclosure.

[0096]   Fig. 5A shows segmentation of 2,700 images spanning all developmental states for training a fully automated neural network that was performed using the GrabCut semi-automated algorithm. Embryo framing of a bounding box (marked by the blue box, left) is performed manually followed by an initial segmentation by GrabCut (marked by the green contour). Adjustment of regions to include within the embryo ROI (marked in red) and regions to exclude from the embryo ROI (marked in yellow) are performed manually (for example, as depicted by the embryo image positioned in the middle of the figure). Embryo segmentation is then adjusted accordingly (as marked by green contour). Multiple iterations of embryo segmentation adjustment are allowed. In some embodiments, once embryo segmentation is approved by user, the morphological operations are performed (filling holes, contour smoothing and rendering) and/or a binary mask of the embryo is defined (as depicted by the embryo image positioned on the right and labeled as "corrected segmentation").

[0097]   Fig. 5B shows a U-Net that was trained on the images segmented by GrabCut and a classifier that was executed for performing embryo segmentation of all images in the database. In some embodiments, such as depicted in Fig. 5B-i, representative embryo segmentations are shown at different developmental states. In some embodiments, such as depicted in Fig. 5B-ii embryo segmentation accuracy is reported via the Dice coefficient on 150 test set embryos, showing nearly perfect segmentation across all developmental states. In some embodiments, empty well images are identified based on small mask area.

[0098]   In some embodiments, a large set of validated embryo images is generated and segmented by binary masks of the embryo ROI, which may be used for training a fully automated classifier. In some embodiments, the embryos are framed manually by a bounding box using the semi-automated interactive GrabCut algorithm, allowing for an initial segmentation of the embryo (Fig. 5B). Next, embryo segmentation may be adjusted interactively by scribbling the regions to include within the embryo region of interest (ROI), and regions to exclude from the embryo ROI, allowing fine tuning of embryo segmentation. Multiple segmentation – scribbling – fine-tuning iterations may be used. Once embryo segmentation is approved, morphological operations may be executed (e.g., filling holes, contour smoothing, and dilation) and a binary mask may be defined.

[0099]   In some embodiments, a U-Net classifier is trained using the images that are labeled interactively via GrabCut. In some embodiments, segmented images are resized to 256x256 pixels and divided into train set (2,350 images), validation set (200 images) and uncontaminated test set (150 images). In some embodiments, training is performed using randomly selected 100 images batch size at 1,500 steps per epoch. In some embodiments, each batch of images underwent random augmentation that included one or more of 0° to 180° rotations, horizontal flips, 0 to 0.1 shearing and 0 to 0.1 zooming. In some embodiments, training exactly the same images on different steps was prevented. In some embodiments, network convergence was reached within 20 epochs. In some embodiments, the U-Net classifier obtained embryo images as an input and generated a binary mask output of the embryo ROI resized to 500x500 pixels (Fig. 5B-i). In some embodiments, the accuracy of embryo segmentation was evaluated on test set embryos using the Dice Coefficient, which measures the overlap between the segmented pixels and the binary mask labels. In some embodiments, the dice coefficient ranged between 0.91 and 0.94 across all developmental states (Fig. 5B-ii). In some embodiments, the ROIs of the embryos were evaluated for all images in the database. In some embodiments, empty well images were identified based on small mask area and/or discarded.

[0100]   In some embodiments, all of the segmented embryo frames were resized 2X along each axis into a 128x128 pixel images. In some embodiments, five consecutive resized and segmented frames of the same focal plane and the same embryo are grouped together into packets. In some embodiments, each packet is designated with an identifier P_n^m, including embryo m and first frame time index n. In some embodiments, each packet is associated with a blastulation (BLAST) and/or successful eventual implantation (KID) label y_m. In some embodiments, the packets serve as the input objects of one or more packet learning machine learning models comprising neural networks.

Embryo Blastulation and Implantation Prediction

MACHINE LEARNING MODEL TRAINING

[0101]   In some embodiments, the present disclosure provides for training one or more machine learning models on embryo image data packets, to predict a probability of at least one of blastulation and implantation in a prenatal embryo.

TRAINING SET PREPARATION AND ANNOTATION

[0102]   In some embodiments, one or more training set are constructed to provide for training one or more machine learning models to predict at least one of embryo blastulation and implantation.

[0103]   In some embodiments, the training sets comprise video image data such as a video segments. In some embodiments, every point in time is represented in the video data by one or more frames arranged in frame packets. In some embodiments, every point in time is represented in the video data by one or more frames wherein each frame depicts a specific focal plane. In some embodiments, each packet comprises a group of frames. In some embodiments, each packet comprises a group of frames wherein the frames may or may not comprise different focal planes. In some embodiments, each packet comprises a plurality of frames depicting a plurality of focal points. In some embodiments, each packet comprises at least 3 frames. In some embodiments, the increment of time between a pair of consecutive frames within a packet is equal. In some embodiments, the increment of time is different between every pair of consecutive frames. In some embodiments, frames comprising each packet are temporally sequential and/or consecutive.

BLASTULATION ANNOTATION

[0104]   In some embodiments, a training set associated with blastulation prediction may comprise a plurality of video segment packets, generated as described in detail above, wherein each video packet may be annotated and/or labeled according to a blastulation outcome of the embryo depicted therein, e.g., 'blastulation-positive' (BLAST_p), 'blastulation-negative' (BALST_n), and 'blastulation-unknwon' (BALST_u). The potential

of an embryo to undergo blastulation marks its developmental quality and is linked with its potential to implant in the uterus.

[0105] In some embodiments, embryo annotations and/or labeling may be performed based on a morphokinetic evaluation of an embryo across a plurality of time point and/or windows.

[0106] Reference is made to Fig. 6, which shows a table of exemplary morphokinetic time windows, in accordance with some embodiments of the present disclosure. Fig. 6 shows, based on the statistics of thousands of embryos, the defined time windows for the specified morphokinetic events.

[0107] Reference is made to Fig. 7, which shows a table of exemplary morphokinetic intervals, in accordance with some embodiments of the present disclosure. Fig. 7 shows, based on the statistics of thousands of embryos, the defined time windows that correspond to the intervals between consecutive morphokinetic events.

[0108] Identification of arrested embryos that cannot advance towards blastulation is based on their morphokinetic profiles and an assessment of the time windows for advancing from one morphokinetic event to the next.

[0109] Figs. 8A-8E show exemplary embodiments of identification of developmentally arrested embryos that fail to reach blastulation (BLAST_n), in accordance with some embodiments of the present disclosure.

[0110] Fig. 8A shows embryos plotted according to their incubation time from fertilization (t_{inc}) versus the latest developmental state reached inside the incubator. Embryos are classified based on the temporal distributions of morphokinetic events. For each developmental state from PNa to Morula, BLAST-unknown (BLAST_u) embryos are located within the corresponding regions that are bounded between the 1st percentile of the current event (bottom dashed line) and the 99th percentile of the of the consecutive event (top dashed line). Developmentally-arrested embryos that failed to advance towards the next developmental state within the defined time windows are located in the corresponding regions which are bottom- bounded by the 99th percentile of the consecutive events. Similarly, embryos are plotted according to the time that lapsed from the time of latest

development event $[t_n - t_{n-1}]$ versus the latest morphokinetic state reached. Embryos are classified based on the temporal distributions of morphokinetic intervals. For each developmental state from PNf to Morula, BLAST_u embryos are located within the corresponding regions that are bounded between the 1st (bottom dashed line) and 99th (top dashed line) percentiles of interval distributions $[t_n - t_{n-1}]$. Developmentally-arrested embryos that failed to advance towards the next morphokinetic event within the corresponding interval time windows are identified as 'arrested' either according to (i) the statistics of morphokinetic events, or (ii) the statistics of morphokinetic intervals, and are labeled BLAST-negative (BLAST_n). BLAST-positive (BLAST_p) are embryos that reached start of blastulation inside the incubator. Ladder-like distributions reflect day-night periodicity.

[0111]   Fig. 8B shows video length distributions wherein that half of the embryos in the database were incubated for four days or longer. Fig. 8C show totals in the dataset for BLAST_p, BLAST_n, and BLAST_u embryos, with further breakdowns for developmental states of BLAST_u and BLAST_n embryos. Fig. 8D shows the differences between different medical centers in blastulation outcome statistics originate from the IVF policies (mainly incubation time).

[0112]   Fig. 8E shows embryos plotted according to their incubation time (t_{inc}, y-axis) and the latest developmental state reached (x-axis), as well as according to the time that lapsed from the time of latest development event ($t_{inc} - t_n$, y-axis) and the latest morphokinetic state reached (x-axis). Developmentally arrested embryos are located above the upper dotted line and below the bottom dotted line. Ladder-like distributions reflect day-night periodicity. Embryos that are identified as arrested either according to (i) the statistics of morphokinetic events or (ii) the statistics of morphokinetic intervals are labeled BLAST_n.

[0113]   Fig. 8F shows that half of the embryos were incubated for four days or longer. Fig. 8G shows totals in the dataset for BLAST_p, BLAST_n, and BLAST_u embryos, with further breakdowns for developmental states of BLAST_u and BLAST_n embryos.

[0114]   In some embodiments, metadata used in the annotation process includes, for each embryo, total time of incubation measured from fertilization, t_{inc}, and latest developmental state reached inside the incubator, S_n. In some embodiments, metadata used in the annotation

21

process includes, for each embryo, the time interval between the total time of incubation measured from fertilization and the latest developmental state reached inside the incubator, $t_{inc} - t_n$.

[0115]   In some embodiments, metadata used in the annotation process includes, for each embryo, two Cartesian coordinates, for example, $[S_n, t\_inc]$ and $[S_n, t_{inc} - t_n]$.

[0116]   In some embodiments, to minimize the potential impact of statistical outliers, time windows were bound between the 1st and the 99th percentiles of the corresponding temporal morphokinetic distributions as measured based on thousands of embryos. Therefore, embryos that are located within the bounded area exhibit normal development and thus hold the potential to proceed if incubation was extended. Embryos located in other regions may be indicated as one of: (i) embryos upper-bounded by the 99th percentile of the next consecutive morphokinetic event, S_{n+1}, which reached state S_n and did not proceed to the next developmental state, but are still within the permitted time window for proceeding to S_{n+1}, (ii) embryos that have missed the permitted time window for advancing towards the next developmental state and are arrested in developmental state S_n, and (iii) embryos that have the capacity to advance to the next developmental state if incubation was extended are located within the orange regions, which are bound between the 1st percentile of the temporal distribution of the morphokinetic event S_n and the 99th percentile of the temporal distribution of the consecutive morphokinetic event S_{n+1}. Embryos that are located in the red regions, which are bottom-bounded by the 99th percentile of the consecutive morphokinetic event S_{n+1}. For example, an embryo that was incubated for 96 hours and reached 4-cells state, is arrested and is labeled 4-cells positive 5-cells negative (4Cp5Cn).

[0117]   A similar statistical analysis may be performed to identify developmentally-arrested embryos based on the temporal distributions of the morphokinetic intervals.

[0118]   In some embodiments, a similar statistical analysis may be performed to identify developmentally-arrested embryos based on the temporal distributions of the morphokinetic intervals.

[0119]   In some embodiments, each embryo is represented by the time that lapsed between time of last morphokinetic event and time of incubation, $t_{inc} - t_n$, which is plotted versus the latest developmental state reached inside the incubator, S_n.

[0120]   In some embodiments, embryos may be indicated as one of: (i) embryos which have reached developmental state S_n and still hold the potential to advance to morphokinetic stage S_{n+1} if incubation is extended and/or continued, (ii) embryos that missed their interval time window and are thus arrested in developmental state S_n.

[0121]   For example, an embryo that had completed 4-cells cleavage and that 36 hours have lapsed since without advancing to the next cleavage event is arrested.

[0122]   In some embodiments, embryos are annotated as follows:

- Embryos found to be developmentally arrested either based on their morphokinetic profiles or based on their interval profiles are annotated as BLAST_n;

- Embryos that reached start-of-blastulation (SB) inside the incubator are labeled BLAST_p; and

- All other embryos are annotated as BLAST-unknown (BLAST_u).

[0123]   In some embodiments, labelling of embryo blastulation is based on morphokinetic histories. In some embodiments, embryos that reached start-of-blastulation (tSB) inside the incubator were labeled blastulation-positive (BLAST_p). In some embodiments, BLAST-negative (BLAST_n) embryos that had been arrested at earlier developmental states and BLAST-unknown (BLAST_u) embryos were identified by projecting their morphokinetic profiles onto the time windows that permit transitioning from one embryo state to the next. In some embodiments, the SHIFRA database contains additional metadata, including maternal age, day-of-transfer and co-transferred embryo statistics (Figs. 9A and 9B).

IMPLANTATION AND CLINICAL PREGNANCY ANNOTATION

[0124]   In some embodiments, embryos in the dataset that are known to have reached implantation in the uterus is labeled as 'Known Implantation Data' (KID). KID status is determined based on established protocols by comparing the number of transferred embryos with the number of implanted embryos as determined by the measured number of gestational sacks on week five of pregnancy. In the case that the number of transferred and implanted embryos was equal, the embryos are labelled KID-positive (KID_p). KID-negative (KID_n) corresponds to the case of no implanted embryos. KID-unknown (KID_u) marks embryos

whose implantation outcome cannot be determined, for example when three embryos were transferred and only one or two were implanted. Clinical pregnancy (CP) accounts for the implantation of a viable embryo as determined by fetal heartbeat ultrasound measurement on week 7 of pregnancy. Positive CP (CP_p) accounts for the case that the number of transferred embryos, gestational sacs and fetal heart beats is the same. Negative CP (CP_n) includes transferred embryos that failed to implant (KID_n) and KID_p embryos with no fetal heartbeat.

[0125] Reference is made to Figs. 9A-9C, which show totals and breakdown of an exemplary training dataset comprising embryo video files and associated clinical metadata, in accordance with some embodiments of the present disclosure.

[0126] Fig. 9A shows dataset totals and breakdowns of time-lapse video files of over 20,000 embryos generated by nine time-lapse (TL) incubators located in four medical centers. Clinical metadata includes information on maternal age, day of transfer (7,824 transferred embryos) and number of co-transferred embryos (5,372 transfers). Fig. 9B shows distributions of maternal age, day of transfer and number of co-transferred cycles vary between medical centers, thus reflecting different IVF policies. Fig. 9C shows known implantation data (KID) labels of positive, negative and unknown implantation outcome are presented for embryos in each clinic.

[0127] Reference is made to Figs. 10A-10D. Fig. 10A shows (i) time-lapse image acquisition of preimplantation embryo development as performed at 18-to-20 minute intervals for up to six days of culture inside a time-lapse incubator. At each time point, a z-stack of seven focal planes 15µm apart is recorded.

[0128] Fig. 10B shows high resolution temporal distributions of the morphokinetic events and intervals between consecutive events of positive and negative blastulation-labelled embryos (BLAST_p and BLAST_n). The temporal overlap between BLAST_p and BLAST_n distributions is quantified by K-S distances (top rows).

[0129] Reference is made to Figs. 11A 11B, 11C, 11D, 11E, and 11F, which show exemplary embodiments of high resolution morphokinetic analysis of embryo preimplantation development, in accordance with some embodiments of the present disclosure.

[0130]   Fig. 11A shows morphokinetic annotation performed by trained embryologists. QA validation was performed blindly by an expert embryologist using a set of randomly selected 253 embryos. The temporal differences between the performed annotations and QA morphokinetic annotations were calculated and the corresponding cumulative distributions are presented, showing almost complete agreement with negligible inconsistencies.

[0131]   Fig. 11B shows high temporal resolution distributions of (i) morphokinetic events and (ii) intervals between consecutive events are evaluated based on annotation of thousands of embryos according to established protocols. The morphokinetic events of the first (tPNf to t2), second (t3 to t4) and third (t5 to t8) cleavage cycles appear to be separated in time.

[0132]   Fig. 11C shows the dataset totals and breakdown, and includes information on time-lapse incubator, medical center, maternal age, co-transferred embryo statistics of 3654 transfer cycles, and known implantation data (KID) labels.

[0133]   Fig. 11D shows high temporal resolution distributions of morphokinetic events which were evaluated based on manual annotation of thousands of embryos according to established protocols. The morphokinetic events of the first (tPNf-t2), second (t3-t4) and third (t5 to t8) cleavage cycles are separated in time.

[0134]   Fig. 11E shows the temporal distributions of the intervals between consecutive morphokinetic events show slow transitions between cell cycles.

[0135]   Fig. 11F illustrates the accuracy of morphokinetic annotation was validated by an expert embryologist. The cumulative distributions of the temporal differences show negligible inconsistencies as evaluated using a set of randomly selected 253 embryos.

[0136]   In some embodiments, the dataset comprises embryos associated with maternal patients having a heterogeneous ethnic and racial backgrounds, who span different maternal age groups, thus decreasing the effect of confounding variables and increasing embryo classification generality. Seven-frame z-stacks, 15μm apart, were recorded at 18-to-20-minute intervals for up to six days of incubation, providing a continuous three-dimensional imaging of preimplantation embryo development. In some embodiments, morphokinetic profiles of 16,000 embryos were annotated based on time-lapse imaging. In some embodiments, the annotations specified the time series of discrete events, e.g., pronuclei

appearance and fading (tPNa/f), cleavage of N cells (tN; N=2 to 9), morula compaction (tM) and start of blastulation (tSB). In some embodiments, morphokinetic annotations are determined via majority voting across multiple qualified and trained embryologists according to established protocols, and further validated blindly by an expert embryologist. In some embodiments, the temporal intervals between consecutive morphokinetic events are included, showing temporal separation between cleavage cycles.

[0137]   In some embodiments, known implantation data (KID) labelling is determined based on embryo transfer statistics and/or the number of gestational sacs and fetal heart beats as measured on week 5 to 7 of pregnancy. In some embodiments, positive and negative implantation outcome (KID_p and KID_n) refer to embryos that were successfully implanted or failed to implant, respectively. In some embodiments, embryos whose implantation outcome was uncertain are labelled KID unknown (KID_u).

BLASTULATION AND IMPLANTATION PREDICTION MACHINE LEARNING MODELS

[0138]   In some embodiments, one or more machine learning models may be trained on the training dataset constructed as detailed above, to predict blastulation and/or implantation in video image data associated with embryogenesis of a target embryo.

[0139]   In some embodiments, a fully automated BLAST classifier and/or prediction model of the present disclosure, termed herein 'SHIFRA$_B$, is disclosed. In some embodiments, a fully automated KID classifier and/or prediction model of the present disclosure, termed herein 'SHIFRA$_K$, is disclosed.

[0140]   In some embodiments, SHIFRA$_B$ comprises a single CNN, whereas seven CNN's of the same architecture, training parameters and loss function are applied in parallel to evaluate KID-labeled packets. In some embodiments, three subnetworks are trained on BLAST-labeled embryos and four subnetworks were trained on KID-labeled embryos. As a result, in some embodiments, BLAST-labeled packets obtained a single value score and KID-labeled packets obtained seven scores. In some embodiments, both BLAST- and KID-labeled packets were grouped according to their time index into a series of cohorts that were separated by two hours. In some embodiments, since each cohort included all the packets

from the preceding 12-hours, packets were shared between six successive cohorts. In some embodiments, the seven scores of each KID-labeled packet were integrated into a single value packet score by training a soft support vector machine (SVM) with a linear kernel within each cohort separately.

[0141]   In some embodiments, the implantation outcome of transferred embryos depends not only on their developmental competence, but also on endometrial receptivity, which was not taken into account in the learning process. Therefore, in some embodiments, packet-learning for KID-prediction is performed using an ensemble of three DNNs that are trained on KID-labeled packets and one DNN that is trained on BLAST-labeled packets (four networks in total). In this manner, BLAST prediction packet learning generates one frame score whereas KID prediction packet learning generates four scores for the first image of each packet that are summed into one final frame score.

[0142]   In some embodiments, SHIFRA$_B$ may comprise two stages. In some embodiments, a first stage may be trained on time-lapse video image data packets of embryos, to output a scalar value for each frame in each packet from the start of fertilization to time of prediction (t_p). In some embodiments, a second stage of the BLAST classifier of the present disclosure may be trained on the output of the first stage, to evaluate a potential of an embryo to blastulate, e.g., based on training a Random Forest algorithm

[0143]   In some embodiments, the predictive strength was quantified using the area under the curve (AUC) of the receiver operating characteristic (ROC). In some embodiments, BLAST prediction AUC that was evaluated for test set embryos, increased monotonically with time of prediction, t_p (e.g., 0.65 at 48 hours, 0.73 at 72 hours, 0.88 at 96 hours and 0.94 at 110 hours). In some embodiments, in order to confirm that SHIFRA$_B$ can significantly predict blastulation also of high-quality embryos, the AUC was calculated for a cohort of embryos that reached at least 8-cells cleavage state (8C_p). The BLAST-prediction AUC of 8C_p embryos was lower than for total embryo population, yet it reached 0.63 at 72 hours, 0.84 at 96 hours and 0.91 at 110 hours. In some embodiments, automated BLAST predictions by SHIFRA$_B$ were compared with the a known manual morphokinetic classifier developed by Milewski et al. for five-cells positive (5C_p) embryos (see R. Milewski et al., A predictive

model for blastocyst formation based on morphokinetic parameters in time-lapse monitoring of embryo development. J Assist Reprod Genet 32, 571-579 (2015).

[0144]   Reference is made to Figs 12A, 12B, 12C, and 12D, which show exemplary embodiments of statistical characteristics of BLAST prediction by SHIFRA$_B$, in accordance with some embodiments of the present disclosure.

[0145]   Fig. 12A shows a BLAST prediction by SHIFRA$_B$ (72 hours) for test-set embryos obtained from different medical centers. Fig. 12B shows cross validation predictions, wherein the cross validation comprises leaving one clinic out, and wherein a classifier is trained on embryos only from three clinics and AUC is evaluated on embryos obtained from the fourth clinic. The AUC of clinics H1 and H3, which contribute most of the BLAST-labeled embryos, are relatively low consistent with high demand for additional embryos for training.

[0146]   Fig. 12C shows BLAST prediction using by SHIFRA$_B$ (72 hours) as evaluated for test set embryos of different maternal age groups.

[0147]   Fig. 12D shows five-fold cross validation of embryo-stage learning, wherein a classifier is trained on 80% of the embryos and tested on the remaining 20%, which demonstrates small differences in AUC and is indicative of lack of overfitting and supports generality of BLAST prediction. In some embodiments, the AUC comprises the area under the ROC curve.

[0148]   To test generality, a five-fold stratified cross-validation was performed, yielding an average $AUC = 0.83 \pm 0.02\ STD$. In addition, the BLAST prediction of test set embryos from each clinic were compared separately ($AUC = 0.78 \pm 0.07\ STD$) and performed leave-one-clinic-out cross-validation where classifiers were trained on embryos from three clinics and tested on the forth. Since H3 clinic is the largest data provider of BLAST-labeled embryos, H3 AUC was smaller than the other clinics.

[0149]   Reference is made to Figs 13A and 13B, which show exemplary embodiments of morphokinetic overlaps between different maternal age embryos, in accordance with some embodiments of the present disclosure. Fig. 13A shows high resolution temporal distributions of the morphokinetic events. Fig. 13B shows high resolution temporal

distributions of intervals between consecutive events of embryos obtained from young (age<32) and older (age>38) women are evaluated based on thousands of annotated profiles. Temporal distributions are highly overlapping as quantified by K- S distances (top rows).

[0150]   Temporal comparison between the morphokinetic events and intervals of embryos that were obtained from young (age<32) and older (age>38) women showed negligible differences (KS < 0.08) except for time of start blastulation (tSB), which occurred on average three hours faster, and tM-tSB interval, which was 1.5 hours shorter in embryos derived from young women. To verify that BLAST prediction was not confounded by maternal age, the test set embryos were divided to four age groups and calculated AUC separately for each. In some embodiments, BLAST prediction AUC was comparable between maternal age groups: $AUC = 0.75 \pm 0.03\ STD$. In a clinic, binary classification of embryos can be obtained by setting the threshold values of SHIFRA$_B$ that define negative prediction (below threshold) and positive prediction (above threshold). The retrospective BLAST prediction statistics are presented by setting two threshold values that support positive predictive value 0.91 (PPV2) and sensitivity 0.98. Collectively, the accuracy, robustness and generality of a fully automated day-3 prediction of embryo blastulation were established and clinical utility was retrospectively demonstrated.

[0151]   In some embodiments, KID prediction is required to overcome two major obstacles: (1) Unlike blastulation, which depends on the capacity of the embryo to develop in the incubator under controlled conditions, implantation also depends on endometrial receptivity – a parameter that is not accounted for during training; and (2) Training is limited to embryos that had been preselected for transfer according to existing morphological and/or morphokinetic protocols. As a result, training is restricted to a dataset of morphokinetically-homogenous KID-labeled embryos.

[0152]   Similarly to SHIFRA$_B$, an automated two-stage KID classifier termed SHIFRA$_K$ may be disclosed herein. In some embodiments, SHIFRA$_K$ comprises packet-learning followed by embryo learning. In some embodiments, and in order to improve predictive strength and robustness, packet learning combined three DNNs which are trained separately on KID-labeled embryos, and one DNN that is trained on BLAST-labeled embryos. In some embodiments, the PPV quantifies the probability that embryos that are predicted positive are

indeed positive. In some embodiments, the sensitivity quantifies the fraction of positive-predicted embryos out of all positive-labeled embryos.

[0153]   Reference is made to Figs. 14A, 14B, and 14C, which are exemplary embodiments of automated predictions of embryo implantation, in accordance with some embodiments of the present disclosure.

[0154]   Fig. 14A shows high resolution temporal distributions of the (i) morphokinetic events and (ii) intervals between consecutive events of positive and negative known implantation data-labelled (KID_p and KID_n) embryos are evaluated based on thousands of annotated profiles. KID_p and KID_n distributions are almost indistinguishable as quantified by K-S distances (top rows).

[0155]   Fig. 14B shows the AUC of automated KID prediction by $SHIFRA_K$ of Day-5 transferred embryos (n=359; 355 transferred blastocysts) increases with prediction time, t_p. Inset: ROC curves of automated KID prediction by $SHIFRA_K$ at 68 hours (left) and at 110 hours (right) are compared with implantation prediction by KIDScore-D3 and KIDScore-D5 manual-morphokinetic algorithms, respectively.

[0156]   Fig. 14C shows confusion matrices demonstrating KID prediction based on retrospective outcome of embryo implantation. In some embodiments, binary classifiers that optimize (i) the positive predictive value (PPV); and/or (ii) the sensitivity were generated by setting $SHIFRA_K$ threshold values.

[0157]   Accordingly, the temporal distributions of morphokinetic events and intervals of KID_p and KID_n embryos are almost fully overlapping. The dataset of over 5,500 KID-labelled embryos was divided into a train-validation set and an uncontaminated test set that consisted of randomly selected 21% of the embryos.

[0158]   Reference is made to Figs. 15A and 15B, which are exemplary embodiments of $SHIFRA_B$ and $SHIFRA_K$ optimization for prediction of embryo blastulation and implantation, in accordance with some embodiments of the present disclosure.

[0159]   Fig. 15A shows scores of KID prediction (110 hours) and BLAST prediction (72 hours) of BLAST_p- KID_n (BpKn) and BLAST_p-KID_p (BpKp) co-labeled embryos (n=396) are weakly-positively correlated (Spearman correlation 0.3). Consistent with their

blastulation labels, BLAST prediction histograms of BpKn and BpKp embryos overlap (top panel) and their average scores are comparable (b-left). However, BpKp embryos are scored 40% higher than BpKn embryos only by $SHIFRA_K$ consistent with their implantation labels (b-right).

[0160]   Fig. 15B shows the temporal features of $SHIFRA_B$ (72 hours) and $SHIFRA_K$ (110 hours) are ranked according to their mean adjusted SHapley Additive exPlantions (SHAP), which scores feature contribution to accurate BLAST and KID prediction. The ten top-ranked BLAST features (> 0.004) and the thirteen top-ranked KID features (>0.007) were derived from the latest time-lapse images (BLAST > 66 hours and KID > 100 hours). SHAP values and feature values of the top ranked temporal features were calculated for BLAST-labeled and KID-labeled train set embryos. ROC curves and AUC obtained by (i) BLAST and (ii) KID classifiers that were trained only on the top ten BLAST features and the top thirteen KID features, are shown respectively.

[0161]   Aneuploidy might impair embryo implantation but permit blastocyst formation. To study the relationship between blastulation and implantation potentials, the classification of BLAST and KID co-labeled embryos was analyzed: 121 BLAST_p – KID_n (B_pK_n) embryos and 275 BLAST_p – KID_p (BpKp) embryos. BLAST and KID prediction scores are weakly-positively correlated (Pearson correlation score 0.3), which is indicative of common visual elements. Consistent with their BLAST_p labels, BLAST prediction score distributions of BpKp and BpKp embryos overlap. In some embodiments, average implantation score of the latter embryos was 40% higher than BpKp embryos, indicating that $SHIFRA_K$ differentiates between blastocysts that have the capacity to implant and blastocysts that don't.

[0162]   Reference is made to Figs. 16A 16B and 16C, which are exemplary embodiments of $SHIFRA_B$ and $SHIFRA_K$ optimization for prediction of embryo blastulation and implantation, in accordance with some embodiments of the present disclosure. Figs. 16A, 16B, and 16C show that the temporal features of $SHIFRA_B$ (72 hours) and $SHIFRA_K$ (110 hours) are ranked according to their mean adjusted SHapley Additive exPlantions (SHAP), which scores feature contribution to accurate BLAST and KID prediction. In some embodiments, the ten top-ranked BLAST features (> 0.004) and (Fig. 16B-ii) the thirteen

top-ranked KID features (> 0.007) were derived from the latest time-lapse images (BLAST > 66 hours and KID > 100 hours; color coded). In some embodiments, the SHAP values and feature values of the top ranked temporal features were calculated for (Fig. 16A-ii) BLAST-labeled and (Fig. 16B-ii) KID- labeled train set embryos. Fig. 16C shows ROC curves and AUC obtained by (i) BLAST and (ii) KID classifiers that were trained only on the top ten BLAST features and the top thirteen KID features, respectively.

[0163] In some embodiments, and in order to gain additional insight into the underlying mechanisms of embryo classification, the visual information that is embedded within the time-lapse images, which SHIFRA$_B$ and SHIFRA$_K$ are sensitive to, is analyzed. To this end, SHAP methodology is employed for quantifying the impact of the temporal features on embryo prediction are identified. In some embodiments, the frames that contribute the most to accurate embryo classification by setting the sign of the SHAP values of each feature according to the BLAST or KID label of the embryos (negative: -1; positive: +1) and averaged across embryos (mean adjusted SHAP). In this manner, positive (negative) SHAP values of temporal features of positively (negatively) labelled embryos contribute oppositely to the mean adjusted SHAP compared with positive (negative) SHAP values of temporal features of negatively (positively) labelled embryos. In some embodiments, the contribution of the temporal features to BLAST prediction and to KID prediction changes greatly between different frames.

[0164] In some embodiments, temporal features of low mean adjusted SHAP scores are associated with early time points. In some embodiments, the top ten BLAST features (adjusted SHAP > 0.004) and top thirteen KID features (adjusted SHAP > 0.007) were associated with latest frames. In some embodiments, the SHAP values and the feature values of the top-ranked BLAST and KID temporal features were correlated across individual embryos (Figs. 16A-ii and 16B-ii). In some embodiments, and in order to study whether the top-ranked temporal features can direct BLAST and KID prediction without including the rest of the temporal features, the embryo-learning BLAST and KID classifiers were trained again using only the top-ranked features. The results were that the top-ranked features were sufficient for reaching high predictive power with comparable AUC values (BLAST: AUC=0.75; KID: 0.7 Fig. 7C-i,ii) as SHIFRA$_B$ (AUC=0.74, Fig. 10C-inset) and SHIFRA$_K$

(AUC=0.71, Fig. 14B-inset). This indicates that the top-ranked temporal features as defined here mark the developmental potential of the embryos to blastulate and to implant.

[0165]   In some embodiments, each of the one or more machine learning models comprises one or more neural networks. In some embodiments, a neural network of the present disclosure may be implemented using the PyTorch framework, and/or trained using Stochastic Gradient Descent with Nesterov of 0.9.

[0166]   In some embodiments, the input objects of the network are packets of five pre-processed frames P_n^m. In some embodiments, training batches include randomly selected $k = 4$ packets obtained from $K = 8$ embryos within 12-hour windows, 20 hour time windows, and/or any range therebetween. In some embodiments, the packets are obtained from randomly selected embryos such that pairs of packets are separated by no longer than 8 hours. In some embodiments, the training batches are Convolutional Neural Network (CNN) training batches. In some embodiments, the training comprises CNN training.

[0167]   In some embodiments, packets of all focal planes are used for training. In some embodiments, packets of the three central focal planes (-15, 0 and +15 μm) are used for training, whereas validation and test set embryos contributed packets only of the middle focal plane. In some embodiments, the residual network architecture comprises thirteen layers that include seven residual blocks and two fully-connected layers. In some embodiments, the layers comprise one or more of a convolution-max pooling layer, residual-max pooling convolution blocks, residual convolution blocks, fully-connected layers, and one or more input neurons that are associated with non-overlapping time windows. In some embodiments, the time windows comprise 6-19 hour time windows, for example, such as 12-hour time windows. In some embodiments, the duration of the time windows ranges between 48 and 120 hours. In some embodiments, packets representing a time point earlier than 48 hours are associated with the first neuron and packets later than 120 hours are associated with the sixth neuron. In some embodiments, the packet score is selected out of the input neurons according to its time index.

[0168]   In some embodiments, the last layer consists of w input neurons that are associated with non-overlapping time windows as defined for BLAST and KID prediction networks. In

some embodiments, packet score output is determined by selection of one of the input neurons according to the time index n of the first frame of the packet.

[0169] In some embodiments, embryo developmental potential is marked by scarce dynamic events that last 30 to 60 min and are thus captured by individual packets. In some embodiments, a high-quality embryo will have only a few packets that are scored high whereas all packets of a low-quality embryo will be scored low. This principle is implemented by weighing logistic loss as follows. In some embodiments, the weighted loss of embryo m is calculated based on all k packets:

$$l^m = \sum_{n=1}^{k} w_n^m Log\left(1 + e^{-y_m \cdot s_n^m}\right),$$

where s_n^m are the packet scores and w_n^m are the softmax weights:

$$w_n^m = \frac{e^{\gamma * s_n^m}}{\sum_{n=1}^{k} e^{-\gamma \cdot s_n^m}}$$

[0170] In some embodiments, the sum of the weights w_n^m across k packets of embryo m is 1. In some embodiments, γ is the softness parameter. In some embodiments, at the limit $\gamma \rightarrow 0$, the weight becomes $\frac{1}{k}$ independent of the scores of the packets. In some embodiments, at the opposite limit of large γ, w_n^m approaches 1 only for the packet of maximal score. The problem of approaching this limit is that it will be increasingly difficult for the network to converge. In some embodiments, the batch loss L is:

$$L = \frac{1}{K} \sum_{m=1}^{K} l^m$$

[0171] The weights are thus optimized to minimize L. In some embodiments, the weights are DNN weights. In some embodiments, the weights are CNN weights. In some embodiments, the performances are optimized by setting the value of γ. For example, in some embodiments, γ is set as $\gamma = 3$. For example, in some embodiments, γ is set as $\gamma = 5$. In some embodiments, for a negative labeled embryo ($y_i = -1$), even if a single packet will obtain a positive score, its weight w_n^m will be highest and the loss of the embryo l^m will be large. As a result, convergence will be approached only if all packets of a negative embryo will obtain negative scores. On the other hand, one packet with a high positive score is

34

sufficient for obtaining a small loss for a positive embryo ($y_i = +1$). For these embryos, the packets with low scores will have small weights and the packet with the highest score will have the highest weight and a small loss will be obtained. In some embodiments, CNN training by BLAST-labeled and by KID-labeled embryos converged within ten epochs.

MACHINE LEARNING MODEL OUTPUT

[0172] In some embodiments an output of the trained machine learning models, e.g., SHIFRA$_B$ and SHIFRA$_K$, comprises a scalar and/or compact and/or reduced-dimensionality representation of each video image data packet. In some embodiments, the output is a single number and/or a vector representing each packet.

[0173] In some embodiments, and as described in greater detail elsewhere herein, the modified soft hinge loss and/or the batch loss are calculated for a plurality of packets. In some embodiments, the training comprises calculating the loss for all the packets together.

[0174] In some embodiments, the scalar representation is a vector. In some embodiments, the conversion comprises labeling the morphokinetic state of each packet. In some embodiments, the conversion comprises labeling the number of cells in each frame. In some embodiments, the conversion comprises generating a vector comprising n characters. In some embodiments, a vector comprises a character for each morphological state. In some embodiments, each character comprises a score between 0 and 1 for each cell. In some embodiments, the score of each character indicates the association of the packet with a specific classification.

[0175] In some embodiments, the vector represents a predictive probability vector. In some embodiments, for an embryo depicted by the packet, each position of a character along the vector indicates the probability of the embryo to be in a specific state in that position along the vector. In some embodiments, the specific state comprises the state of the embryo for a specific position within the vector. In some embodiments, the vector can be converted into a single number that represents the morphokinetic state of the embryo as a function of time.

[0176] In some embodiments, the conversion of the image data to compact representation includes using an autoencoder. In some embodiments, the conversion comprises training a network to compress an image to compact representation. In some embodiments, the

conversion comprises training a network to decompress the compact representation to restore the image data.

[0177]   In some embodiments, the vector is converted to a matrix. For example, an image of 500 by 500 can be compressed to 100 by 1, and thereby stored as a column of a compact-representation matrix. In some embodiments, the compact representation matrix comprises a plurality of columns wherein each column comprises a two-dimensional compact representation of an image.

[0178]   In some embodiments, the embryo depicted by the compact representation is evaluated. In some embodiments, the compact representation of each embryo comprises a plurality of vectors. For example, in some embodiments the compact representation of each embryo comprises 6 vectors: 3 vectors are fed into a machine learning model trained to identify blastulation packets and 3 vectors are fed into a machine learning model trained to identify implantation packets. In some embodiments, the plurality of vectors are summed together into a single vector which is evaluated for blastulation and/or implantation potential of the represented embryo.

[0179]   In some embodiments, the compact representation comprises a matrix which is applied to a machine learning model trained to evaluate the blastulation and/or implantation potential of the embryo represented by the matrix.

[0180]   In some embodiments, packet learning for BLAST prediction and for KID prediction is performed using the same DNNs as described above. In some embodiments, network training for BLAST prediction is performed using $w = 25$ non-overlapping time windows (0-4; 4-8; 8-12; 12-16; 16-20; 20-24; 24-27; 27-30; 30-5 33; 33-36; 36-42; 42-46; 46-48; 48-51; 51-56; 56-64; 64-72; 72-76; 76-80; 80-85; 85-90; 90-95; 95-105; 105-115; >115). In some embodiments, network training for KID prediction is performed using $w = 16$ non-overlapping time windows (0-12; 12-24; 24-30; 30-36; 36-42; 42-48; 48-56; 56-64; 64-72; 72-76; 76-80; 80-85; 85-95; 95-115; 115-120; >120). In some embodiments, DNN training using BLAST-labeled and KID- labeled embryos typically converged within 20-to-60 epochs.

[0181]   Fig. 17 shows the input of the deep neural network m are the preprocessed packets P_n^m of embryo m and time index n. The DNN consists of 13 layers. There are m input

neurons to the last layer, which determine the time windows for each packet according to its time index n. The output scalar neuron of the network is the packet score. In some embodiments, conv comprises Convolution. In some embodiments, RCB comprises Residual convolution block.

[0182] In some embodiments, the developmental competence of each embryo was evaluated based on the scores of the packets that belong to it. Consistent with the postulation that high-developmental competence is marked by rare visual features, threshold values were applied across all BLAST- and KID-labeled packet cohorts, thus removing low-score noisy packets and highlighting high-score packets. One hundred equally-separated threshold values were defined for each 12-hours packet cohort, ranging between the lowest and the highest train-set packet scores. For each threshold value, embryo scores were calculated based on the validation-set packets that belong to it as follows. Packets of lower scores were discarded and packets of higher scores were summed after threshold-subtraction. For each time of prediction, the selected threshold value generated the maximal AUC as calculated across the validation set embryos.

EMBRYO DEVELOPMENTAL COMPETENCE PREDICTION

[0183] In some embodiments, the developmental potential of embryos is scored using a second trained classifier. Each embryo is represented by a vector of frame scores obtained by packet learning. In some embodiments, the temporal features were generated by interpolation of the vectors of frame scores, thus obtaining a synchronized representation of all embryos. In some embodiments, different classifiers are trained independently in order to allow embryo prediction at different time points (time of prediction, t_p). In some embodiments, a classifier is trained on all train-set labeled embryos of video length greater than a given t_p using the temporal features earlier than t_p. In some embodiments, the BLAST prediction is performed using a Random Forest classifier. In some embodiments, the KID prediction was performed using logistic regression. In both cases, training parameters were optimized via grid-search five-fold cross validation.

EXPERIMENTAL RESULTS

[0184]   The ambiguity in identifying the actual visual elements that direct neural network prediction is one of the major drawbacks in deep learning. In some embodiments, in order to obtain a mechanistic insight into how embryo prediction by SHIFRA$_B$ and SHIFRA$_K$ work, the top-ten positive SHAP images and top-ten negative SHAP images for each of the top-ranked BLAST and KID temporal features are present. In some embodiments, embryos are marked such that the cleavage-stage embryos with 4 cells or less, embryos with asymmetric blastomeres, and highly-fragmented embryos in the images that contributed the most to positive and negative BLAST prediction by SHIFRA$_B$. In some embodiments, temporal features 1 (72 hours), 2 (70 hours) and 4 (71 hours) were most sensitive to 4-cells cleavage stage embryos and were identified as SHAP-negative images. In some embodiments, temporal features 1, 3 (71.7 hours) and 8 (66.3 hours) were most sensitive to uneven blastomere size and temporal features 2 and 4 were most sensitive to embryo fragmentation, which also obtained negative-SHAP scores.

[0185]   In some embodiments, the cleavage and morula stage embryos were marked with non-compacted blastomeres and compacted morulae in the images that contributed the most to positive and negative KID prediction by SHIFRA$_K$. In some embodiments, the temporal features 10 (106.7 hours) and 13 (107.3 hours) were most sensitive to the appearance of non-compacted blastometres and temporal features 3 (109.7 hours), 9 (100.3 hours) and 11 (103.7 hours) were most sensitive to morula. In both cases, these morphological characteristics directed negative-SHAP KID prediction. In some embodiments, the morphokinetic and morphological characteristics are depicted only by a small fraction of the images of the temporal features, indicating that the determinant visual elements that direct BLAST and KID prediction are not distinguished by human level perception.

[0186]   Reference is made to Figs. 18A and 18B, which show exemplary embodiments of top positive versus negative SHAP-scored embryo frames directing BLAST prediction, in accordance with some embodiments of the present disclosure.

[0187]   Figs. 18A and 18B shows Top ten SHAP-positive versus top ten SHAP-negative embryo frames are shown for the selected temporal features by mean adjusted SHAP for SHIFRA$_B$ BLAST prediction at 72 hours. In some embodiments, cleavage stage embryos

with no more than 4 cells , blastomere asymmetry and high fragmentation embryos are marked.

[0188]   Reference is made to Figs. 19A and 19B, which show exemplary embodiments of top positive versus negative SHAP-scored embryo frames directing KID prediction, in accordance with some embodiments of the present disclosure.

[0189]   Figs. 19A and 19B show top ten SHAP-positive versus top ten SHAP-negative embryo frames are shown for the selected temporal features by mean adjusted SHAP for SHIFRA$_K$ KID prediction at 110 hours. In some embodiments, cleavage stage or murula stage embryos exposing non-compacted blastomeres and morula stage embryos are marked.

[0190]   Fig. 10C shows the area-under-curve (AUC) of automated blastulation prediction by SHIFRA$_B$ of test set embryos (n=1,621) and high quality embryos that reached 8-cells (8C_p; n=1,456) with video length $\geq$ 110 hours increase monotonically with time of prediction, t_p. Left inset: Receiver operating characteristic (ROC) curves and AUC values calculated at $t_p = 72$ $hours$ of test set embryos and 8C_p embryos with sufficiently long videos. Right inset: A comparison between SHIFRA$_K$ and Milewski BLAST classifier at $t_p = 86$ $hours$ (8). Fig. 10D shows confusion matrices demonstrating retrospective BLAST prediction. Binary classifiers were generated by setting SHIFRA$_B$ threshold values to optimize (i) PPV and (ii) sensitivity, wherein K-S: Kolmogorov Smirnov, ROC: Receiver operating characteristic, AUC: Area under the (ROC) curve, and PPV: positive predictive value.

[0191]   Fig. 20A shows fully automated prediction of embryo blastulation by SHIFRA$_B$ that is performed directly on the time-lapse video files of the embryos. With prediction time, the AUC is monotonically increasing as the visual information encoded in the video gains association with embryo quality. Fig. 20B shows the ROC curves of BLAST prediction at 72 hours of (i) all test set embryos (n=852), (ii) high quality embryos that reached 8cells (8Cp, n=775), and (iii) very high quality embryos that reached morula compaction (MORULAp, n=730).

[0192]   IVF cycles typically consist of multiple fertilized oocytes whose number and quality tend to decline at advanced maternal age. The decision which embryo(s) to transfer

and when is critically important for reaching live birth while minimizing health risks and shortening time to pregnancy by decreasing the number of cycles. During incubation, each embryo can either be transferred, discarded, further cultured or frozen as a reserve for subsequent transfers. This complex decision-making process requires a comprehensive strategy that optimizes implantation potential and takes into account outcomes of previous cycles, maternal age, sperm quality, and clinical background. The present disclosure comprises scoring the developmental potential of individual embryos to form a blastocyst and implant in the uterus. Video files of over 11,000 embryos were collected, wherein the embryos were cultured in nine incubators and located in five medical centers during the past five years. In some embodiments, seven-frame Z-stacks, 15 μm apart, were recorded at ~20 minute intervals for up to six days of incubation, thus providing a continuous three-dimensional imaging of preimplantation embryo development. In addition to the video files, the SHIFRA database includes metadata for each embryo, including maternal age, co-transferred embryo statistics, and implantation outcome (Figs. 11A-11F). In some embodiments, morphokinetic events of all embryos were annotated by expert embryologists in accordance with established protocols and validated via a quality assurance protocol.

[0193]   Reference is made to Figs. 21A and 21B, which shows exemplary embodiments of first direct unequal cleavage (DUC1) embryos that are selected by SHIFRA$_K$, in accordance with some embodiments of the present disclosure. Fig. 21A shows DUC1 referring to the abnormal division of the first cell into three blastomeres leading to chromosomal aberrations. Fig. 21B shows DUC1 embryos that are scored low relative to non-DUC1 embryos (both KID_n and KID_p embryos), and are thus deselected by SHIFRA$_K$ at 90 hours. In some embodiments, analysis included all DUC1 embryos in the SHIFRA database compared with test-set non-DUC1 embryos in order to improve statistical significance. Out of 1,131 KID_p embryos in the SHIFRA database, only three were DUC1 embryos and only one was incubated for over 90 hours.

[0194]   In some embodiments, the temporal distributions of the morphokinetic events and the consecutive intervals between them were evaluated based on the profiles of thousands of embryos. In some embodiments, the small fraction of direct unequal cleavage (DUC) embryos was identified and labelled as well. In some embodiments, the embryos were divided into train, validation and test sets. In some embodiments, test sets comprise of

randomly selected 18% of all embryos. In some embodiments, test sets were strictly maintained uncontaminated and were used only after training processes were completed. In some embodiments, the capacity of an embryo to undergo blastulation marks its developmental quality and is linked with its potential to implant in the uterus. Hence, classification of BLAST-positive (BLAST_p) embryos that can reach start-of-blastulation (SB) and BLAST-negative (BLAST_n) embryos that are developmentally arrested is of high clinical value (Figs. 8A-8G). A comparison between the temporal distributions of BLAST_p and BLAST_n embryos shows high overlap across all morphokinetic events and intervals.

[0195] Reference is made to Figs. 22A, 22B, and 22C, which show statistical characteristics of BLAST prediction by SHIFRA$_B$, in accordance with some embodiments of the present disclosure. Fig. 22A shows the BLAST prediction ROC curves at 72 hours of the train, in which the validation and test sets are overlapping, thus excluding overfitting by SHIFRA$_B$. Fig. 22B shows positive and negative predictive values (PPV and NPV) that are plotted as a function of classification threshold at 72 hours. Fig. 22C demonstrates binary prediction of embryo blastulation by SHIFRA$_B$, where a threshold value of 42 was set such that embryos with scores higher than the threshold are classified positive. At this setting, PPV=0.91 and NPV=0.31.

[0196] In some embodiments, a deep neural network was trained directly on the video files using BLAST-labelled embryos, which is not limited to a morphokinetic representation. In some embodiments, a fully automated classifier SHIFRA$_B$ was generated, which predicts embryo blastulation. A detailed description of the CNN design (Figs. 3 and Fig. 17) and the learning process is provided in the present disclosure. In some embodiments, the overlap between the receiver operating characteristic (ROC) curves of the train, validation and test set embryos excludes overfitting.

[0197] In some embodiments, the BLAST-prediction, which is measured by the area under the ROC curve (AUC) for test-set embryos with sufficiently long video recordings, increases monotonically from 0.68 at 60 hours and 0.75 at 72 hours to 0.87 at 90 hours. In some embodiments, the BLAST prediction is also demonstrated for high-quality embryos that reached 8-cells state (AUC=0.65) and very- high quality embryos that reached morula compaction (AUC=0.61), demonstrating that SHIFRA$_B$ can discriminate between morulae

that can form a blastocyst and developmentally arrested morulae. In some embodiments, a binary classifier that supports SET methodology is derived by selecting a threshold value 42, which optimizes positive predictive value (PPV, Fig. 22B). In some embodiments, 399 embryos are classified positive and 469 embryos are classified negative, with 91% precision.

[0198] Reference is made to Figs. 23A, 23B, and 23C, which are exemplary demonstrations of fully automated predictions of embryo implantation (SHIFRA$_K$), in accordance with some embodiments of the present disclosure. Fig. 23A shows the temporal distributions of (i) the morphokinetic events and (ii) the time intervals between consecutive events are evaluated based on thousands of annotated profiles, revealing significant overlaps between positive and negative known implantation data labelled embryos (KID_p and KID_n; KS test scores < 0.34, top rows). Fig. 23B shows fully automated prediction of embryo implantation by SHIFRA$_K$ is performed directly on the time-lapse video files of the embryos at time of prediction. The drop in the AUC at 80-82 hours is due to a decrease in the number of transferred embryos beyond day-3 that are available for CNN training. In some embodiments, the KID prediction by SHIFRA$_K$ is superior compared KIDScore-D3 morphokinetic classifier. Fig. 23C shows the ROC curves of KID prediction by SHIFRA$_K$ at (i) 60 hours (n=626), (ii) 74 hours (n=274) and (iii) 90 hours (n=204).

[0199] Reference is made to Figs. 24A, 24B, 24C, 24D, and 24E, which show the SHIFRA$_B$ and SHIFRA$_K$ databases are optimized to predict blastulation and implantation with partial dependence on embryo state, in accordance with some embodiments of the present disclosure.

[0200] Fig. 24A shows the prediction scores of embryo implantation (SHIFRA$_K$), which are linearly correlated with the prediction scores of blastulation (SHIFRA$_B$) as computed for co-labeled embryos at 72 hours (Spearman correlation coefficient 0.77). The BLAST_n-KID_n (BnKn) embryos are scored low by both classifiers. Relative to BLAST_p-KID_n (BpKn) embryos, the distribution of BLAST_p-KID_p (BpKp) embryos is shifted towards higher implantation scores (right panel) but not towards higher blastulation scores (top panel).

[0201] Fig. 24B shows that the average prediction scores yield BnKn< BpKn< BpKp by SHIFRA$_B$ and by SHIFRA$_K$, but BpKn< BpKp is statistically significant only by SHIFRA$_K$,

indicating that SHIFRA$_K$ is capable of correctly evaluating the potential even of blastocysts to implant. Fig. 24C shows that the probability distributions of the developmental states of positive and negative (i) BLAST-labeled and (ii) KID-labeled embryos are shown between 60 and 90 hours from fertilization.

[0202]   Fig. 24D shows that the information entropies, $H_B(t) > H_K(t)$, as calculated for the embryonic states of BLAST and KID labeled embryos, vary little between 60 and 90 hours. Fig. 24E shows the adjusted mutual information (AMI) between the distributions of SHIFRA$_B$ and SHIFRA$_K$ classification scores, and the distributions of states of BLAST-labeled and KID-labeled embryos are plotted as a function of prediction time. As a control, the AMI calculated relative to a random distribution of states are zero.

[0203]   Reference is made to Figs. 25A and 25B, which are exemplary embryonic state distributions statistical analysis, in accordance with some embodiments of the present disclosure.

[0204]   Fig. 25A shows the probability distributions of the developmental states of all test set embryos are shown between 60 and 90 hours from fertilization. Fig. 25B shows the information entropy $H(t)$, as calculated for the embryonic state distributions, varies little between 60 and 90 hours. Hence, the effective number of states that the embryos are found in, estimated by $2^{H(t)}$, is maintained constant.

[0205]   Compared with BLAST classification, prediction of known implantation data (KID) is inherently hindered due to two reasons: (1) Unlike blastulation, which depends on the capacity of the embryo to develop in the incubator under controlled conditions, implantation also depends on uterus receptivity, which is not taken into account in the training process, and (2) KID-labelled embryos had been preselected for transfer according to morphological and/or morphokinetic parameters. This is reflected by the high temporal overlap between the distributions of the morphokinetic events and intervals of KID-positive (KID_p) and KID-negative (KID_n) embryos. Hence, the embryos available for training of KID prediction form a homogenous group compared with the morphokinetic variation that is characteristic of BLAST-labeled embryos (Figs. 10A-10D, Figs. 20A-20B, and Fig. 24C-24D) and the entire population of embryos (Figs. 10A-10F and Figs. 25A-25B).

[0206] Reference is made to Figs. 26A and 26B, which are exemplary statistical characteristics of KID prediction by SHIFRA$_K$, in accordance with some embodiments of the present disclosure. Positive and negative predictive values (PPV and NPV) are plotted as a function of SHIFRA$_K$ classification threshold at (depicted by Fig. 26A-i) 72 and at (depicted by Fig. 26B-i) 90 hours. Binary classification of embryo implantation is demonstrated by setting threshold values (depicted by Fig. 26A-ii) 80 at 72 hours, and (depicted by Fig. 26B-ii) 115 at 90 hours. Embryos with scores higher than thresholds are classified positive. At these settings, (depicted by Fig. 26A-iii) PPV=0.64 and NPV=0.72 at 72 hours, and (depicted by Fig. 26B-iii) PPV=0.81 and NPV=0.72 at 90 hours.

[0207] Reference is made to Fig. 27, which is a table of the KIDScore-D3, in accordance with some embodiments of the present disclosure. Fig. 27 shows the sensitivity and specificity of KIDScore-D3 as calculated for the test set KID-labelled embryos.

[0208] To overcome these limitations, the KID-classifier SHIFRA$_K$ integrates multiple subnetworks that were trained in parallel on BLAST-labeled and on KID-labeled embryos. KID prediction is comparable when evaluated on the train, validation and test sets, both at 72 hours and at 90 hours' prediction time, thus excluding overfitting. In some embodiments, the AUC increased with prediction time from 0.65 at 60 hours, 0.70 at 74 hours and 0.74 at 90 hours (Figs. 23B and 23C). In some embodiments, at 66 hours, KID prediction by SHIFRA$_K$ outperforms the current state-of-the-art morphokinetic classifier KIDScore-D3.

[0209] Reference is made to Figs. 28A and 28B, which are exemplary embryo transfer and implantation statistics, in accordance with some embodiments of the present disclosure. Fig. 28A shows a histogram of the time of embryo transfer shows a separation between embryos that were transferred on day-2, day-3, day-4 and day-5. Embryos from all 3654 transfer cycles from five hospitals are included. Fig. 28B shows that while (i) the number of BLAST-labeled embryos is remains almost constant until day-5, (ii) the number of KID-labeled embryos drops sequentially on days 2 to 5 due to embryo transfers.

[0210] The decrease in AUC at 80-82 hours is due to the reduction in the number of transferred embryos and specifically KID-labeled embryos that were available for training. In some embodiments, the high AUC at 90 hours is consistent with a high PPV compared with 72 hours whereas NPV is the same. Similar to BLAST prediction, a binary KID

classifier that supports SET methodology was derive by selecting threshold 80 at 72 hours and 115 at 90 hours. PPV increased from 0.64 (72 hours) to 0.81 (90 hours) while NPV was 0.72.

[0211]   In some embodiments, direct unequal cleavage (DUC) embryos were not removed from the training, validation and test sets, in order to maintain full automation of KID prediction by SHIFRA$_K$. Compared with late DUC events, the first DUC (DUC1) is associated with the most significant decrease in embryo blastulation, implantation and euploid rate, and holds the most far-reaching clinical implications. DUC1 embryos were deselected for transfer: the prevalence of DUC1 embryos is 4.7% in the SHIFRA database but only 2.1% of transferred embryos are DUC1. The implantation rate of DUC1 and non-DUC1 transferred embryos was 15% (17 embryos) and 37%, respectively. Satisfyingly, SHIFRA$_K$ scores both KID_n and KID_p DUC1 embryos significantly lower than non-DUC1 embryos. In fact, the KID_p DUC1 embryo is scored lower than KID_n non-DUC1 embryos, indicating that DUC1-embryos are strongly deselected by SHIFRAK. Since KID prediction is evaluated based on the image packets of the 12-hours preceding prediction time, deselection of DUC1 embryos by SHIFRAK on day-3 or later is not based on the images of first and second cleavage events, but rather on visual features that propagated forward.

[0212]   Reference is made to Figs. 29A, 29B, 29C, and 29D, which shows that classification of embryo blastulation and implantation is robust to differences in maternal age, in accordance with some embodiments of the present disclosure. Fig. 29A shows a comparison between the temporal distributions of all morphokinetic events of embryos of young (age<32) and older (age>38) women as evaluated based on thousands of annotated profiles shows very high overlaps (KS test scores < 0.13, top row). On average, only Morula compaction (tM) and start-of-blastulation (tSB) appear earlier in embryos obtained from young women.

[0213]   Fig. 29B shows the fraction of successfully implanted embryos obtained from young women (722 out of 1702 transferred embryos) is 2.7 fold greater than embryos obtained from older women (351 out of 2249 transferred embryos). Fig. 29C and Fig. 29D show the ROC curves of BLAST prediction and KID prediction, respectively, at 72 hours are plotted for embryos obtained from young versus older women. The AUC (legends) of

embryos of young women is comparable with the AUC of the entire pool of embryos, yet the AUC of embryos from older women is higher, indicating that both SHIFRA$_B$ and SHIFRA$_K$, which were trained on the entire pool of embryos, detect visual features that are associated with advanced maternal age.

[0214] In some embodiments, the KID predictive strength increased with time of prediction t_p, as evaluated for the same cohort of day-5 transferred test-set embryos. In some embodiments, the AUC increases slowly from 48 to 84 hours and more rapidly from 84 hours onward. Compared with the manual-morphokinetic KIDScore™ decision support tools, embryo implantation prediction by SHIFRA$_K$ is as accurate as KIDScore-D3 on day-3 and more accurate than KIDScore-D5 on day-5 as evaluated for the same test-set embryos. In some embodiments, Day-5 predictive strength of SHIFRA$_K$ remains high despite the fact that 98% of the transferred embryos were blastocysts (very high-quality embryos) and endometrial receptivity was likely an important factor.

[0215] Prediction of embryo implantation was performed by KIDSCore-D3 on day-3 (66 hours' prediction time) and by KIDScore-D5 on day-5 (110 hours' prediction time) according to the manufacturer's protocols. Specificity and sensitivity were evaluated as a function of KIDScore values.

[0216] Reference is made to Fig. 30A, 30B, 30C, and 30D, which show exemplary embodiments of statistical characteristics of KID prediction by SHIFRA$_K$, in accordance with some embodiments of the present disclosure.

[0217] Fig. 30A shows KID prediction by SHIFRA$_K$ is shown for test-set Day-5 transferred embryos obtained from different medical centers. Fig. 30B shows Leave one clinic out cross validation predictions, where a classifier is trained on embryos only from three clinics and AUC is evaluated on embryos obtained from the fourth clinic, are presented. In some embodiments, AUC of clinic H3, which contributes most of the KID-labeled embryos, is relatively low consistent with high demand for additional embryos for training.

[0218] Fig. 30C shows Day-5 KID prediction using SHIFRA$_K$ is evaluated for test set embryos of different maternal age groups. Fig. 30D shows five-fold cross validation of embryo-stage learning, where a classifier is trained on 80% of the embryos and tested on the

remaining 20%, demonstrates small differences in AUC, which is indicative of lack of overfitting and supports generality of KID prediction.

[0219]   In some embodiments, generality of $SHIFRA_K$ is verified first via five-fold cross-validation of day-5 transferred embryos, showing negligible variation of KID prediction: $AUC = 0.69 \pm 0.02\ STD$. In some embodiments, KID prediction of test set embryos that are obtained from different clinics are compared, giving $AUC = 0.64 \pm 0.04\ STD$. AUC values by H1 (above average) and H2 (below average) clinics were likely skewed due to a highly uneven ratio between KID_p and KID_n train-set embryos. In some embodiments, generality was tested via leave-one-clinic-out cross-validation: $AUC = 0.64 \pm 0.04\ STD$. The below-average AUC by H3 is consistent with the smallest available train-set once H3 embryos were removed. The fraction of implanted embryos out of all transferred embryos (taking into account also KID_u transfers) was 2.7 fold higher for young women (age < 32; 39%) than older women (age > 38, 14%). In some embodiments, and in order to verify that $SHIFRA_K$ is not biased by maternal age, the embryos were divided into four age groups and tested KID-prediction on each. Satisfyingly, variation in day-5 KID prediction across age groups was small: average $AUC = 0.75 \pm 0.03\ STD$. In some embodiments, retrospective KID prediction statistics were demonstrated by setting the $SHIFRA_K$ threshold values that support PPV 0.59 and sensitivity 0.89. Thus, automated prediction of embryo implantation with superior predictive strength of day-5 transfers was confirmed, and generality across clinics and age groups and retrospectively demonstrate clinical utility was verified.

[0220]   Throughout the pre-menopause life span of the woman, oocytes are continuously exposed to various stress mediators. Hence, maternal age is an important determinant of female fertility. In some embodiments, a broad comparison was performed between the morphokinetic statistics of embryos that were obtained from young (age<32) versus older (age>38) women. In some embodiments, the temporal morphokinetic differences associated with maternal age appear negligible between PNa and 9C states (KS test < 0.08). In some embodiments, only morula compaction (tM) and start-of-blastulation (tSB) become retarded in advanced age. In some embodiments, the fraction of successfully implanted embryos out of all transferred embryos is 2.7 fold higher for young compared with older women, likely due to the accumulation of chromosomal aberrations and aneuploidy. To verify that

47

SHIFRA$_B$ and SHIFRA$_K$ are robust to the reduced developmental potential associated with advanced maternal age, BLAST-prediction and KID-prediction were compared at 72 hours. In some embodiments, the AUC values of BLAST and KID prediction of embryos from older women are higher than the AUC values of embryos from young women and the entire pool of embryos (Figs. 20B-i, and 23C-ii). This indicates that SHIFRA$_B$ and SHIFRA$_K$ are sensitive to visual features that are encoded within the video files of advanced maternal age.

[0221] In some embodiments, the set of 186 BLAST and KID co-labeled embryos was analyzed to study the relationship between blastulation and implantation potentials. Developmentally impaired embryos, for example due to chromosomal aberrations, can form blastocysts despite their compromised implantation potential. The positive correlation between BLAST and KID prediction scores (Spearman correlation 0.77) is indicative of common visual features. BLAST_n-KID_n (BnKn) embryos are of poor developmental quality and are scored lower than BLAST_p-KID_n (BpKn) and BLAST_p-KID_p (BpKp) embryos not only by SHIFRA$_B$ but also by SHIFRA$_K$. In some embodiments, the SHIFRA$_K$ but not SHIFRA$_B$ scores BpKp embryos higher than BpKn embryos in a statistically-significant manner, demonstrating that SHIFRA$_K$ can differentiate between blastocysts that have the capacity to implant in the uterus and the blastocysts that don't. This is indicative of visual features that are identified by SHIFRA$_K$, which are exclusively associated with implantation and not with blastulation.

[0222] In some embodiments, a careful analysis of the distributions of the embryonic states of BLAST-labelled and KID-labelled embryos between 60 and 90 hours shows agreement between the most probable states, consistent will all embryos pooled together, but the distributions of BLAST-labelled embryos are broader. In some embodiments, a measure of the number of states that BLAST-labelled and KID-labelled embryos are found in, is given by $2^{H_B(t)}$ and $2^{H_K(t)}$, respectively, where $H_B(t)$ and $H_K(t)$ are the corresponding information entropies. In some embodiments, between 60 and 90 hours, the BLAST-labeled embryos occupied 5.3 to 6.3 states, similar to all embryos pooled together, whereas KID-labeled embryos occupied only 3.3 to 4.7 states due to the preselection of embryos for transfer according to common morphological and/or morphokinetic parameters.

[0223] In some embodiments, the adjusted mutual information (AMI) was calculated between the distributions of the developmental states of the embryos at time of prediction and their BLAST and KID prediction scores in order to obtain insight into how embryonic developmental potential is evaluated. In some embodiments, the AMI quantifies the dependence between the state of the embryo at prediction time and the score it obtained by $SHIFRA_B$ and by $SHIFRA_K$. In some embodiments, as a control, the AMI for both classifiers is zero for a random state distribution.

[0224] In some embodiments, $SHIFRA_B$ is more sensitive to the developmental states of the embryos than $SHIFRA_K$, yet in both cases the AMI is smaller than 0.2, indicating that visual features other than the developmental states of the embryos have a greater effect on embryo classification. In some embodiments, the dependence of BLAST classification on the embryo state is increasing between 60 and 78 hours. In some embodiments, for KID classification, the AMI is highest between 72 and 76 hours and between 84 and 88 hours when the embryos approach morula compaction.

[0225] As will be appreciated by one skilled in the art, aspects of the present invention may be embodied as a system, method or computer program product. Accordingly, aspects of the present invention may take the form of an entirely hardware embodiment, an entirely software embodiment (including firmware, resident software, micro-code, etc.) or an embodiment combining software and hardware aspects that may all generally be referred to herein as a "circuit," "module" or "system." Furthermore, aspects of the present invention may take the form of a computer program product embodied in one or more computer readable medium(s) having computer readable program code embodied thereon.

[0226] Any combination of one or more computer readable medium(s) may be utilized. The computer readable medium may be a computer readable signal medium or a computer readable storage medium. A computer readable storage medium may be, for example, but not limited to, an electronic, magnetic, optical, electromagnetic, infrared, or semiconductor system, apparatus, or device, or any suitable combination of the foregoing. More specific examples (a non-exhaustive list) of the computer readable storage medium would include the following: an electrical connection having one or more wires, a portable computer diskette, a hard disk, a random access memory (RAM), a read-only memory (ROM), an

erasable programmable read-only memory (EPROM or Flash memory), an optical fiber, a portable compact disc read-only memory (CD-ROM), an optical storage device, a magnetic storage device, or any suitable combination of the foregoing. In the context of this document, a computer readable storage medium may be any tangible medium that can contain or store a program for use by or in connection with an instruction execution system, apparatus, or device.

[0227]   A computer readable signal medium may include a propagated data signal with computer readable program code embodied therein, for example, in baseband or as part of a carrier wave. Such a propagated signal may take any of a variety of forms, including, but not limited to, electro-magnetic, optical, or any suitable combination thereof. A computer readable signal medium may be any computer readable medium that is not a computer readable storage medium and that can communicate, propagate, or transport a program for use by or in connection with an instruction execution system, apparatus, or device.

[0228]   Program code embodied on a computer readable medium may be transmitted using any appropriate medium, including but not limited to wireless, wireline, optical fiber cable, RF, etc., or any suitable combination of the foregoing.

[0229]   Computer program code for carrying out operations for aspects of the present invention may be written in any combination of one or more programming languages, including an object-oriented programming language such as Java, Smalltalk, C++ or the like and conventional procedural programming languages, such as the "C" programming language or similar programming languages. The program code may execute entirely on the user's computer, partly on the user's computer, as a stand-alone software package, partly on the user's computer and partly on a remote computer or entirely on the remote computer or server. In the latter scenario, the remote computer may be connected to the user's computer through any type of network, including a local area network (LAN) or a wide area network (WAN), or the connection may be made to an external computer (for example, through the Internet using an Internet Service Provider).

[0230]   Aspects of the present invention are described herein with reference to flowchart illustrations and/or block diagrams of methods, apparatus (systems) and computer program products according to embodiments of the invention. It will be understood that each block

of the flowchart illustrations and/or block diagrams, and combinations of blocks in the flowchart illustrations and/or block diagrams, can be implemented by computer program instructions. These computer program instructions may be provided to a hardware processor of a general-purpose computer, special purpose computer, or other programmable data processing apparatus to produce a machine, such that the instructions, which execute via the processor of the computer or other programmable data processing apparatus, create means for implementing the functions/acts specified in the flowchart and/or block diagram block or blocks.

[0231] These computer program instructions may also be stored in a computer readable medium that can direct a computer, other programmable data processing apparatus, or other devices to function in a particular manner, such that the instructions stored in the computer readable medium produce an article of manufacture including instructions which implement the function/act specified in the flowchart and/or block diagram block or blocks.

[0232] The computer program instructions may also be loaded onto a computer, other programmable data processing apparatus, or other devices to cause a series of operational steps to be performed on the computer, other programmable apparatus or other devices to produce a computer implemented process such that the instructions which execute on the computer or other programmable apparatus provide processes for implementing the functions/acts specified in the flowchart and/or block diagram block or blocks.

[0233] The flowcharts and block diagrams in the Figures illustrate the architecture, functionality, and operation of possible implementations of systems, methods and computer program products according to various embodiments of the present invention. In this regard, each block in the flowchart or block diagrams may represent a module, segment, or portion of code, which comprises one or more executable instructions for implementing the specified logical function(s). It should also be noted that, in some alternative implementations, the functions noted in the block may occur out of the order noted in the figures. For example, two blocks shown in succession may, in fact, be executed substantially concurrently, or the blocks may sometimes be executed in the reverse order, depending upon the functionality involved. It will also be noted that each block of the block diagrams and/or flowchart illustration, and combinations of blocks in the block diagrams and/or flowchart illustration,

can be implemented by special purpose hardware-based systems that perform the specified functions or acts, or combinations of special purpose hardware and computer instructions.

[0234] The descriptions of the various embodiments of the present invention have been presented for purposes of illustration but are not intended to be exhaustive or limited to the embodiments disclosed. Many modifications and variations will be apparent to those of ordinary skill in the art without departing from the scope and spirit of the described embodiments. The terminology used herein was chosen to best explain the principles of the embodiments, the practical application or technical improvement over technologies found in the marketplace, or to enable others of ordinary skill in the art to understand the embodiments disclosed herein.

[0235] In the description and claims of the application, each of the words "comprise" "include" and "have", and forms thereof, are not necessarily limited to members in a list with which the words may be associated. In addition, where there are inconsistencies between this application and any document incorporated by reference, it is hereby intended that the present application controls.

CLAIMS

What is claimed is:

1. A system comprising:

at least one hardware processor; and

a non-transitory computer-readable storage medium having stored thereon program code, the program code executable by the at least one hardware processor to:

receive a plurality of video segments, each depicting prenatal embryogenesis of a corresponding embryo;

divide each of said video segments into a plurality of consecutive packets, wherein each of said plurality of packets comprises a specified number of frames;

train a first machine learning model on a training set comprising

(i)   said packets, and

(ii)   labels indicating a developmental parameter associated with each of said corresponding embryos; and

train a second machine learning model on a training set comprising:

(iii) sets of outputs of said first machine learning model, wherein each of said sets is associated with said packets comprising one of said video segments, and

(iv)   labels indicating said developmental parameter associated with said corresponding embryo depicted in said one of said video segments.

2. The system of claim 1, wherein, with respect to each of said packets, said output of said trained first machine learning model is a numerical representations indicating a probability associated with said developmental parameter.

3. The system of claim 2, wherein said numerical representation is one of: a scalar representation, a vector representation, and a matrix representation.

4. The system of any one of claims 1-3, wherein said numerical representation reflects a dimensionality reduction.

5.   The system of any one of claims 1-4, wherein said trained second machine learning model predicts a developmental potential associated with each of said corresponding embryos.

6.   The system of any one of claims 1-5, further comprising, at an inference stage:

   (i)   applying said trained first machine learning model to target packets associated with a target video segment depicting prenatal embryogenesis of a target embryo, to obtain said numerical representations for each of said target packets; and

   (ii)   applying said trained second machine learning model to said obtained numerical representations, to predict a developmental potential of said target embryo.

7.   The system of any one of claims 1-6, wherein said first machine learning model comprises at least two machine learning models, and wherein:

   (i)   with respect to a first of said machine learning models, said developmental parameter indicated by said labels is a blastulation state; and

   (ii)   with respect to a second of said machine learning models, said developmental parameter indicated by said labels is an implantation state.

8.   The system of any one of claims 1-6, wherein said developmental parameter comprises at least one of: morphological stage, cleavage stage, number of cells, cell fragmentation, cell symmetry, inner cell mass, trophectoderm, pronuclei symmetry, pronuclei movement, pronuclei location, cell location, and cell movement.

9.   The system of claim 8, wherein said trained second machine learning model predicts at least one of: morphological stage, cleavage stage, number of cells, cell fragmentation, cell symmetry, inner cell mass, trophectoderm, pronuclei symmetry, pronuclei movement, pronuclei location, cell location, and cell movement.

10.   The system of any one of claims 1-6, wherein said first machine learning model comprises a self-supervised algorithm.

11.   A method comprising:

54

receiving a plurality of video segments, each depicting prenatal embryogenesis of a corresponding embryo;

dividing each of said video segments into a plurality of consecutive packets, wherein each of said plurality of packets comprises a specified number of frames;

training a first machine learning model on a training set comprising

(i)   said packets, and

(ii)   labels indicating a developmental parameter associated with each of said corresponding embryos; and

training a second machine learning model on a training set comprising:

(iii)   sets of outputs of said first machine learning model, wherein each of said sets is associated with said packets comprising one of said video segments, and

(iv)   labels indicating said developmental parameter associated with said corresponding embryo depicted in said one of said video segments.

12.   The method of claim 11, wherein, with respect to each of said packets, said output of said trained first machine learning model is a numerical representations indicating a probability associated with said developmental parameter.

13.   The method of claim 12, wherein said numerical representation is one of: a scalar representation, a vector representation, and a matrix representation.

14.   The method of any one of claims 11-13, wherein said numerical representation reflects a dimensionality reduction.

15.   The method of any one of claims 11-14, wherein said trained second machine learning model predicts a developmental potential associated with each of said corresponding embryos.

16.   The method of any one of claims 11-15, further comprising, at an inference stage:

(i)   applying said trained first machine learning model to target packets associated with a target video segment depicting prenatal embryogenesis of a target embryo, to obtain said numerical representations for each of said target packets; and

(ii)   applying said trained second machine learning model to said obtained
         numerical representations, to predict a developmental potential of said target
         embryo.

17. The method of any one of claims 11-16, wherein said first machine learning model
comprises at least two machine learning models, and wherein:

(i)   with respect to a first of said machine learning models, said developmental
         parameter indicated by said labels is a blastulation state; and

(ii)   with respect to a second of said machine learning models, said developmental
         parameter indicated by said labels is an implantation state.

18. The method of any one of claims 11-16, wherein said developmental parameter
comprises at least one of: morphological stage, cleavage stage, number of cells, cell
fragmentation, cell symmetry, inner cell mass, trophectoderm, pronuclei symmetry,
pronuclei movement, pronuclei location, cell location, and cell movement.

19. The method of claim 18, wherein said trained second machine learning model
predicts at least one of: morphological stage, cleavage stage, number of cells, cell
fragmentation, cell symmetry, inner cell mass, trophectoderm, pronuclei symmetry,
pronuclei movement, pronuclei location, cell location, and cell movement.

20. The method of any one of claims 11-16, wherein said first machine learning model
comprises a self-supervised algorithm.

21. A computer program product comprising a non-transitory computer-readable storage
medium having program code embodied therewith, the program code executable by at least
one hardware processor to:

     receive a plurality of video segments, each depicting prenatal embryogenesis of a
     corresponding embryo;

     divide each of said video segments into a plurality of consecutive packets, wherein
     each of said plurality of packets comprises a specified number of frames;

     train a first machine learning model on a training set comprising

     (i)   said packets, and

(ii)   labels indicating a developmental parameter associated with each of said corresponding embryos; and

train a second machine learning model on a training set comprising:

(iii)   sets of outputs of said first machine learning model, wherein each of said sets is associated with said packets comprising one of said video segments, and

(iv)   labels indicating said developmental parameter associated with said corresponding embryo depicted in said one of said video segments.

22. The computer program product of claim 21, wherein, with respect to each of said packets, said output of said trained first machine learning model is a numerical representations indicating a probability associated with said developmental parameter.

23. The computer program product of claim 22, wherein said numerical representation is one of: a scalar representation, a vector representation, and a matrix representation.

24. The computer program product of any one of claims 21-23, wherein said numerical representation reflects a dimensionality reduction.

25. The computer program product of any one of claims 21-24, wherein said trained second machine learning model predicts a developmental potential associated with each of said corresponding embryos.

26. The computer program product of any one of claims 21-25, further comprising, at an inference stage:

(i)   applying said trained first machine learning model to target packets associated with a target video segment depicting prenatal embryogenesis of a target embryo, to obtain said numerical representations for each of said target packets; and

(ii)   applying said trained second machine learning model to said obtained numerical representations, to predict a developmental potential of said target embryo.

27. The computer program product of any one of claims 21-26, wherein said first machine learning model comprises at least two machine learning models, and wherein:

(i)   with respect to a first of said machine learning models, said developmental parameter indicated by said labels is a blastulation state; and

(ii)   with respect to a second of said machine learning models, said developmental parameter indicated by said labels is an implantation state.

28. The computer program product of any one of claims 21-26, wherein said developmental parameter comprises at least one of: morphological stage, cleavage stage, number of cells, cell fragmentation, cell symmetry, inner cell mass, trophectoderm, pronuclei symmetry, pronuclei movement, pronuclei location, cell location, and cell movement.

29. The computer program product of claim 28, wherein said trained second machine learning model predicts at least one of: morphological stage, cleavage stage, number of cells, cell fragmentation, cell symmetry, inner cell mass, trophectoderm, pronuclei symmetry, pronuclei movement, pronuclei location, cell location, and cell movement.

30. The computer program product of any one of claims 21-26, wherein said first machine learning model comprises a self-supervised algorithm.

FIG. 1

Incubator	
TL1	4844
TL2	610
TL3	2914
TL4	673
TL5	3752
TL6	2215
TL7	1455
TL8	2651
TL9	1121
Total	**20,235**

Medical center	
H1	6,304
H2	4,624
H3	7,839
H4	1,486
Total	**20,253**

Maternal age	
20-30	4,134
30-35	5,800
35-40	4,868
40+	3,548
Total	**18,350**

Transfer rate	
1	2,749
2	2,170
3	399
4	53
Total	**5,371**

KID	
KIDp	1,199
KIDn	4,326
KIDu	979
Total	**6,504**

FIG. 2A

Incubator	
TL1	3,104
TL2	242
TL3	1,975
TL4	305
TL5	1,524
TL6	870
TL7	1,074
TL8	1,641
TL9	359
Total	11,094

Medical center	
H1	2,394
H2	1,839
H3	5,094
H4	1,074
H5	693
Total	11,094

Maternal age	
20-30	2,511
30-35	3,397
35-40	2,952
40+	2,215
Total	11,075

Transfer rate	
1	2,010
2	1,361
3	255
4	27
Total	3,653

KID	
KIDp	1,131
KIDn	3,382
KIDu	896
Total	5,409

FIG. 2B

Embryo video file → Embryo ROI cropping → Down sampling → Grouping five-frame packets p_n^m

Discard empty frames

FIG. 3

FIG. 4A

FIG. 4B

WO 2020/157761   PCT/IL2020/050120

6 / 62
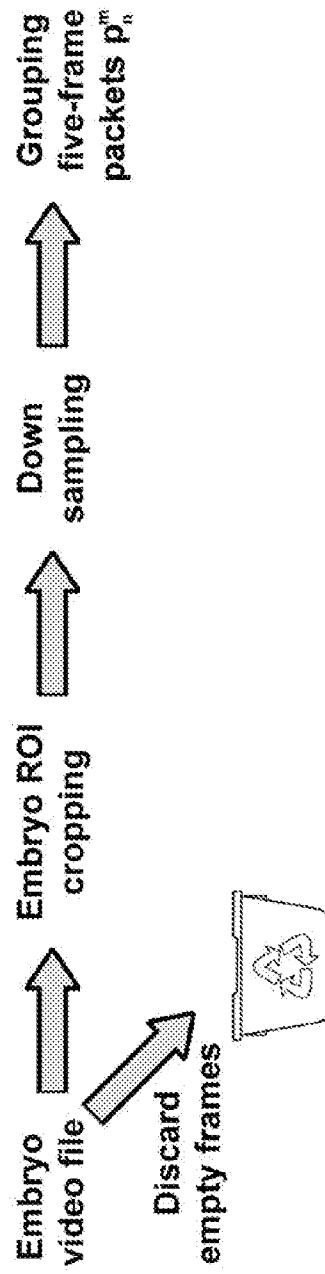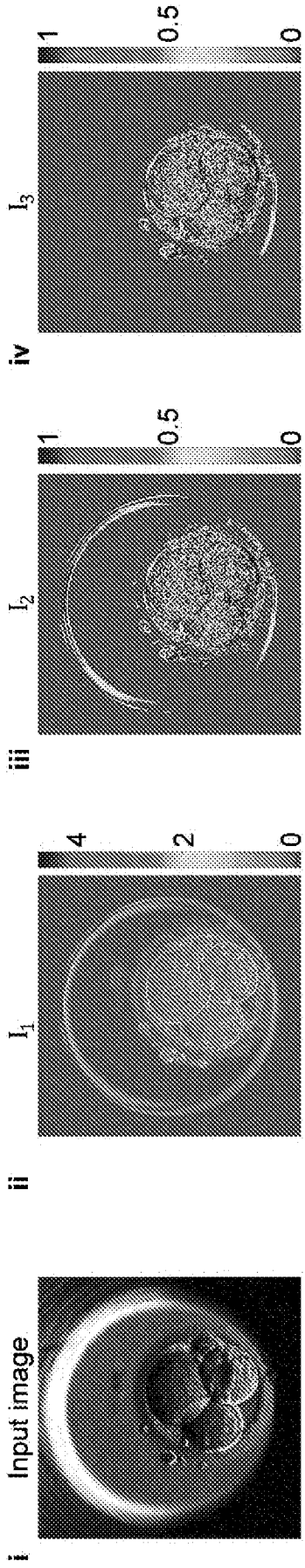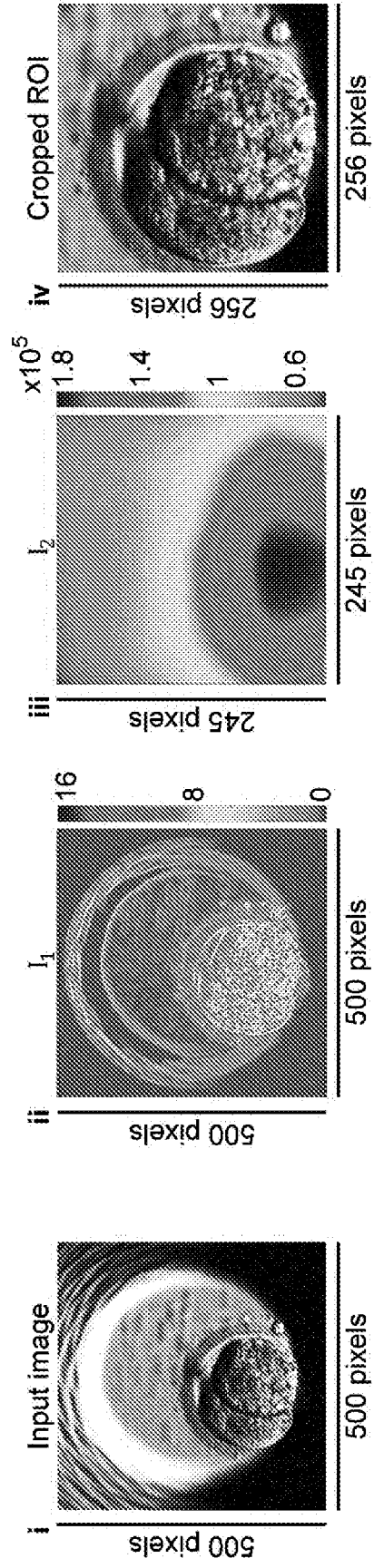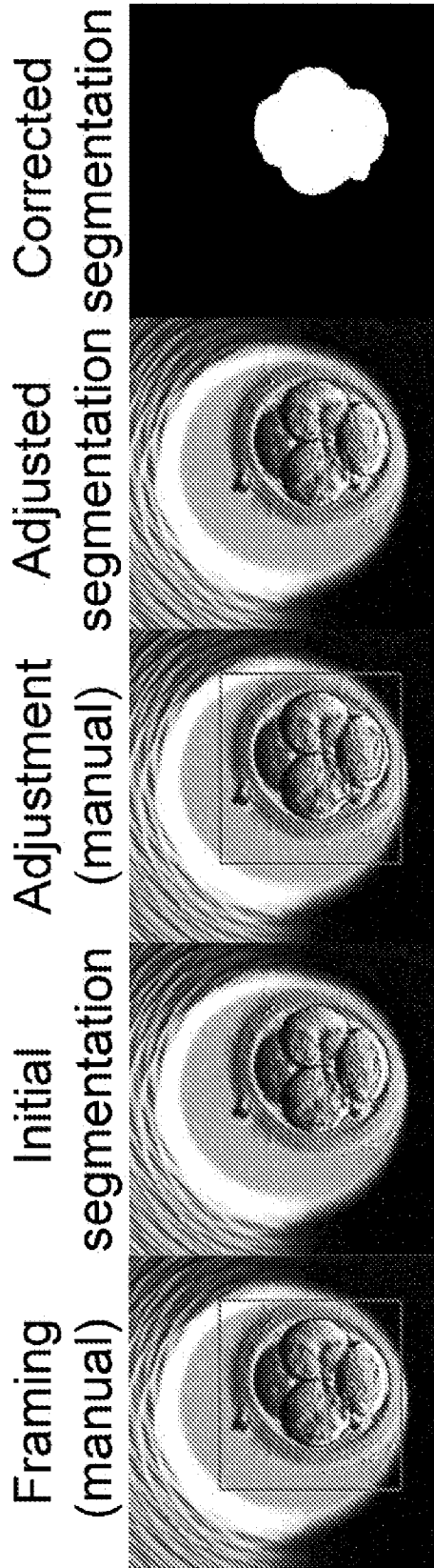
FIG. 5A
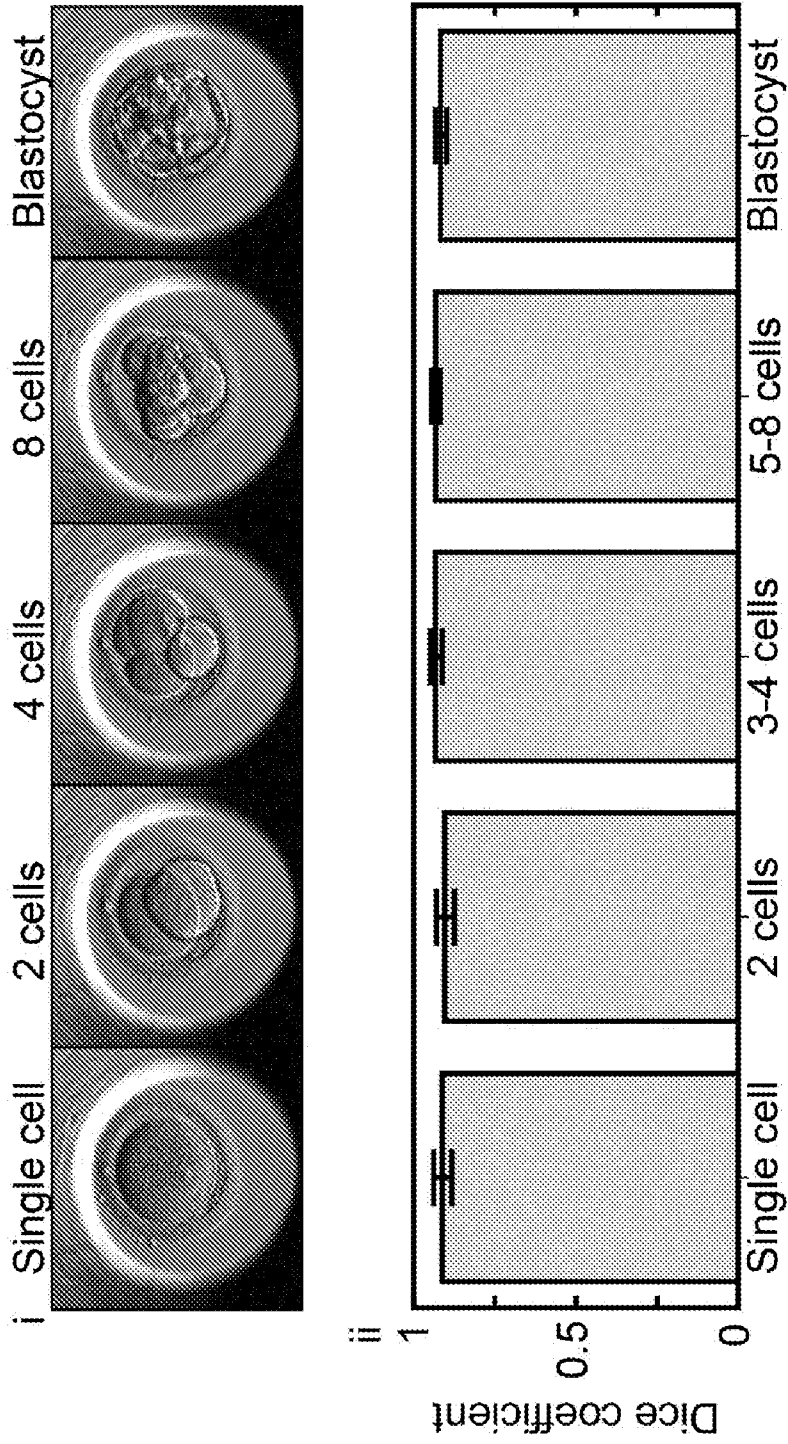
FIG. 5B

Morphokientic event	tPNa	tPNf	t2	t3	t4	t5	t6	t7	t8	t9	tM	tSB
1st percentile [D:HH:MM]	0:02:11	0:18:05	0:20:18	1:02:07	1:06:17	1:10:42	1:13:27	1:17:43	1:19:20	1:23:57	2:17:28	3:05:33
99th percentile [D:HH:MM]	0:20:57	1:22:31	2:01:08	2:15:21	2:23:31	3:13:01	3:18:42	4:01:17	4:05:09	4:14:34	5:04:15	5:11:39

FIG. 6

| Morphokientic intervals	tPNf-tPNa	t2-tPNf	t3-t2	t4-t3	t5-t4	t6-t5	t7-t6	t8-t7	t9-t8	tM-t9	tSB-tM
1st percentile [D:HH:MM]	12:16:04	1:12:10	0:00:00	0:00:00	0:00:00	0:00:00	0:00:00	0:00:00	0:00:00	0:00:00	0:16:00
99th percentile [D:HH:MM]	1:07:07	0:05:00	0:21:37	0:17:30	1:03:50	0:21:55	0:22:40	0:22:00	1:03:40	1:21:50	1:10:04

FIG. 7
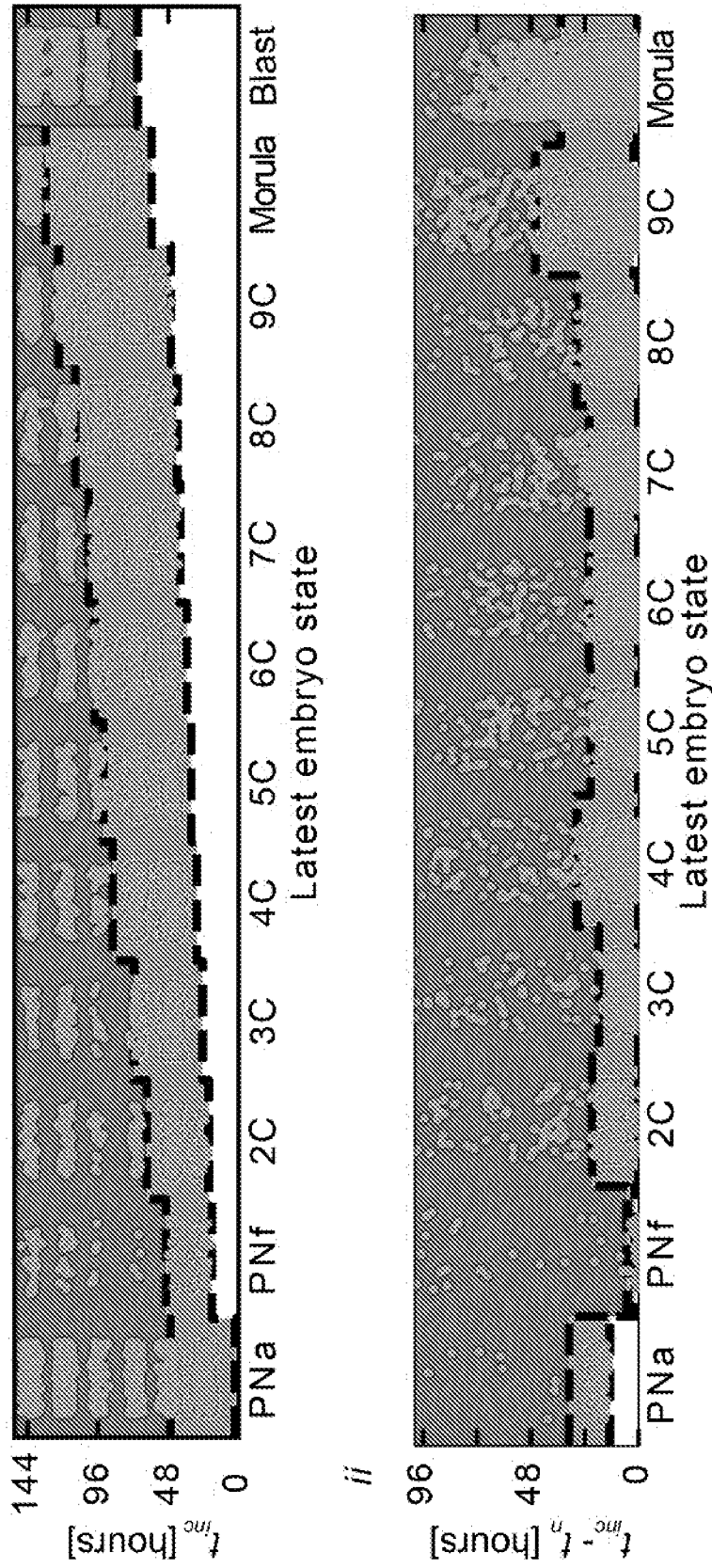
FIG. 8A
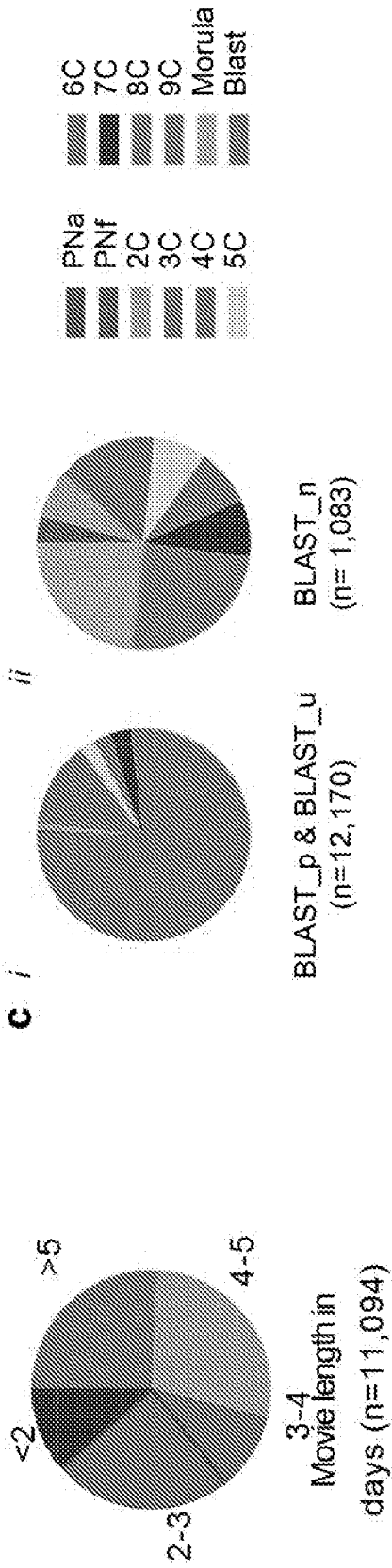
FIG. 8C

FIG. 8D

FIG. 8B
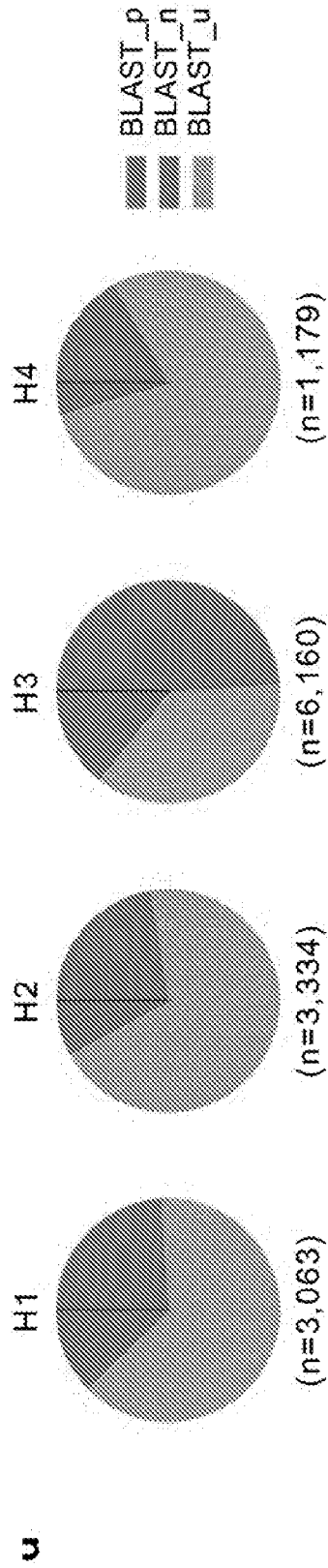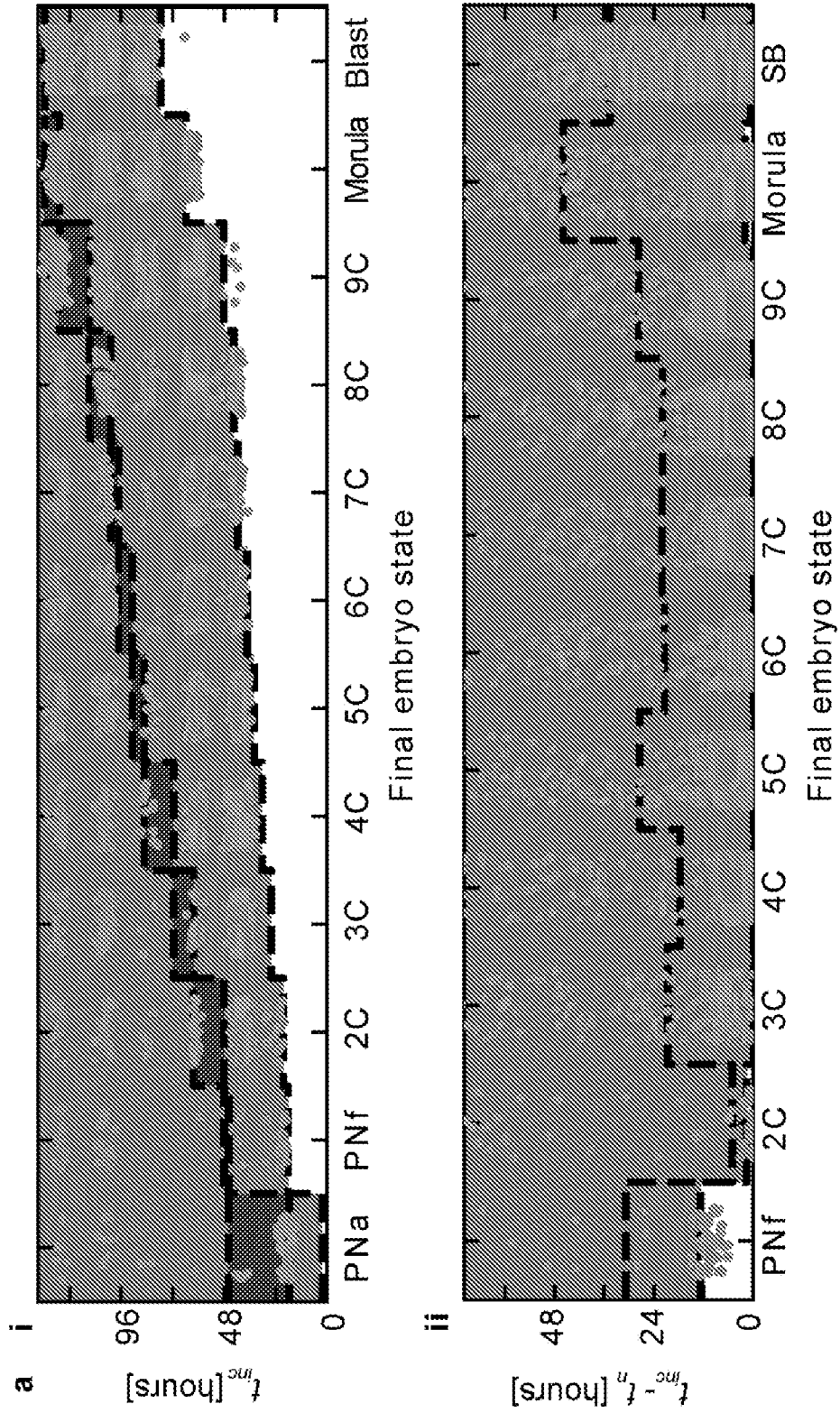
FIG. 8E
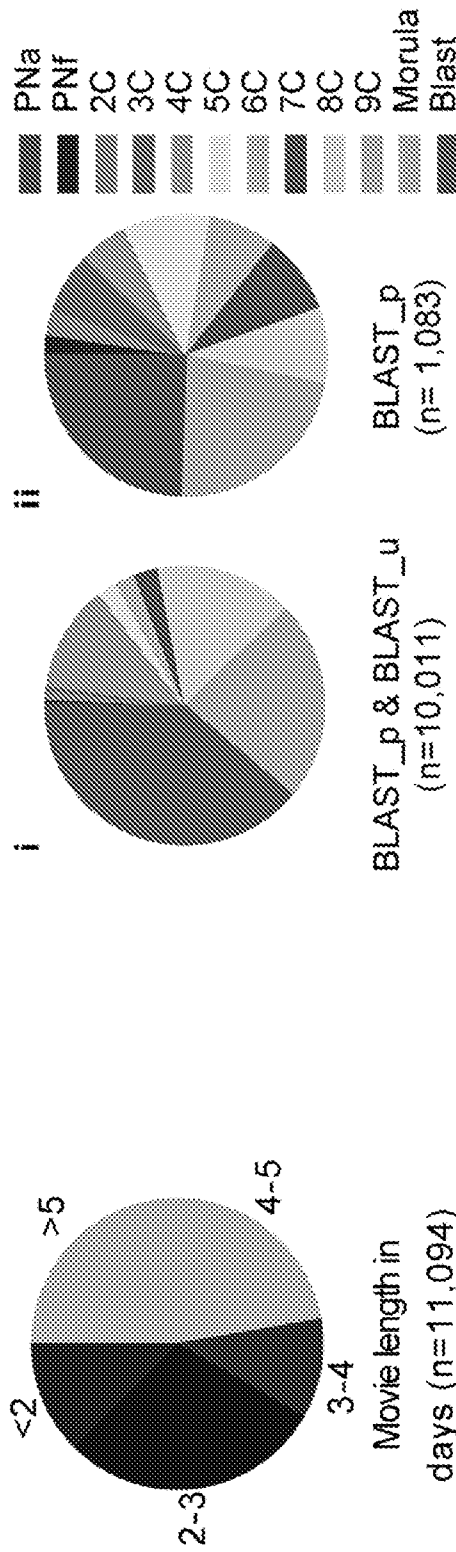
FIG. 8G

FIG. 8F
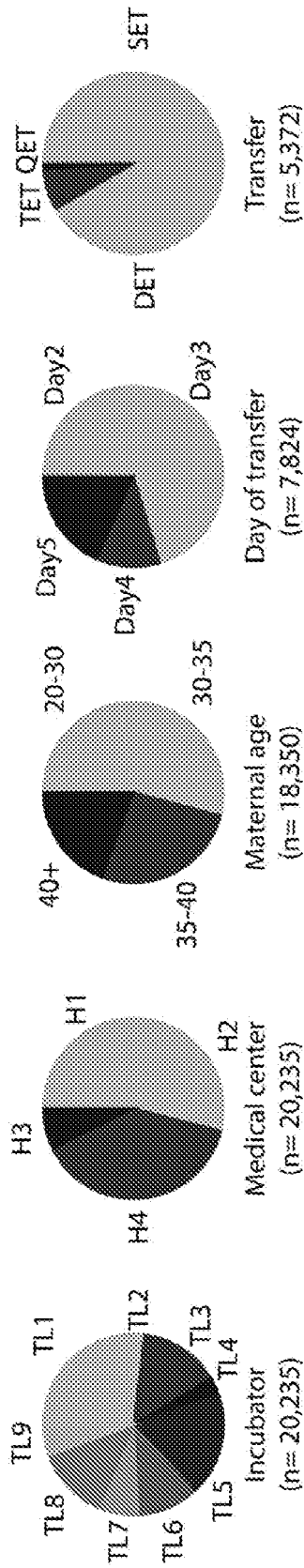
FIG. 9A
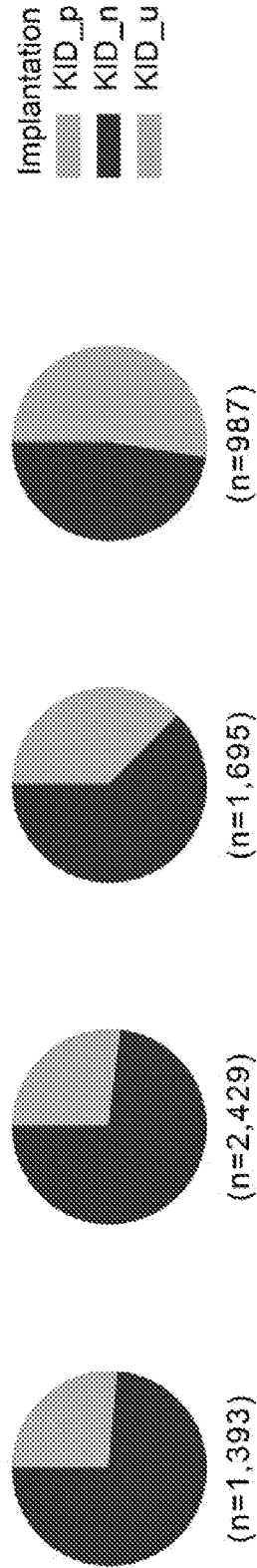
FIG. 9C
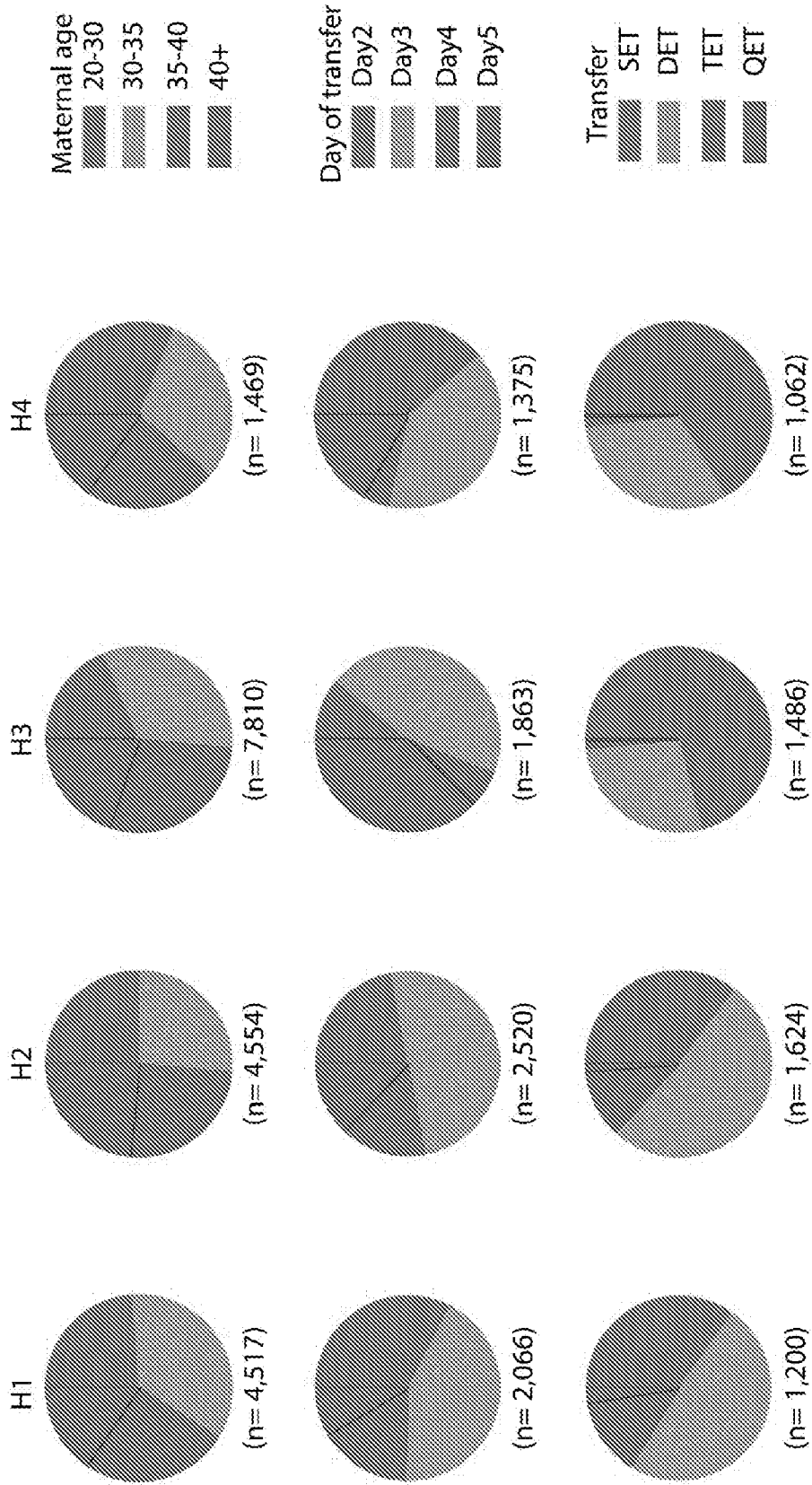
FIG. 9B
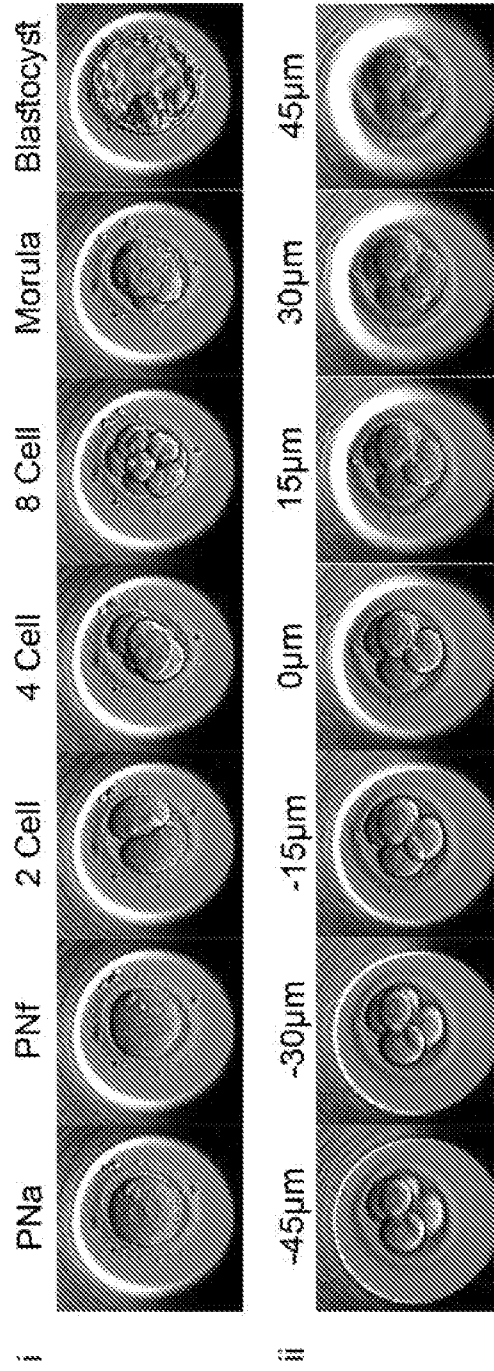
FIG. 10A
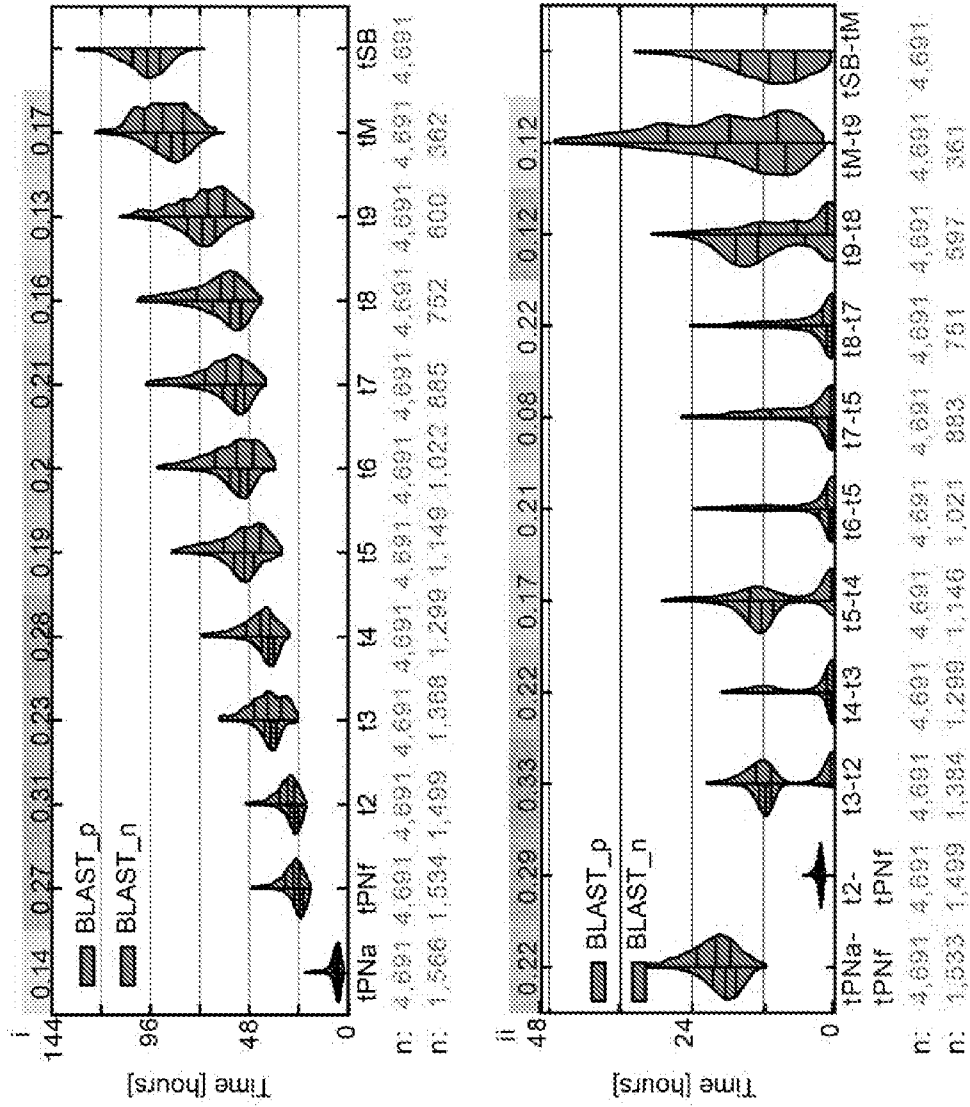
FIG. 10B
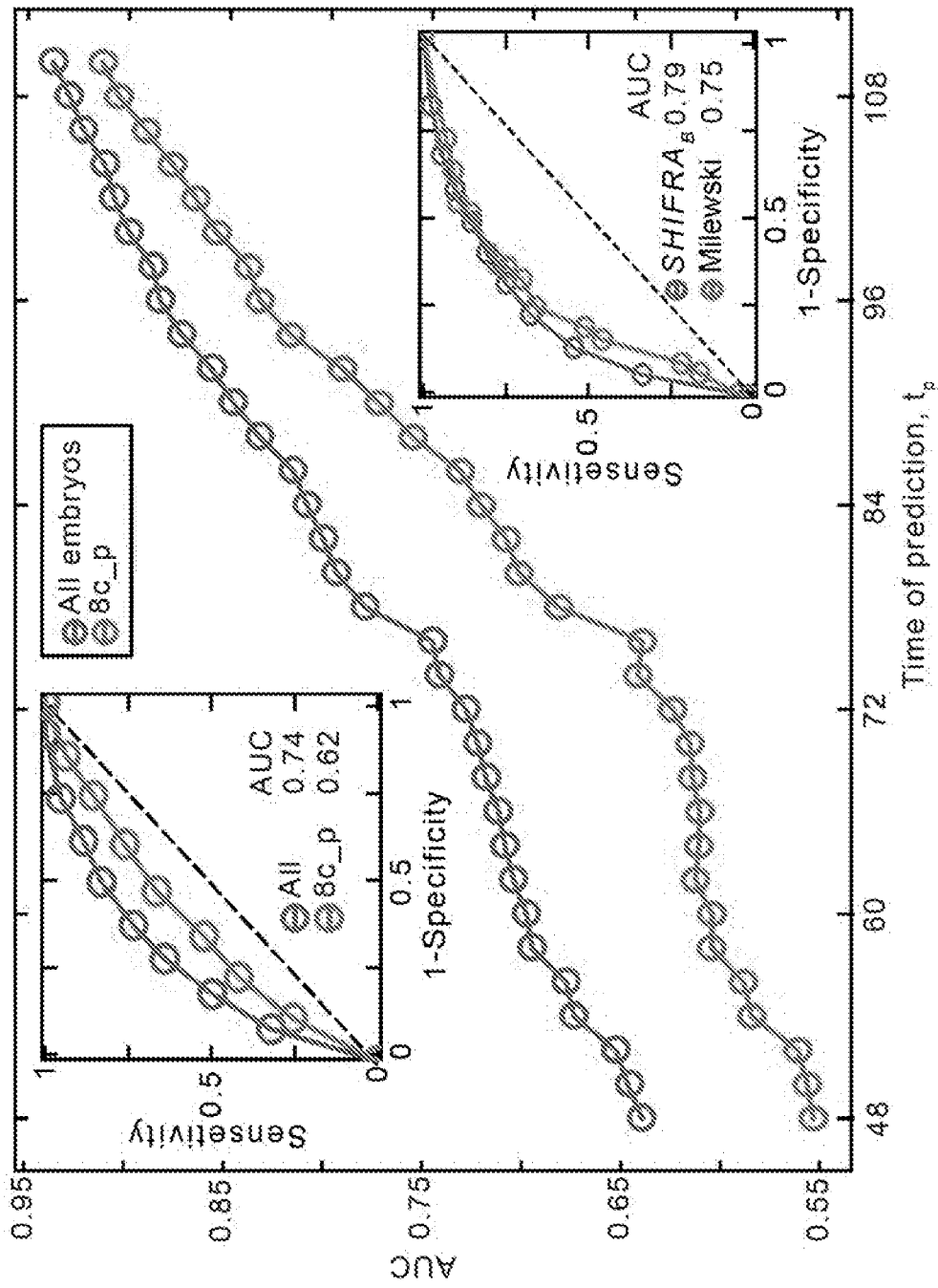
FIG. 10C
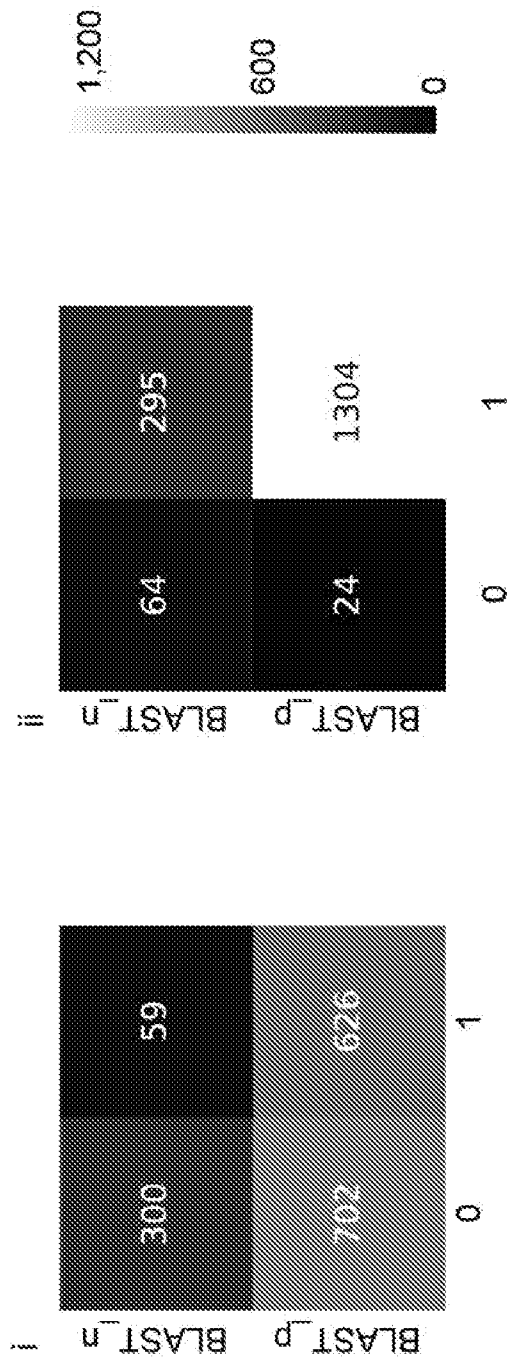
FIG. 10D
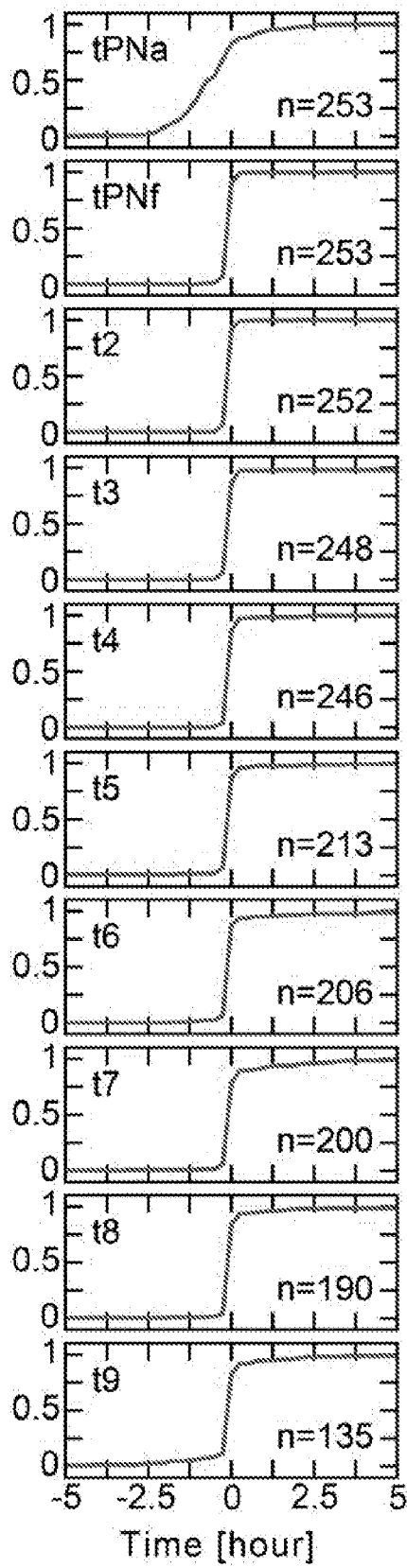
FIG. 11A
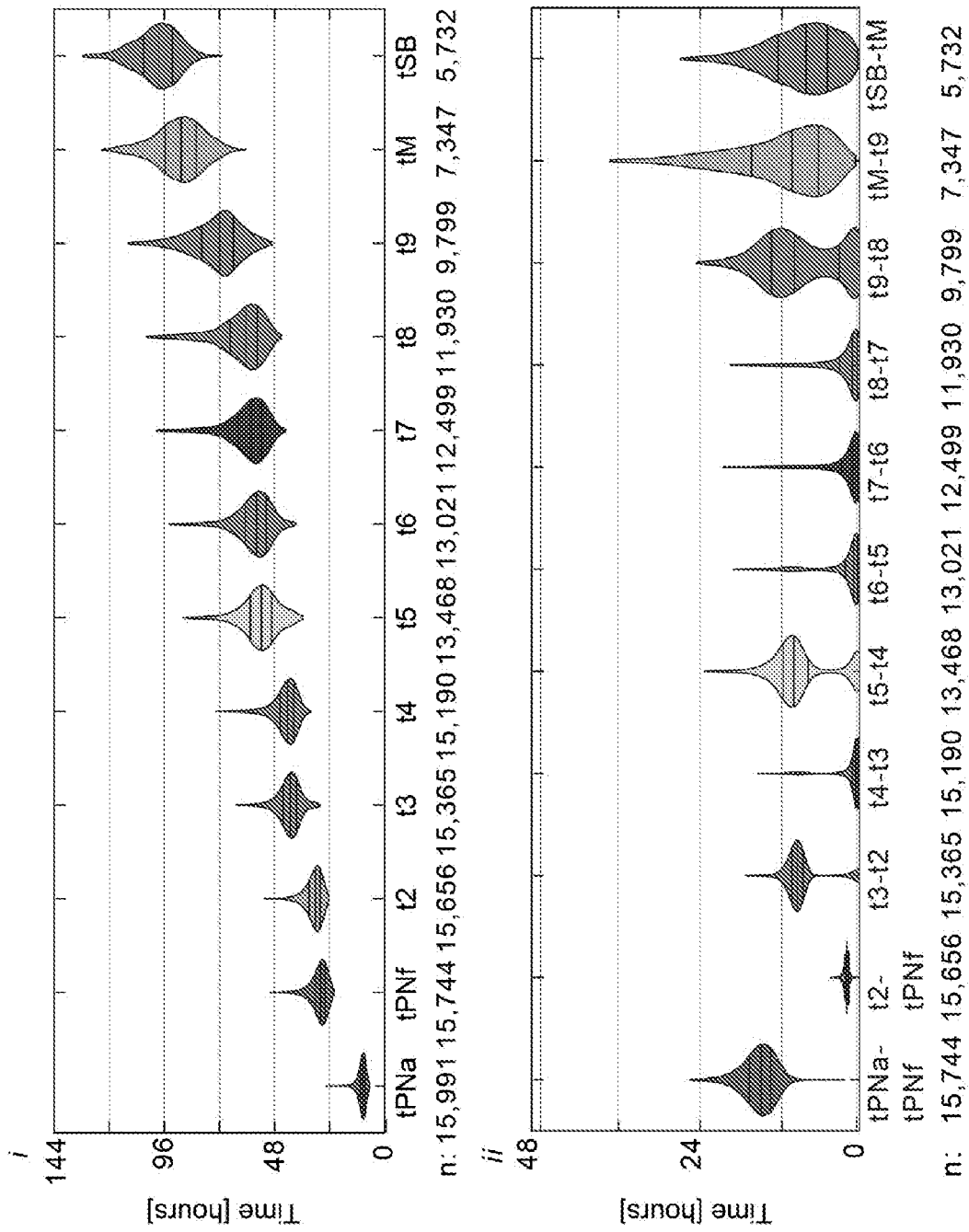
FIG. 11B
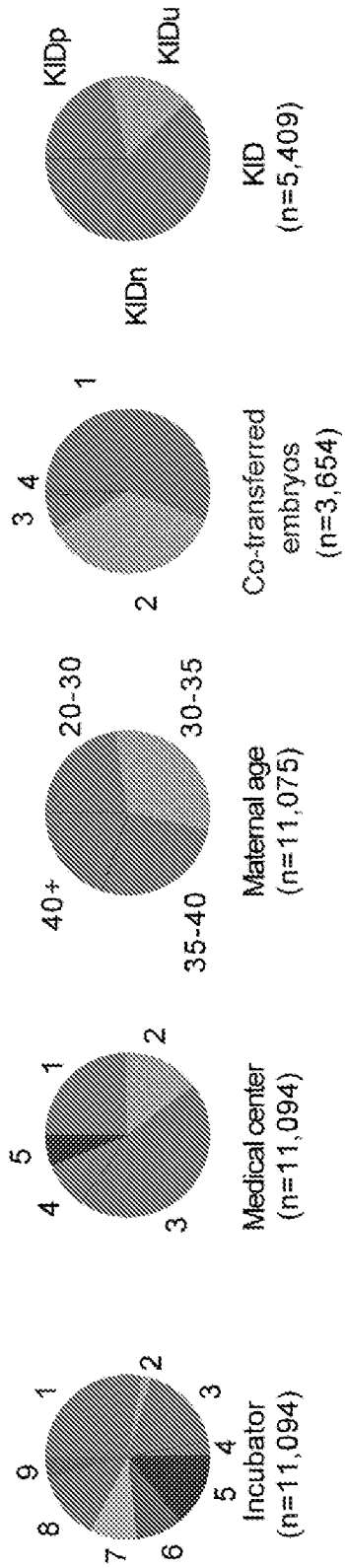
FIG. 11C
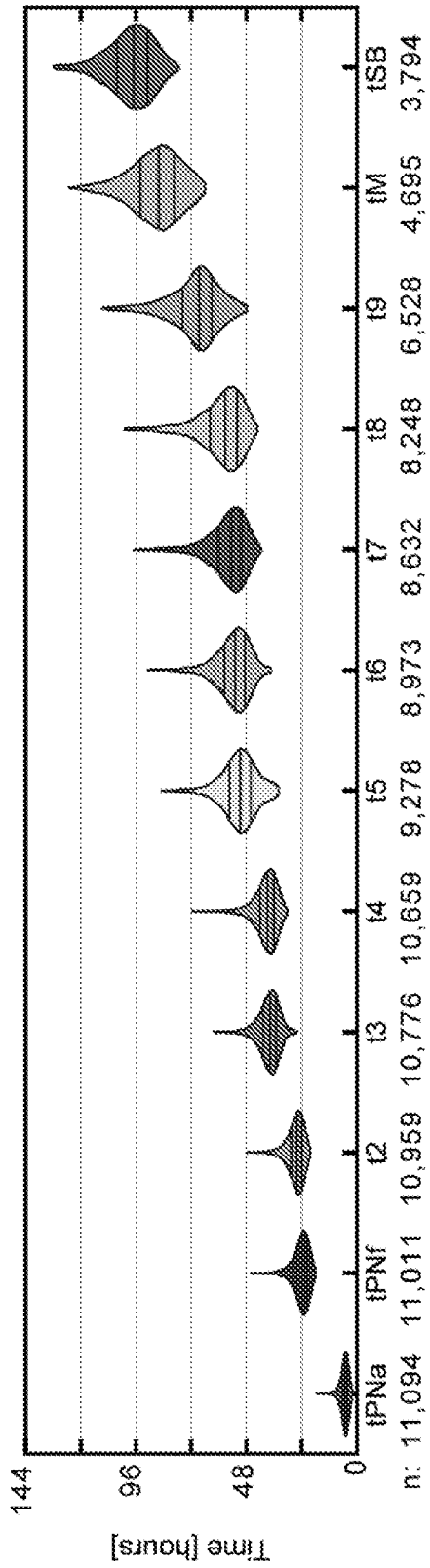
FIG. 11D
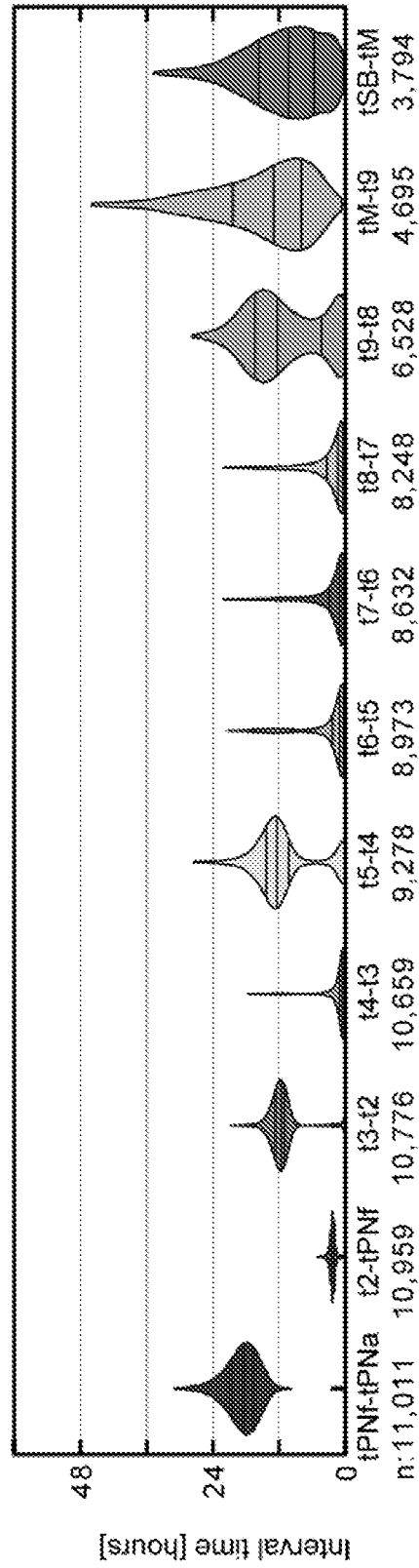
FIG. 11E
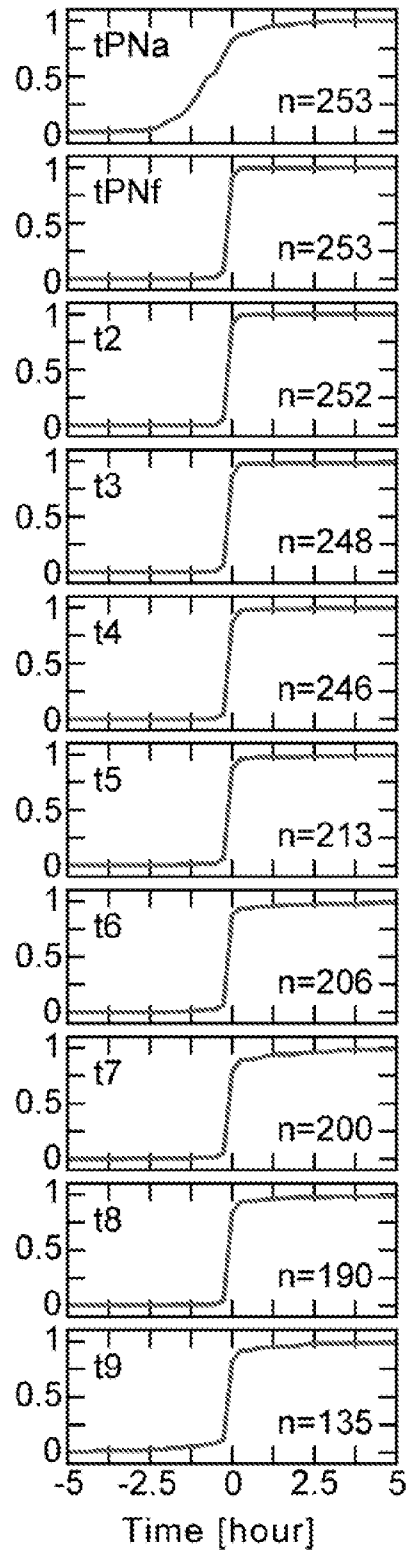
FIG. 11F
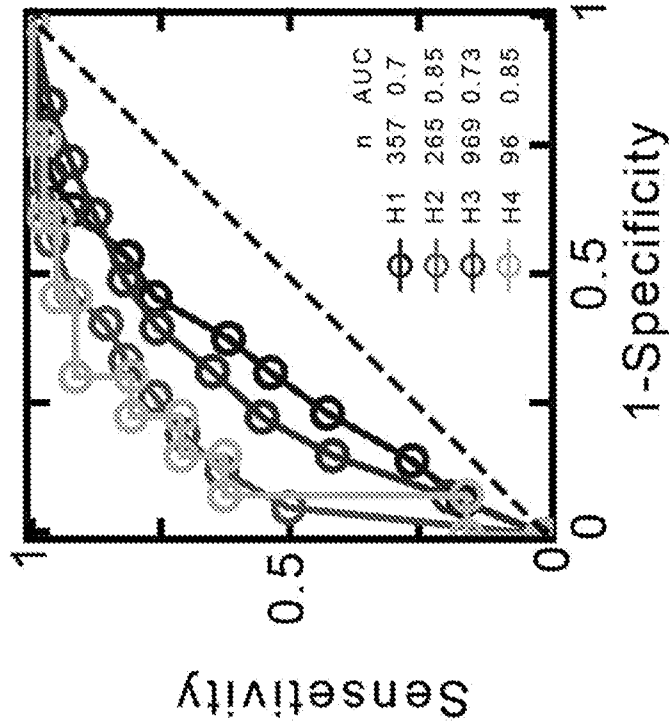
FIG. 12B
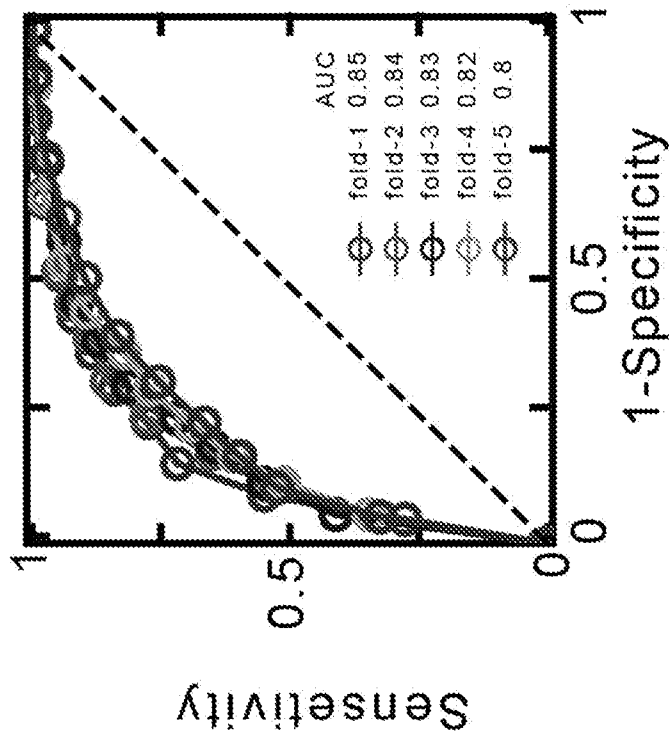
FIG. 12A
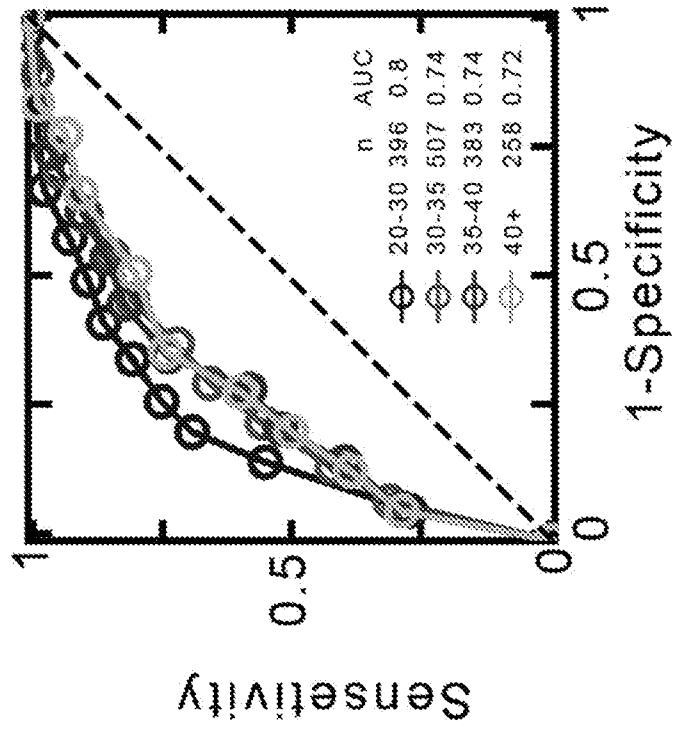
FIG. 12D
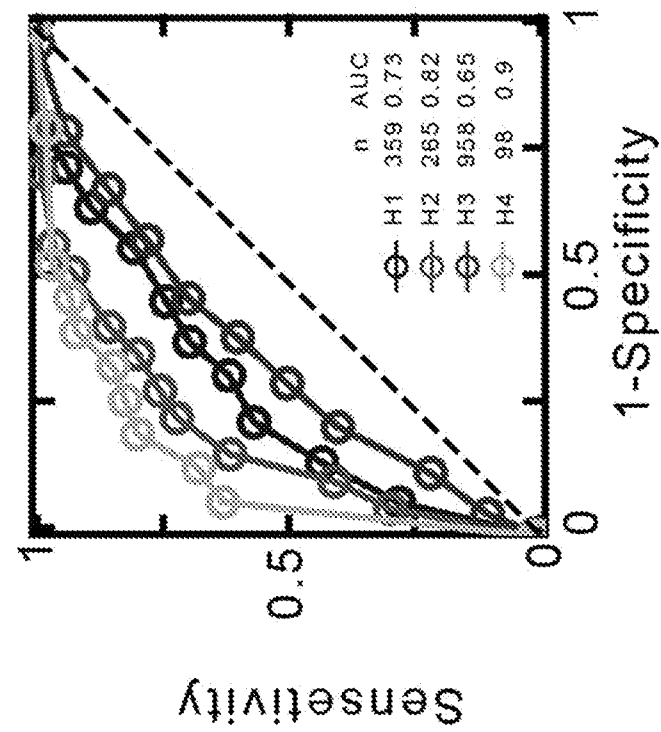
FIG. 12C
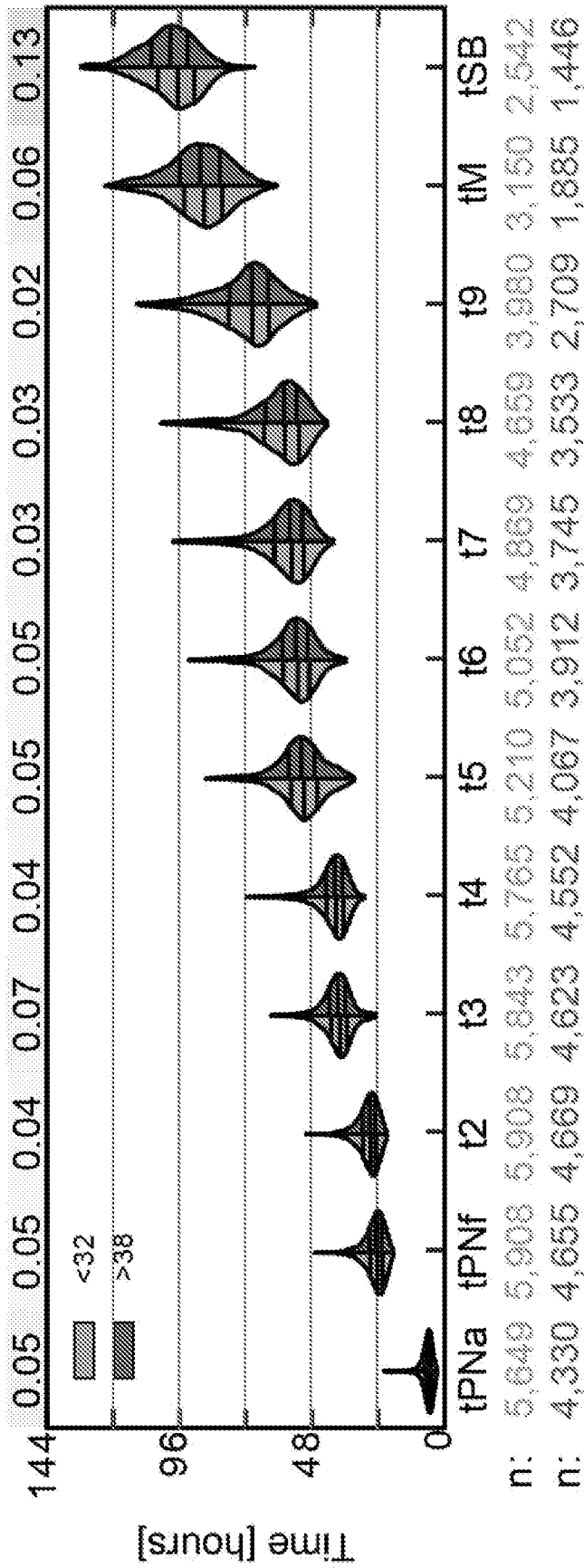
FIG. 13A
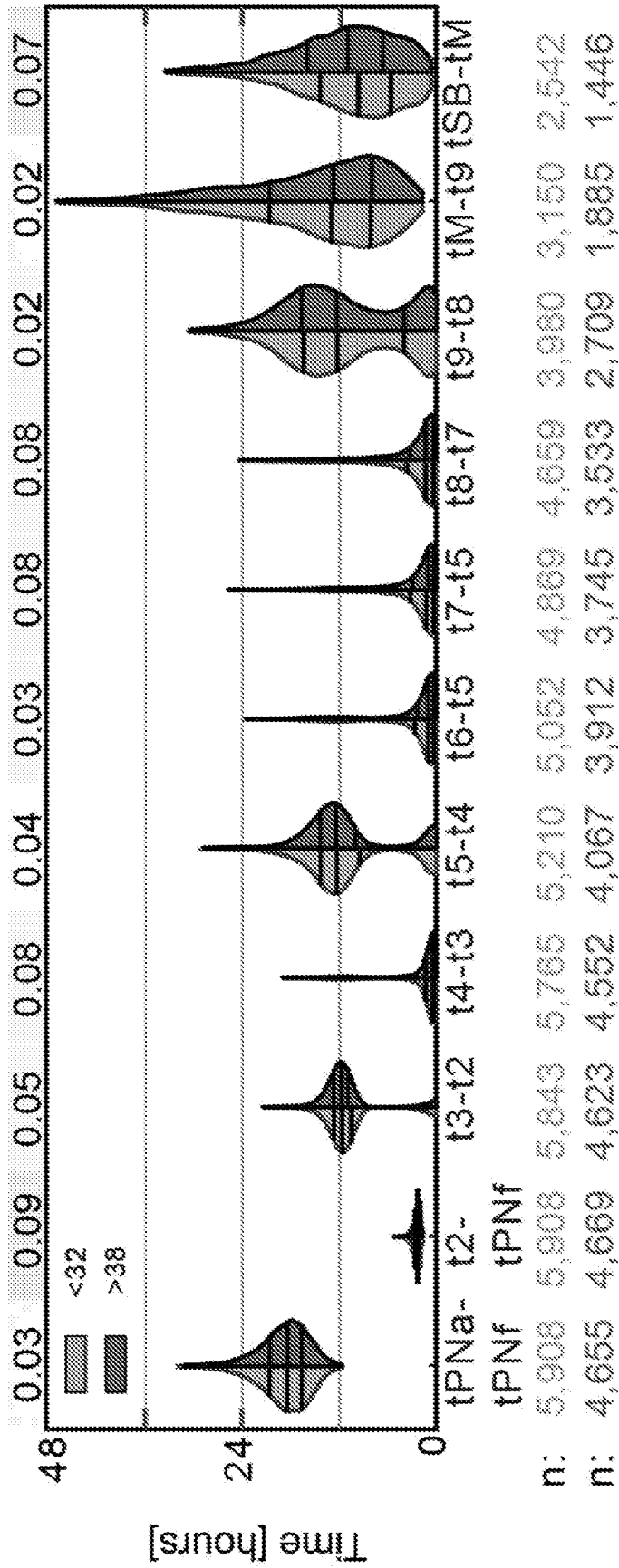
FIG. 13B
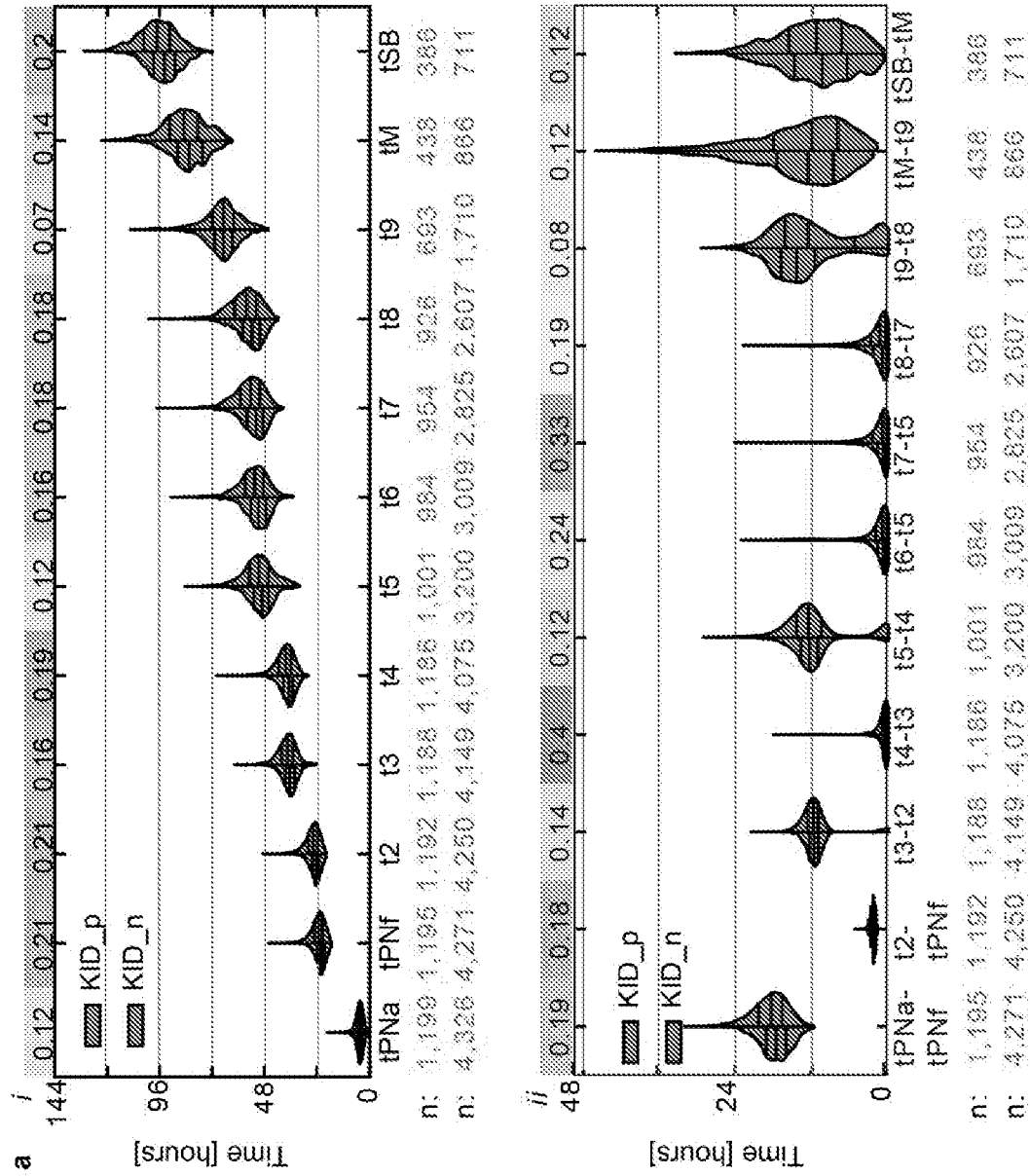
FIG. 14A
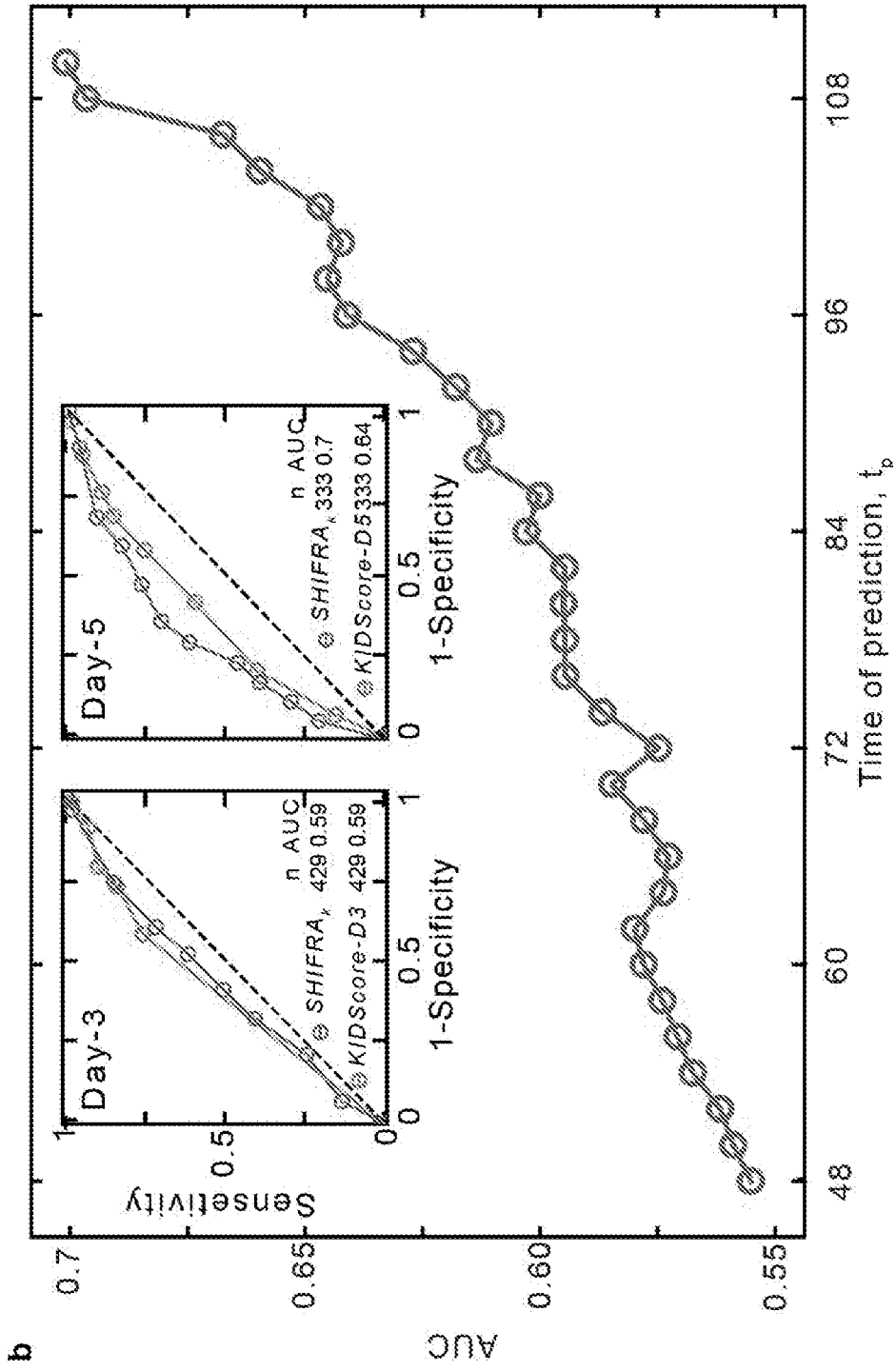
FIG. 14B
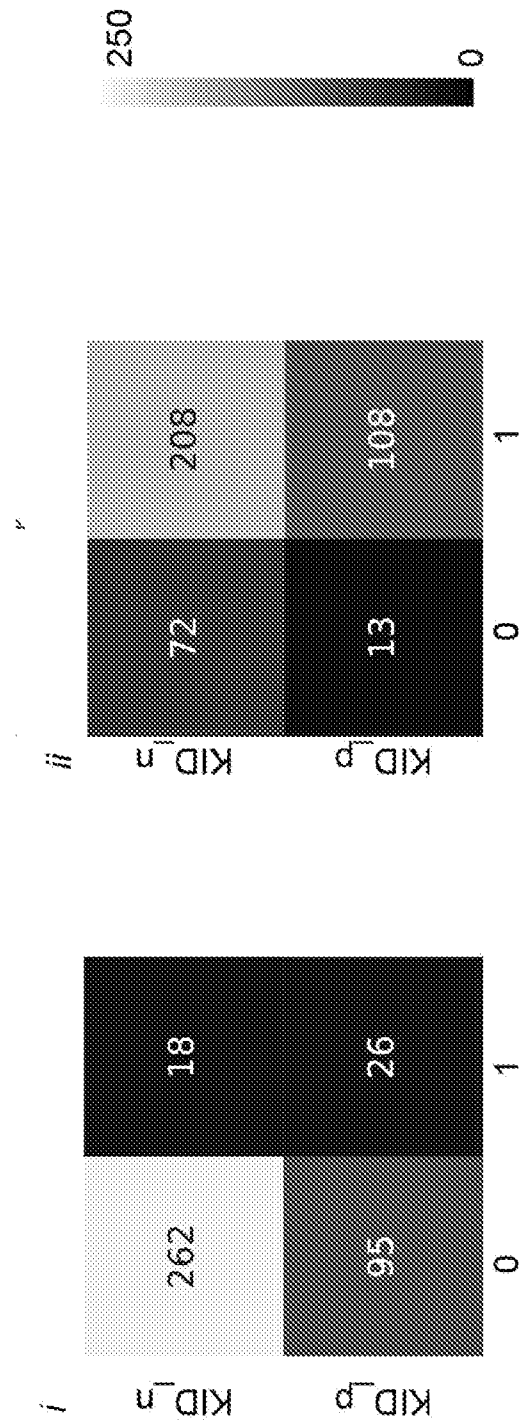
FIG. 14C
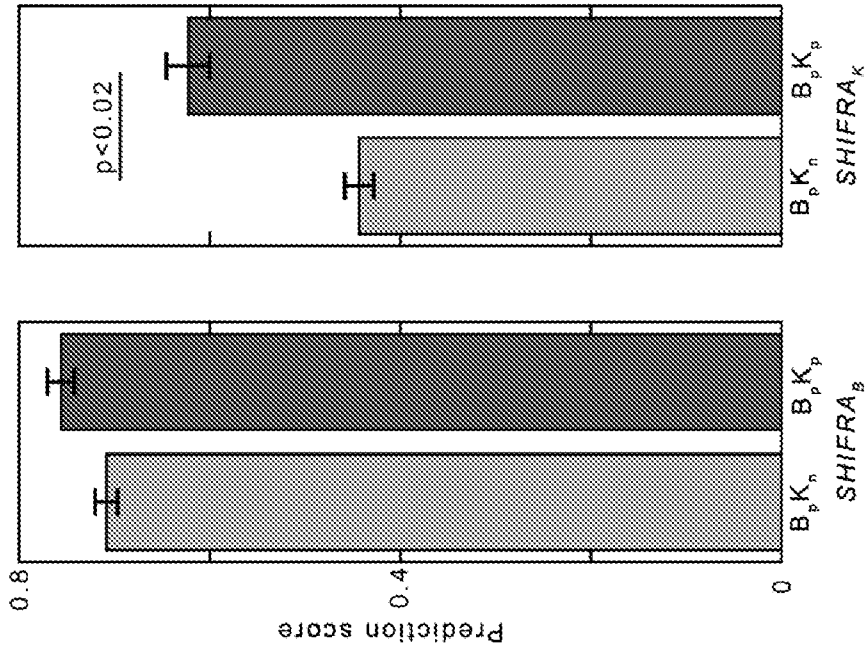
FIG. 15B
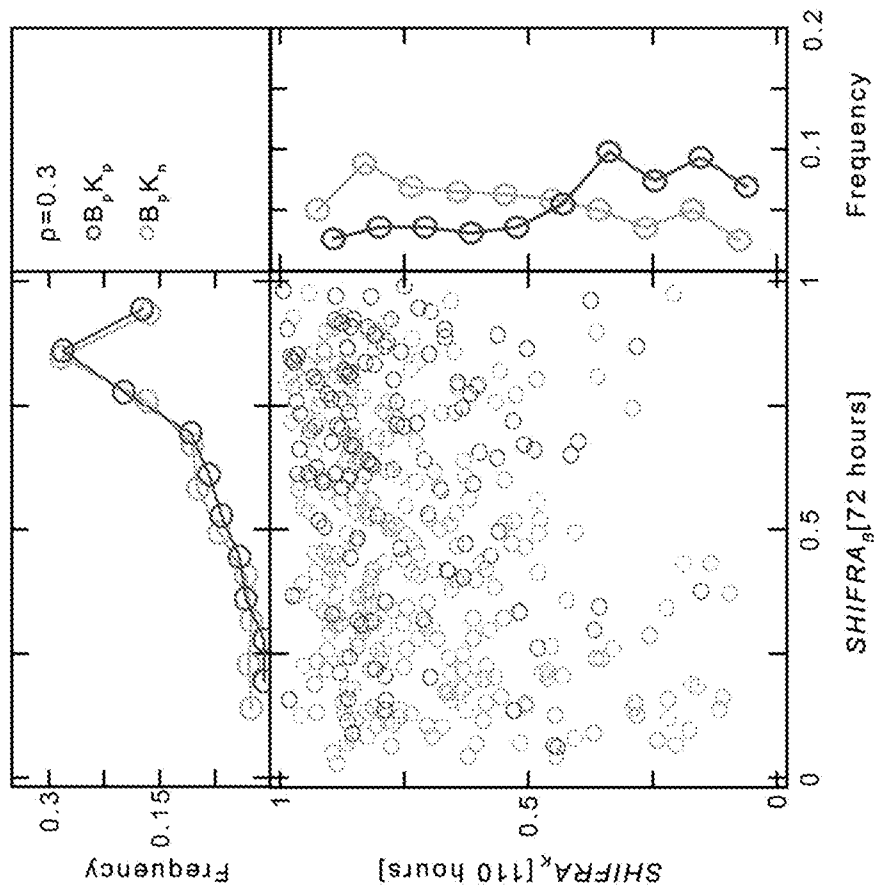
FIG. 15A
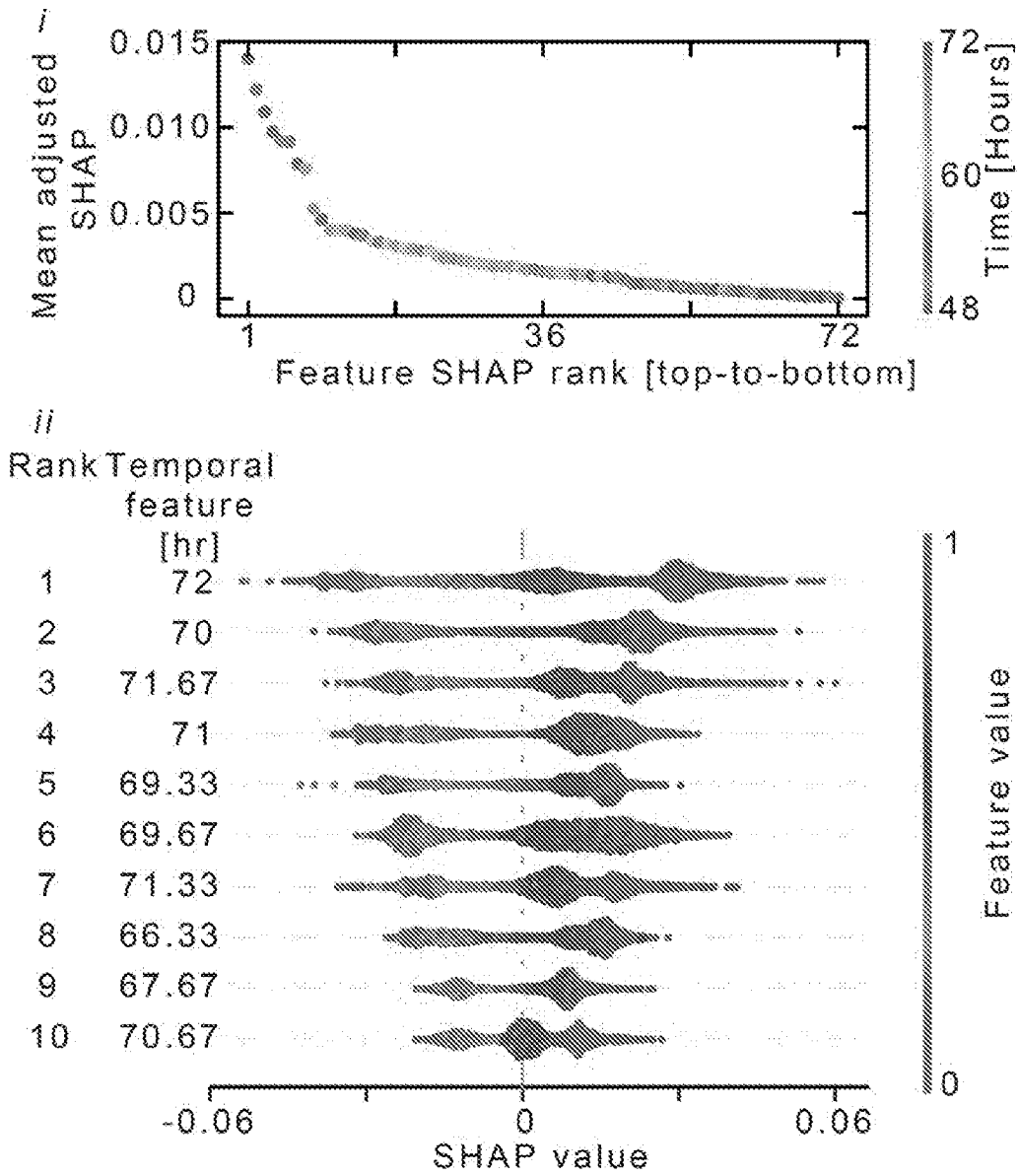
FIG. 16A
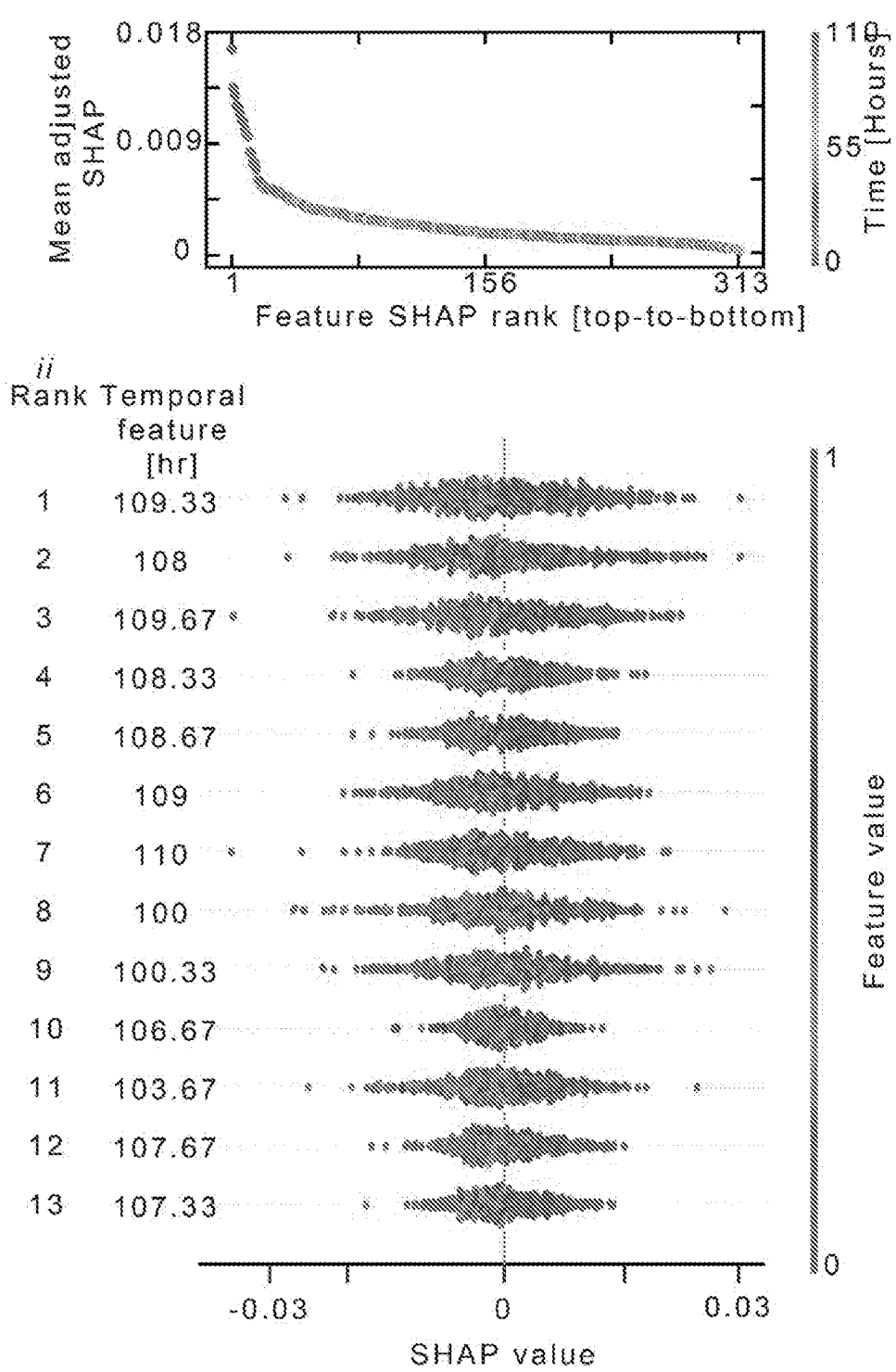
FIG. 16B
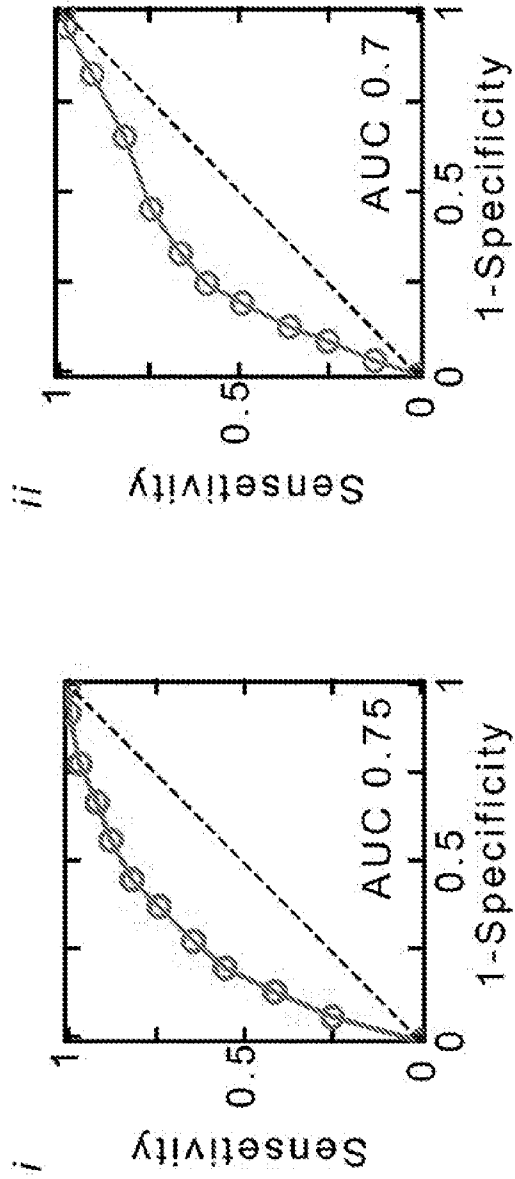
FIG. 16C
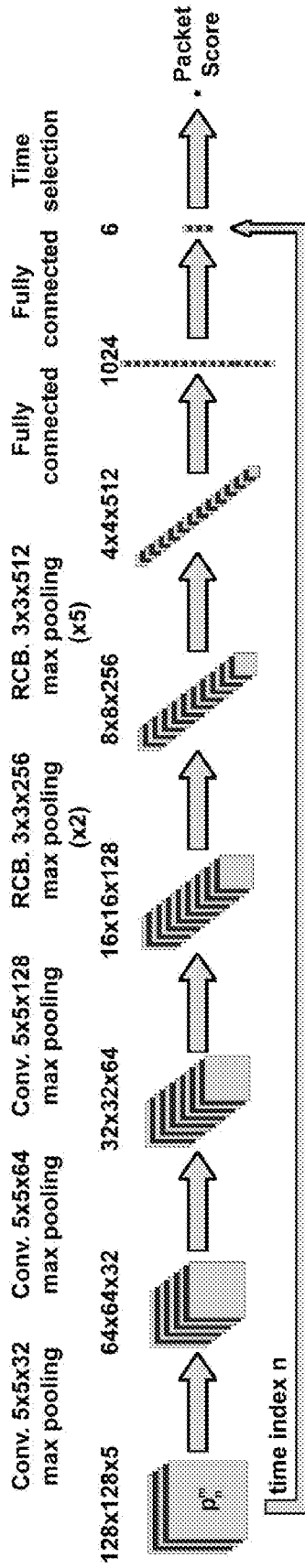
FIG. 17
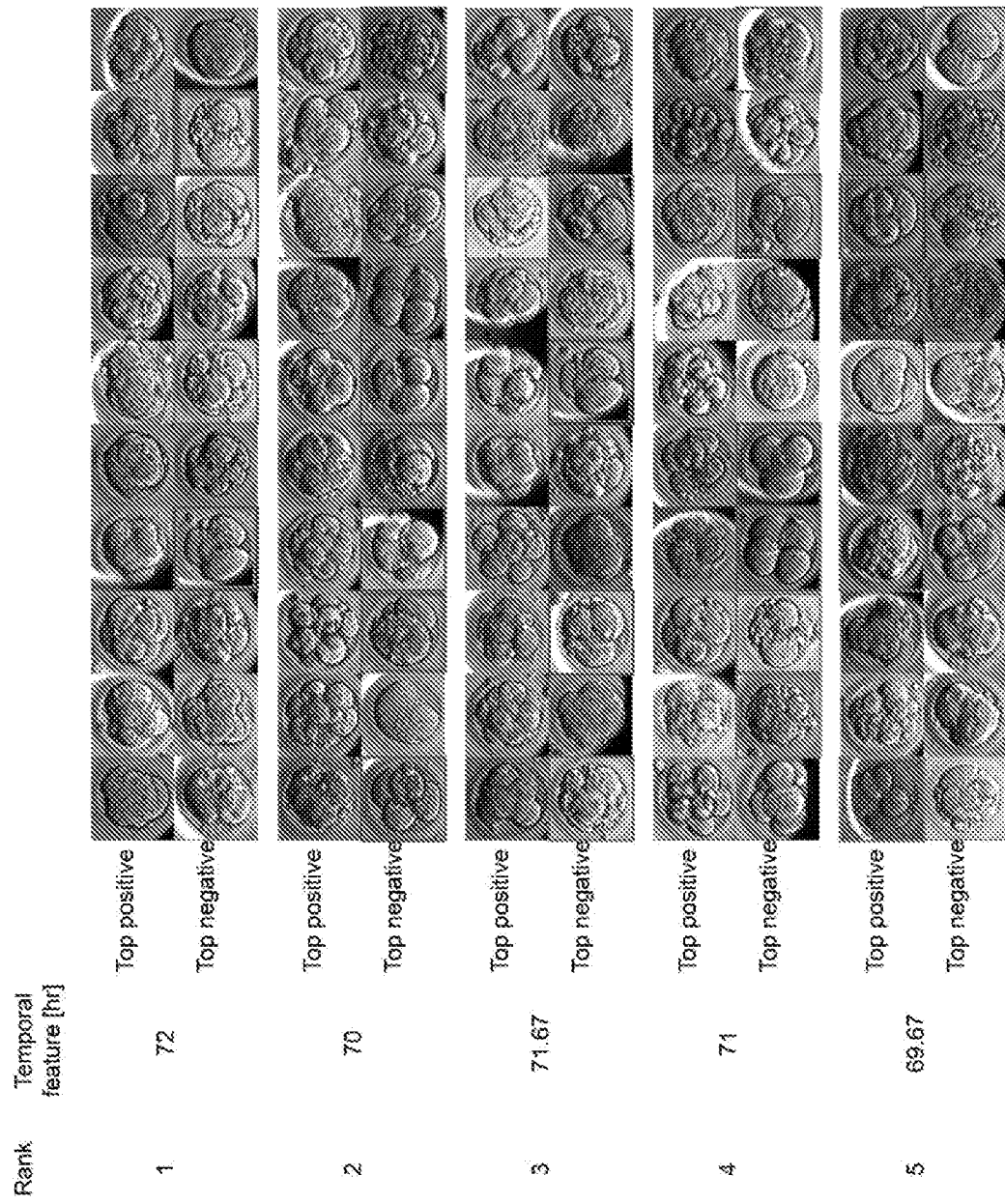
FIG. 18A (continued in 18B)
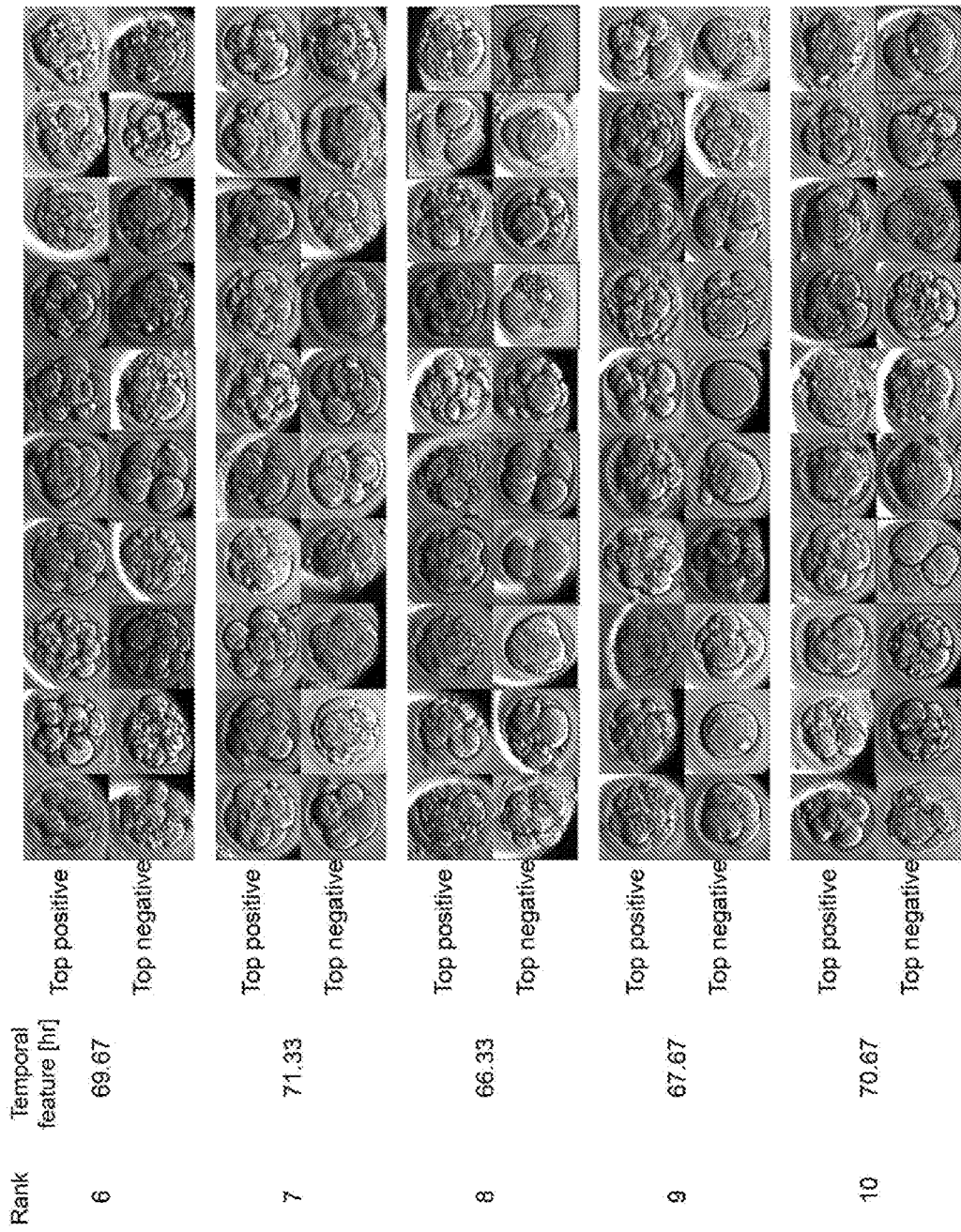
FIG. 18B (continued from 18A)

Rank	Temporal feature [hr]	
1	109.33	
2	108	
3	109.67	
4	108.33	
5	108.67	
6	109	

FIG. 19A (continued in19)

Rank	Temporal feature [hr]		
7	110	Top positive / Top negative	
8	100	Top positive / Top negative	
9	100.33	Top positive / Top negative	
10	106.67	Top positive / Top negative	
11	103.67	Top positive / Top negative	
12	107.67	Top positive / Top negative	
13	107.33	Top positive / Top negative	
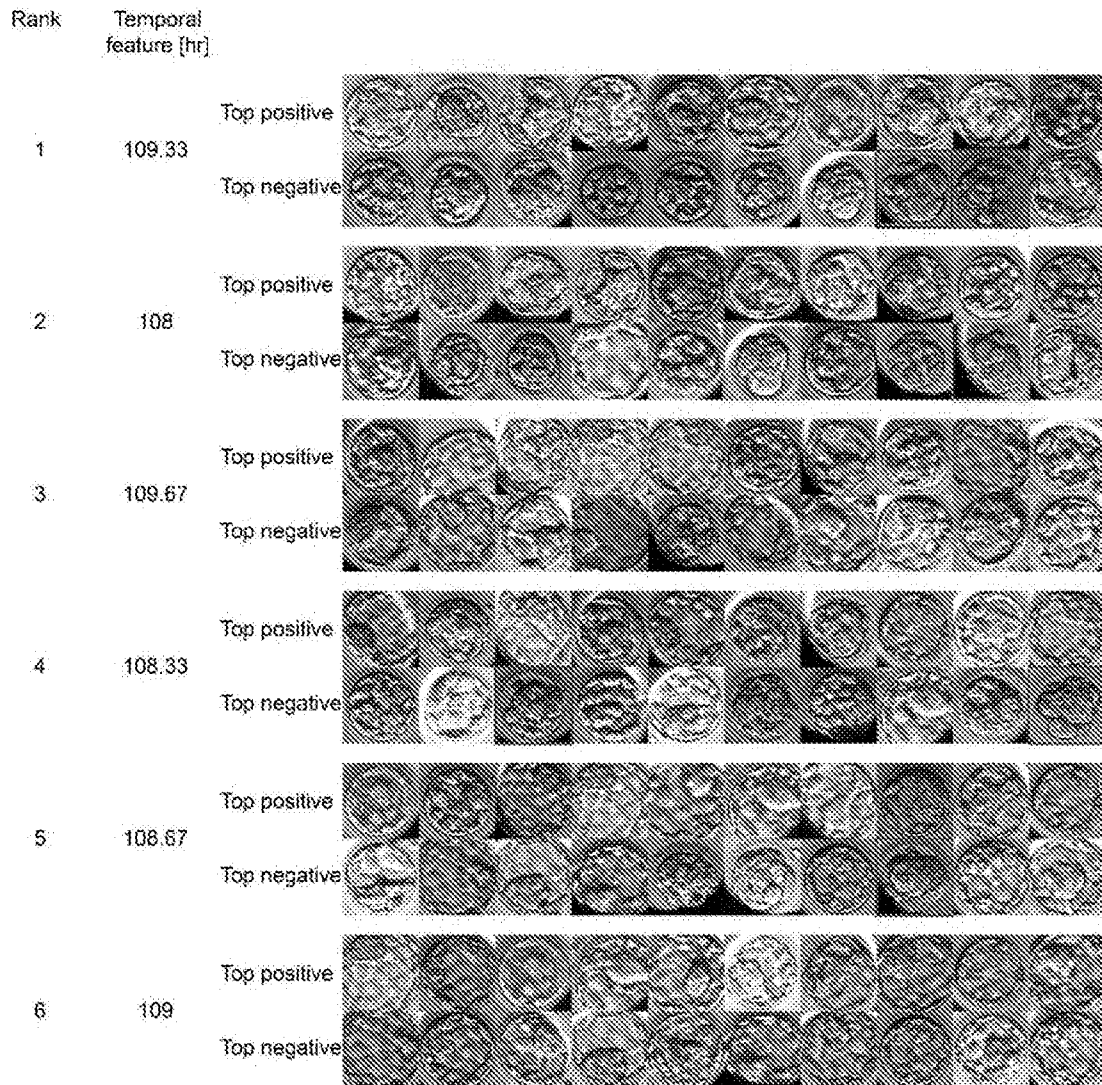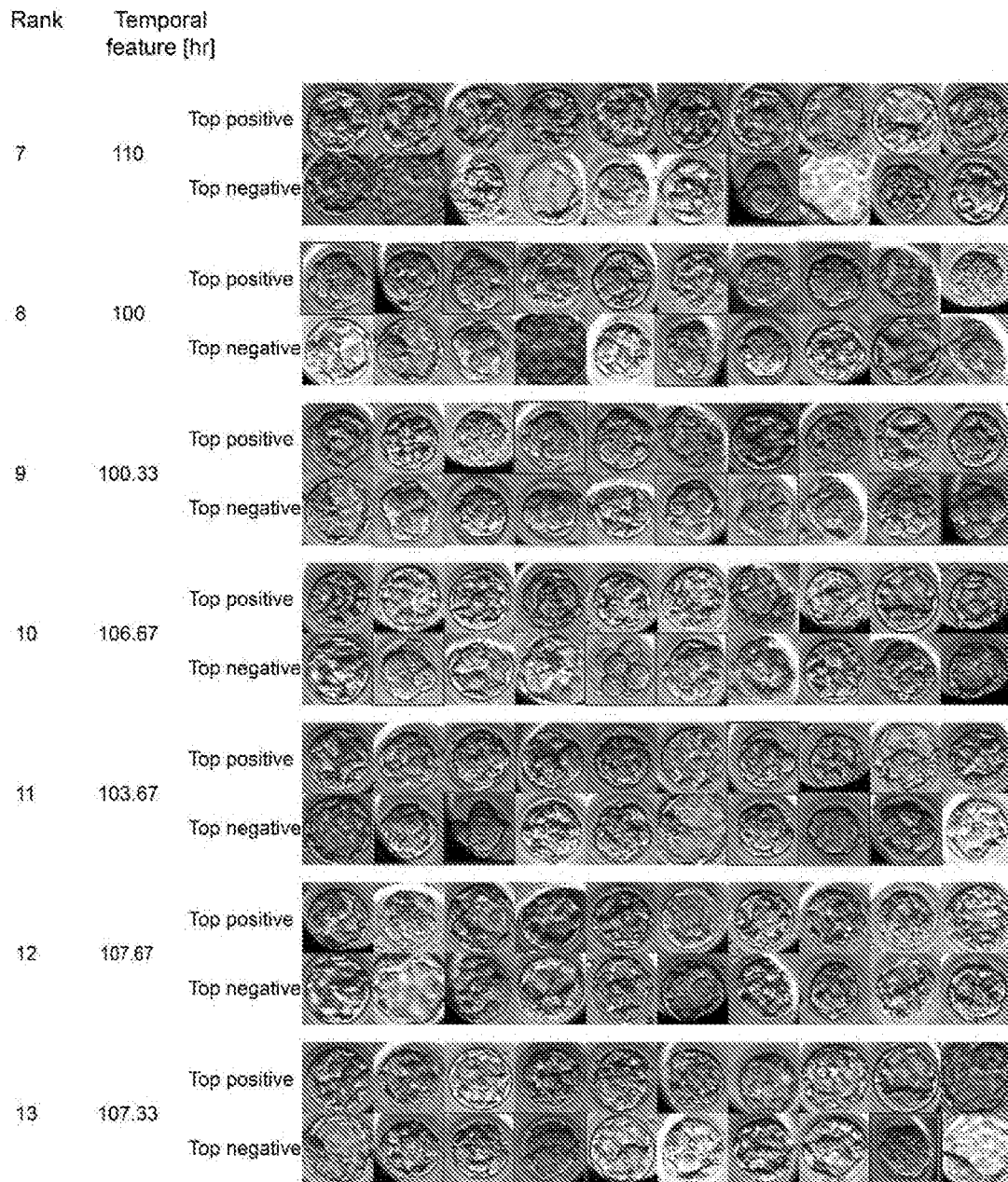
FIG. 19B (continued from 19A)
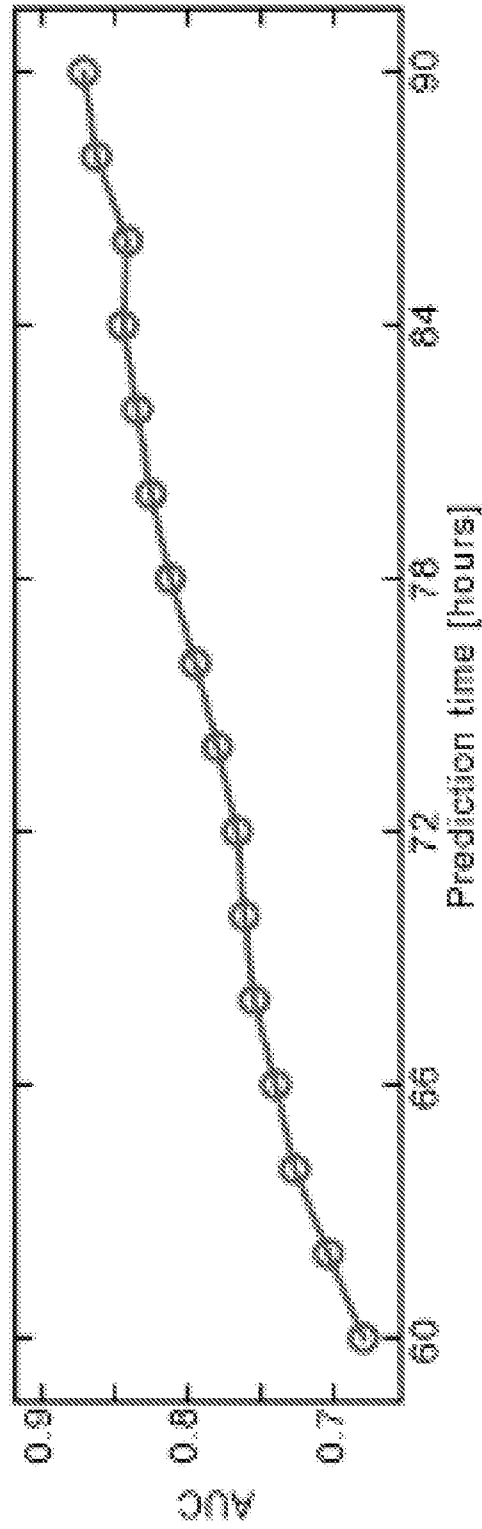
FIG. 20A
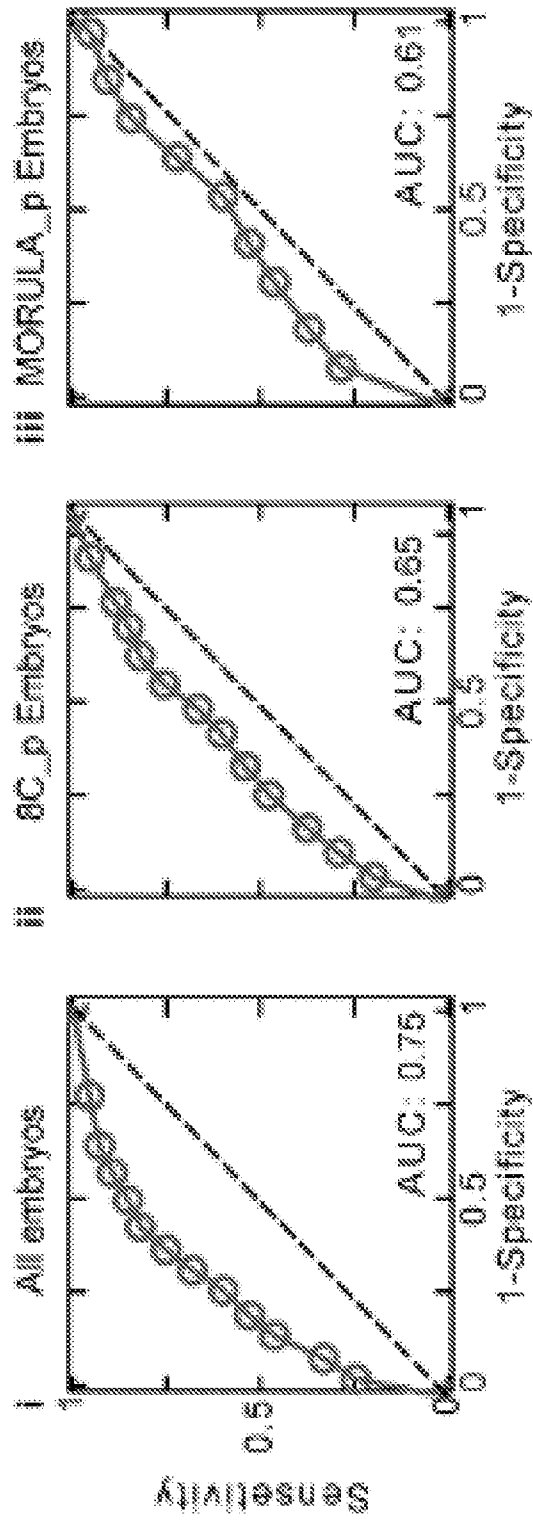
FIG. 20B
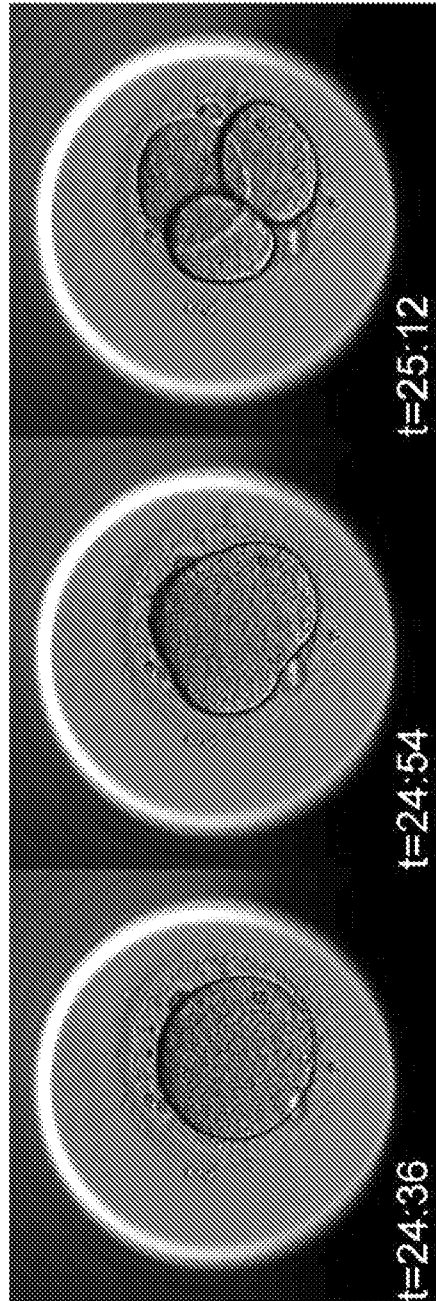
FIG. 21A
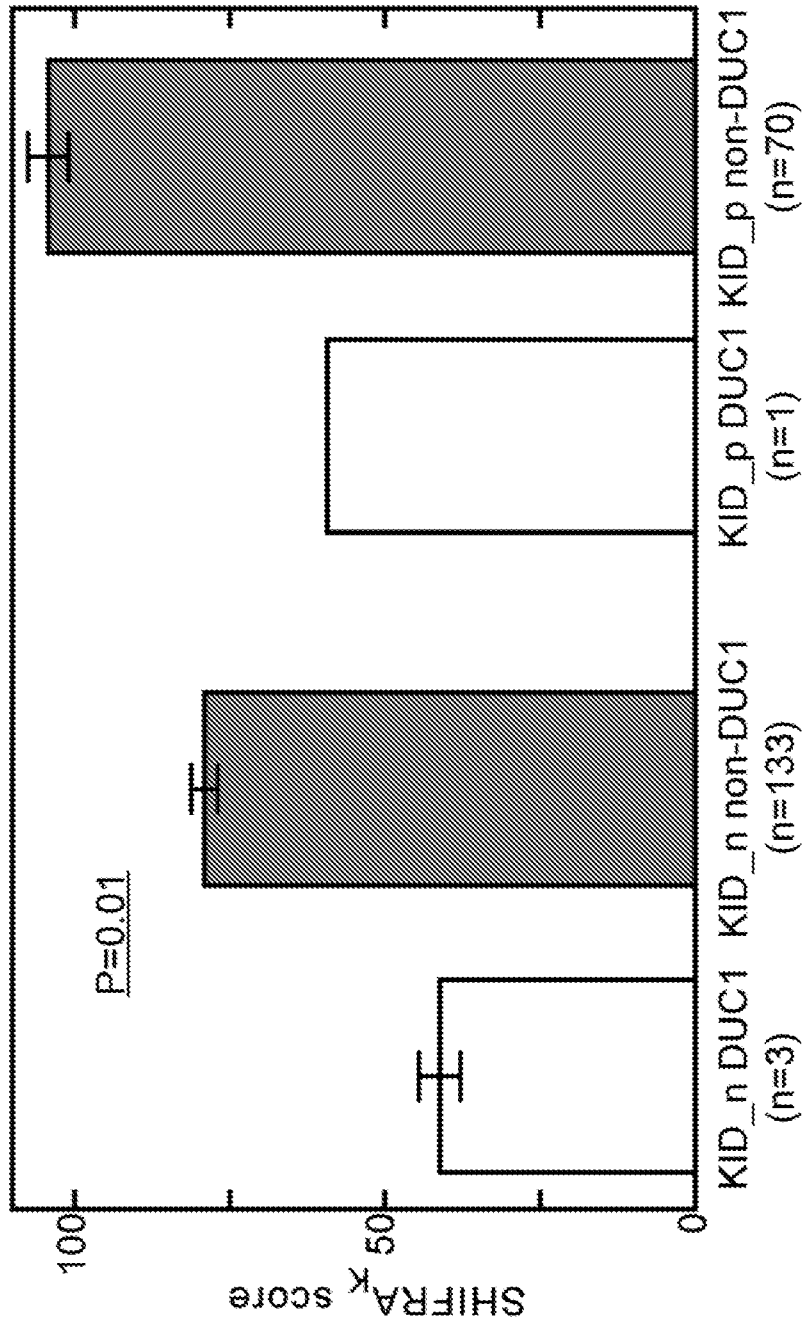
FIG. 21B
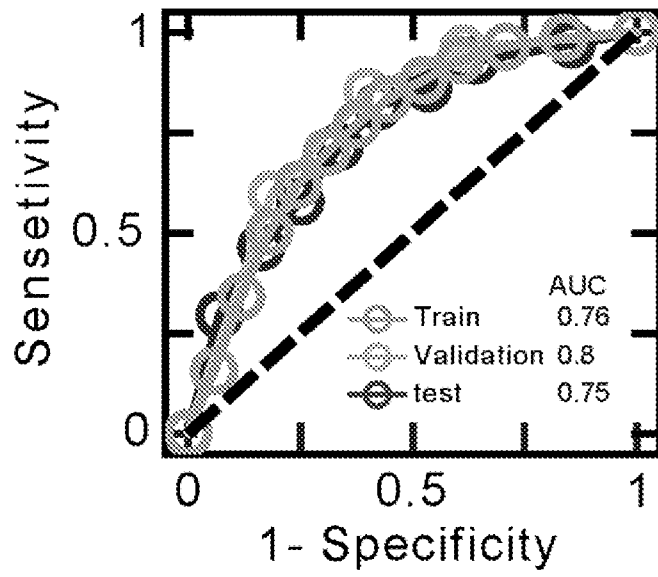
FIG. 22A
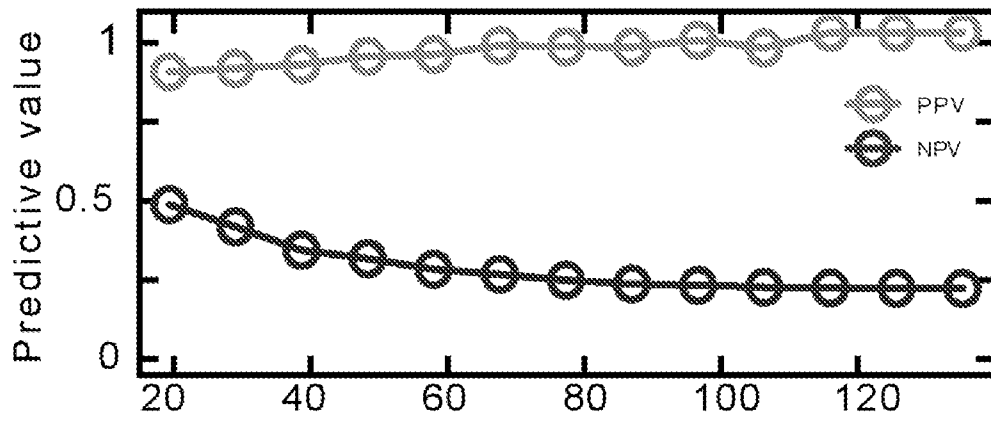
FIG. 22B
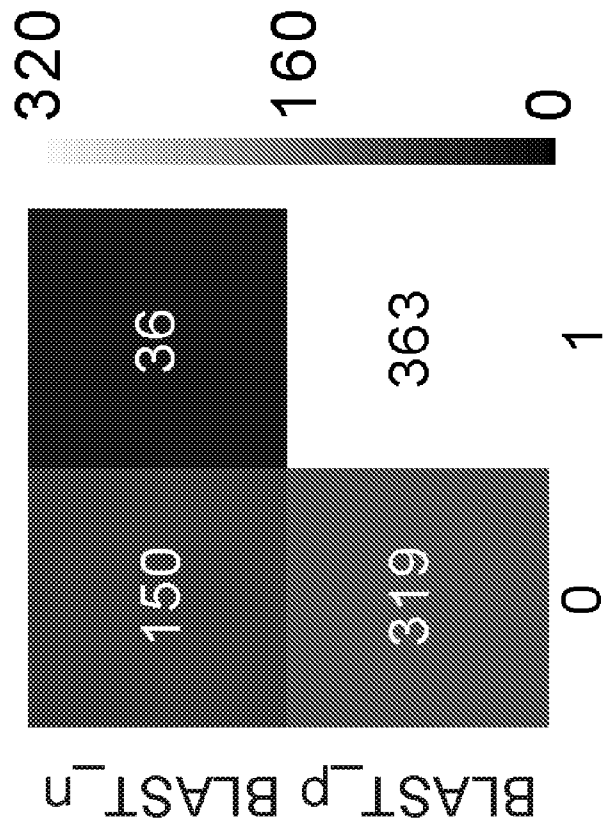
FIG. 22C
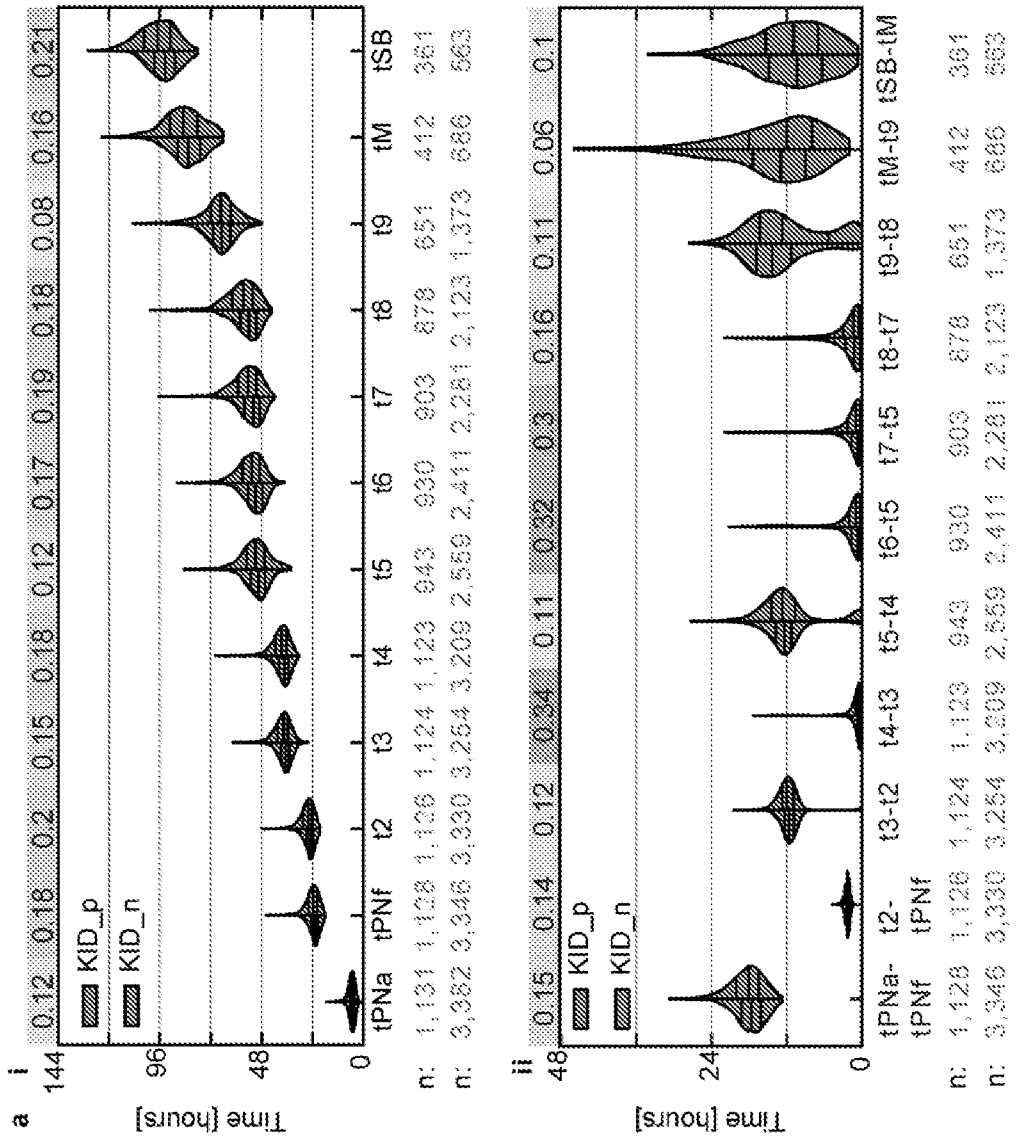
FIG. 23A
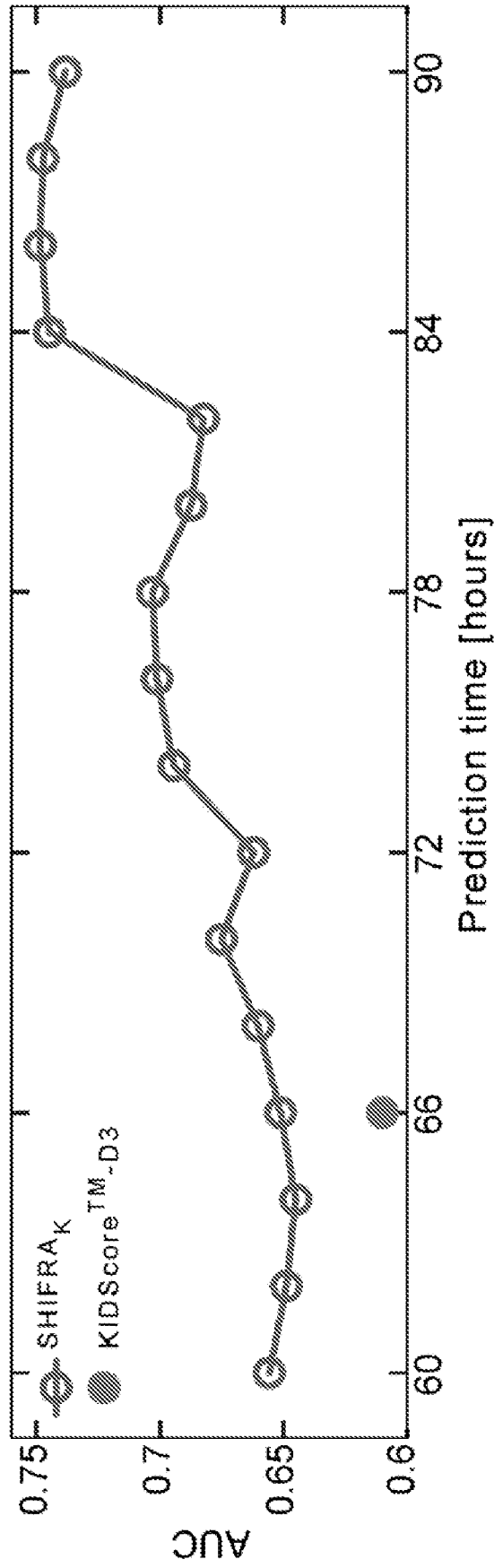
FIG. 23B
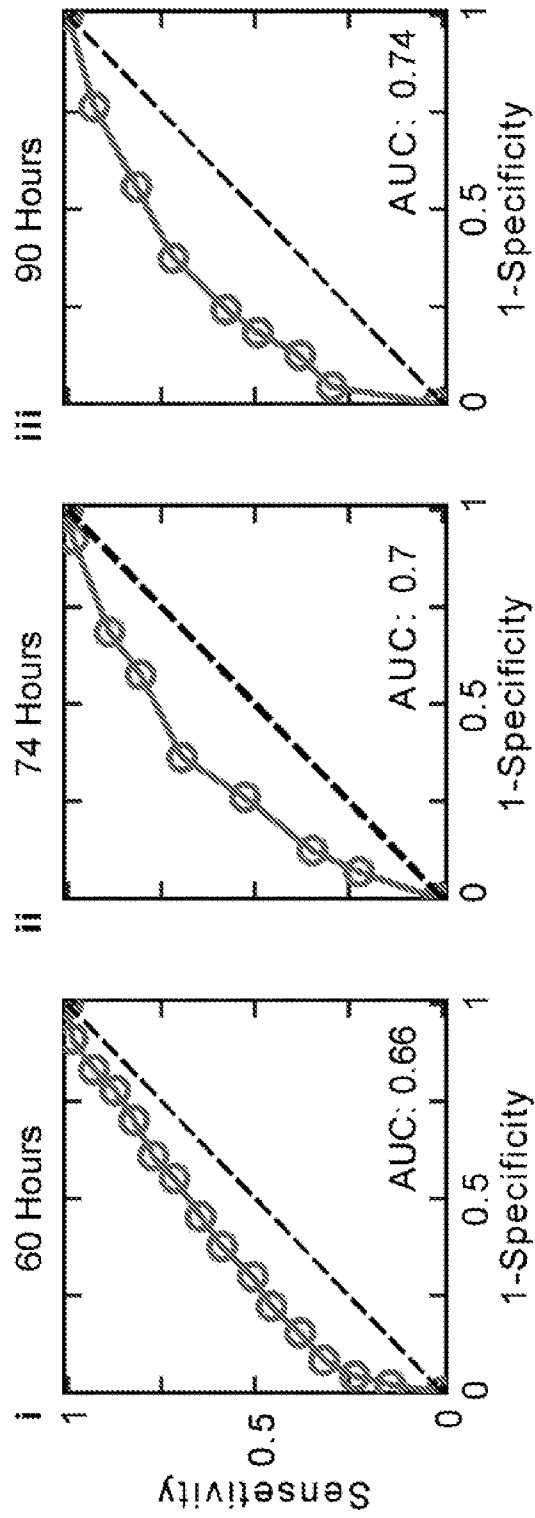
FIG. 23C
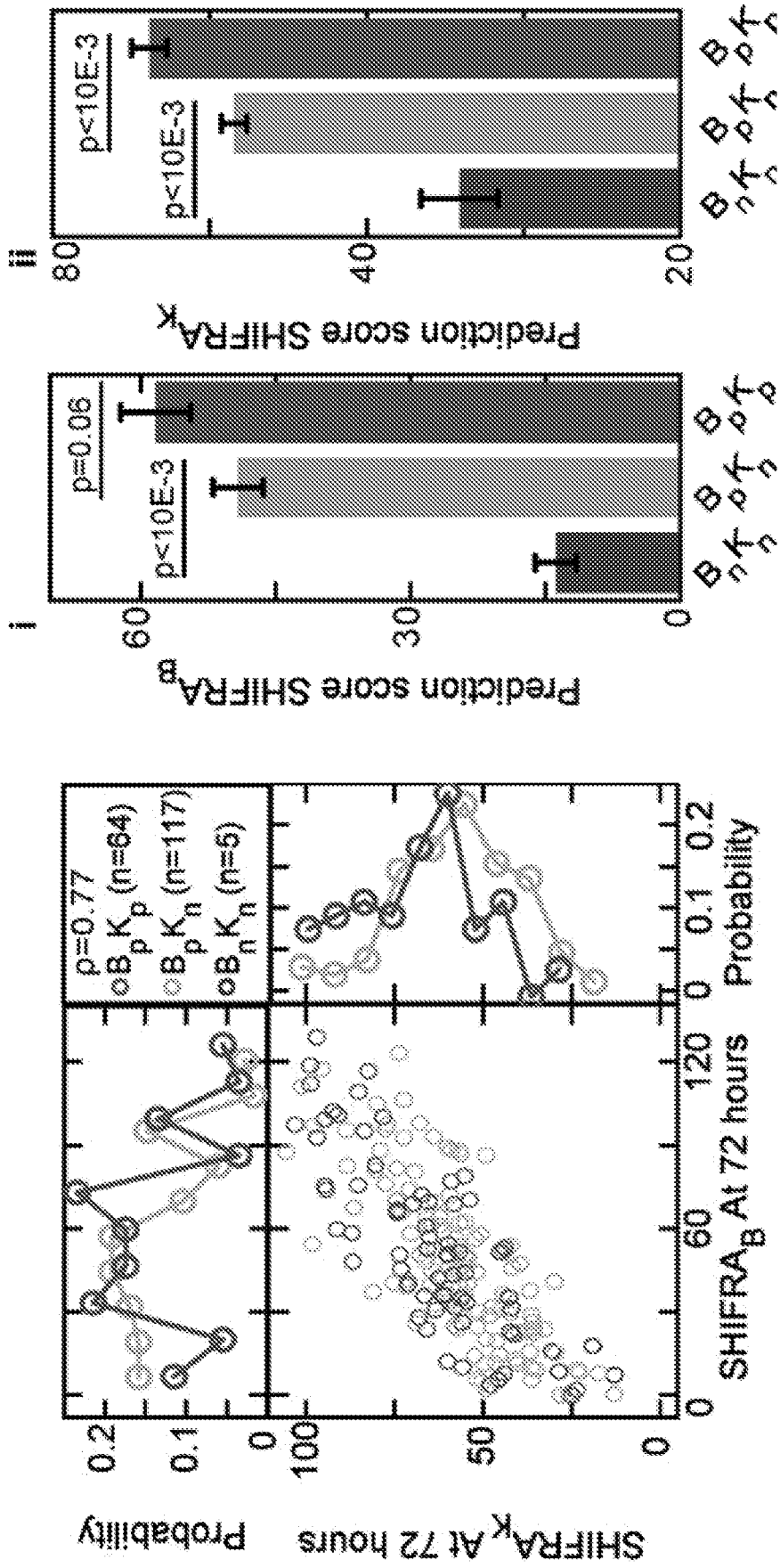
FIG. 24B

FIG. 24A
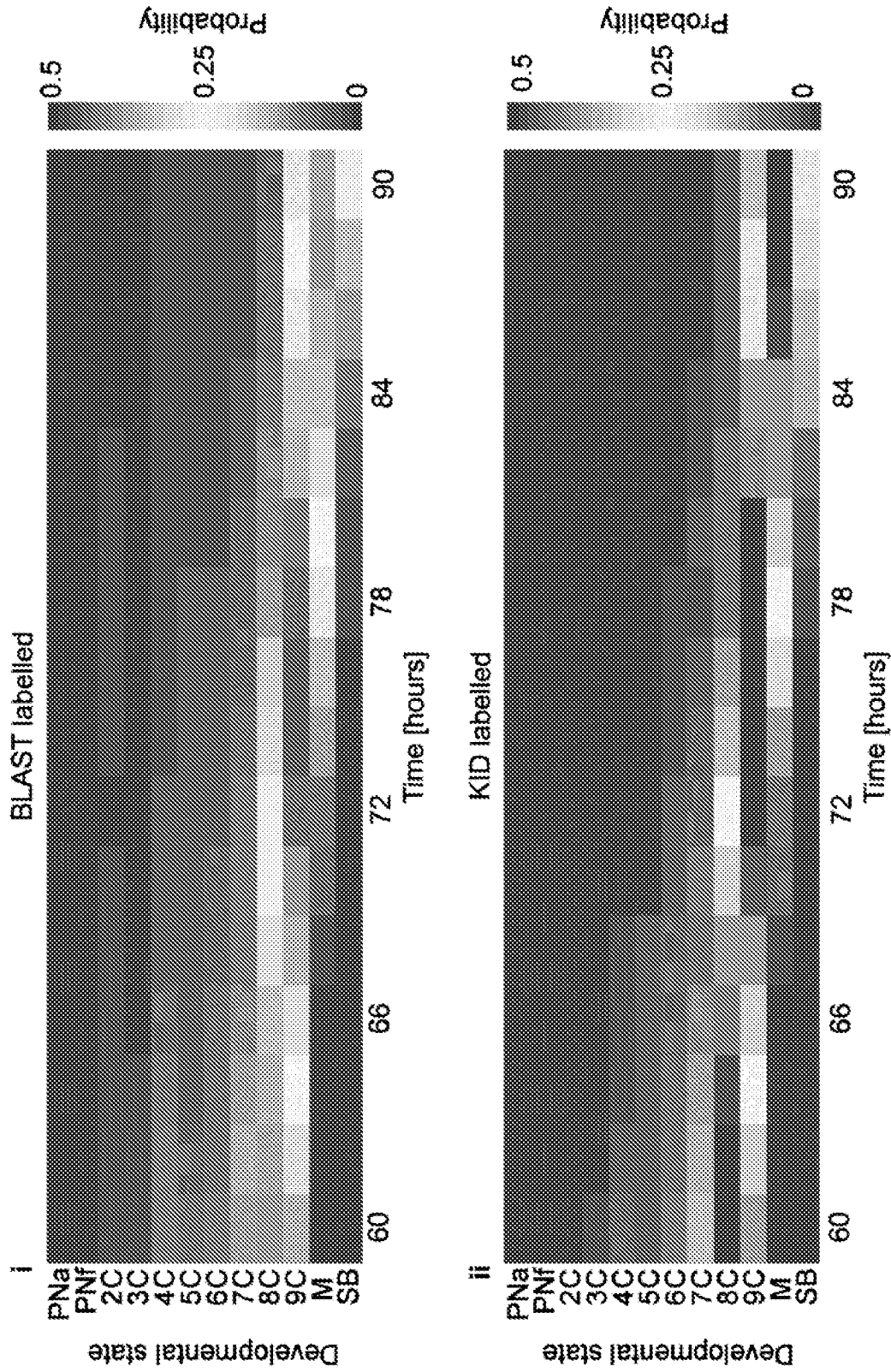
FIG. 24C
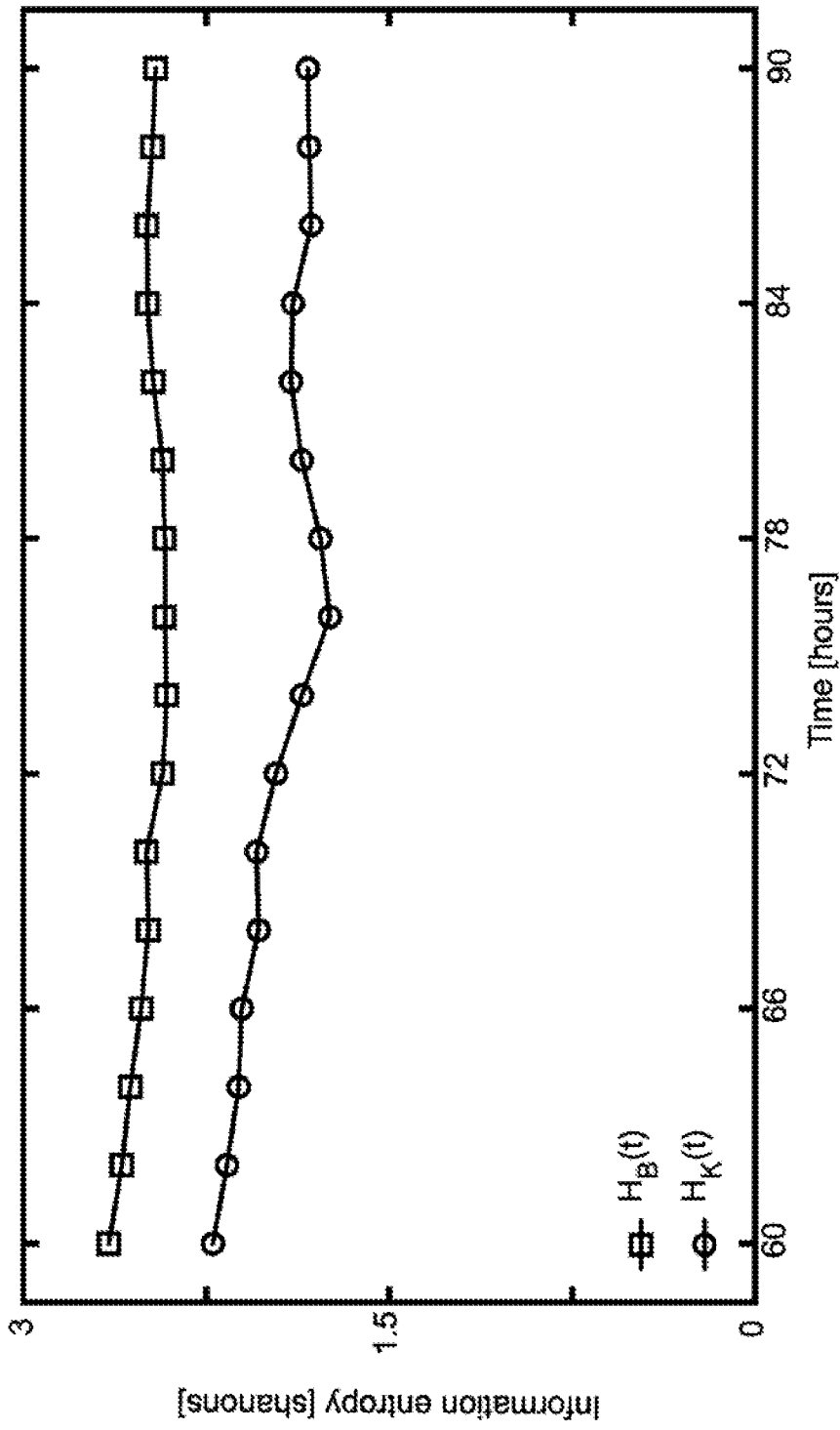
FIG. 24D
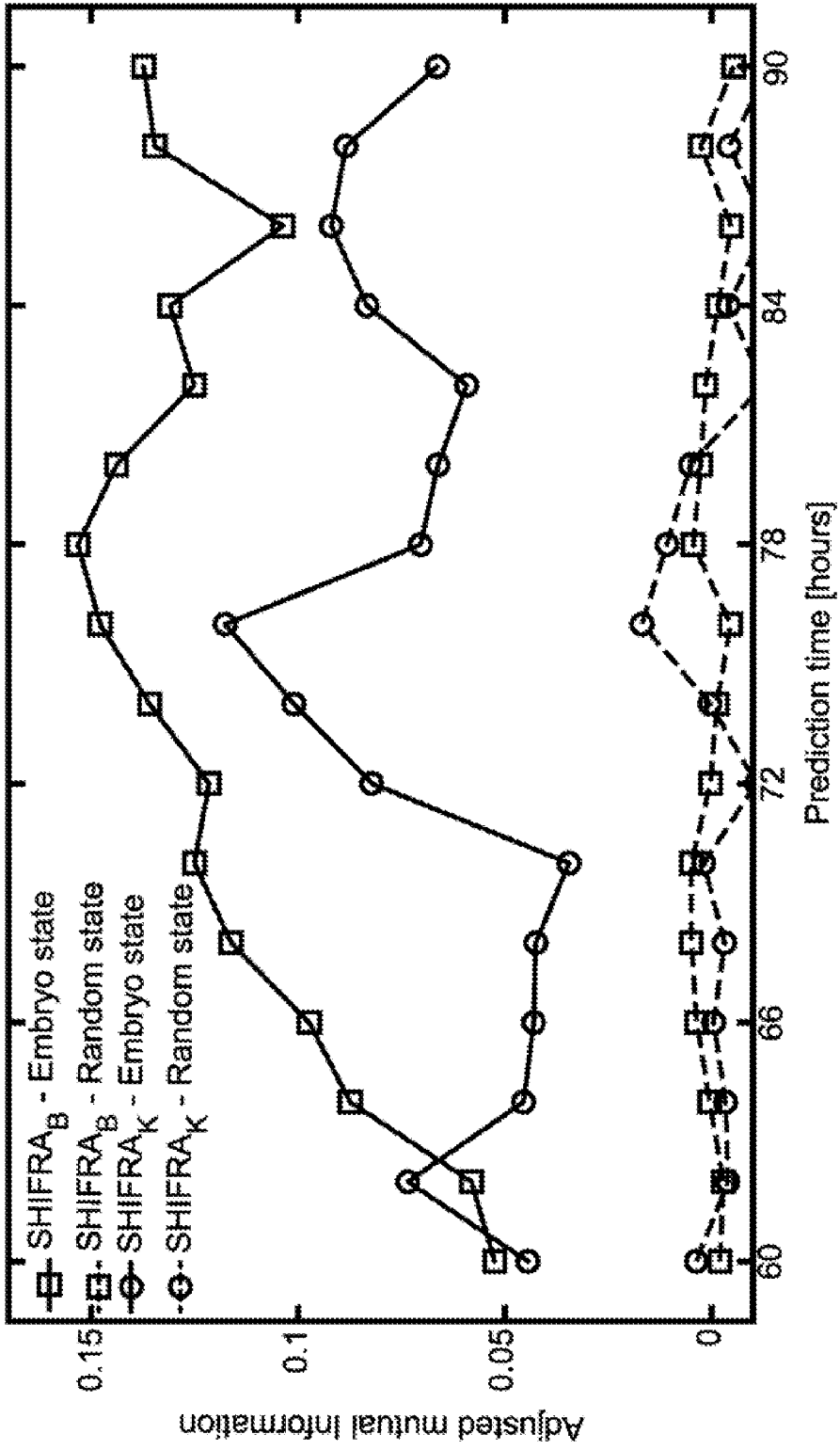
FIG. 24E
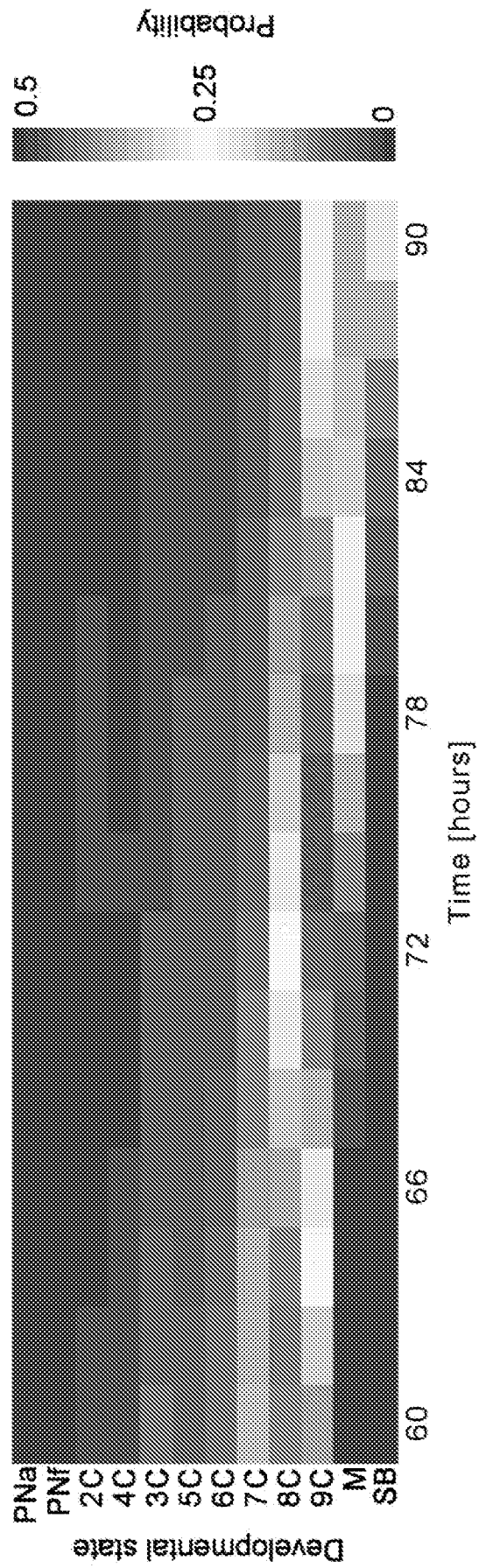
FIG. 25A
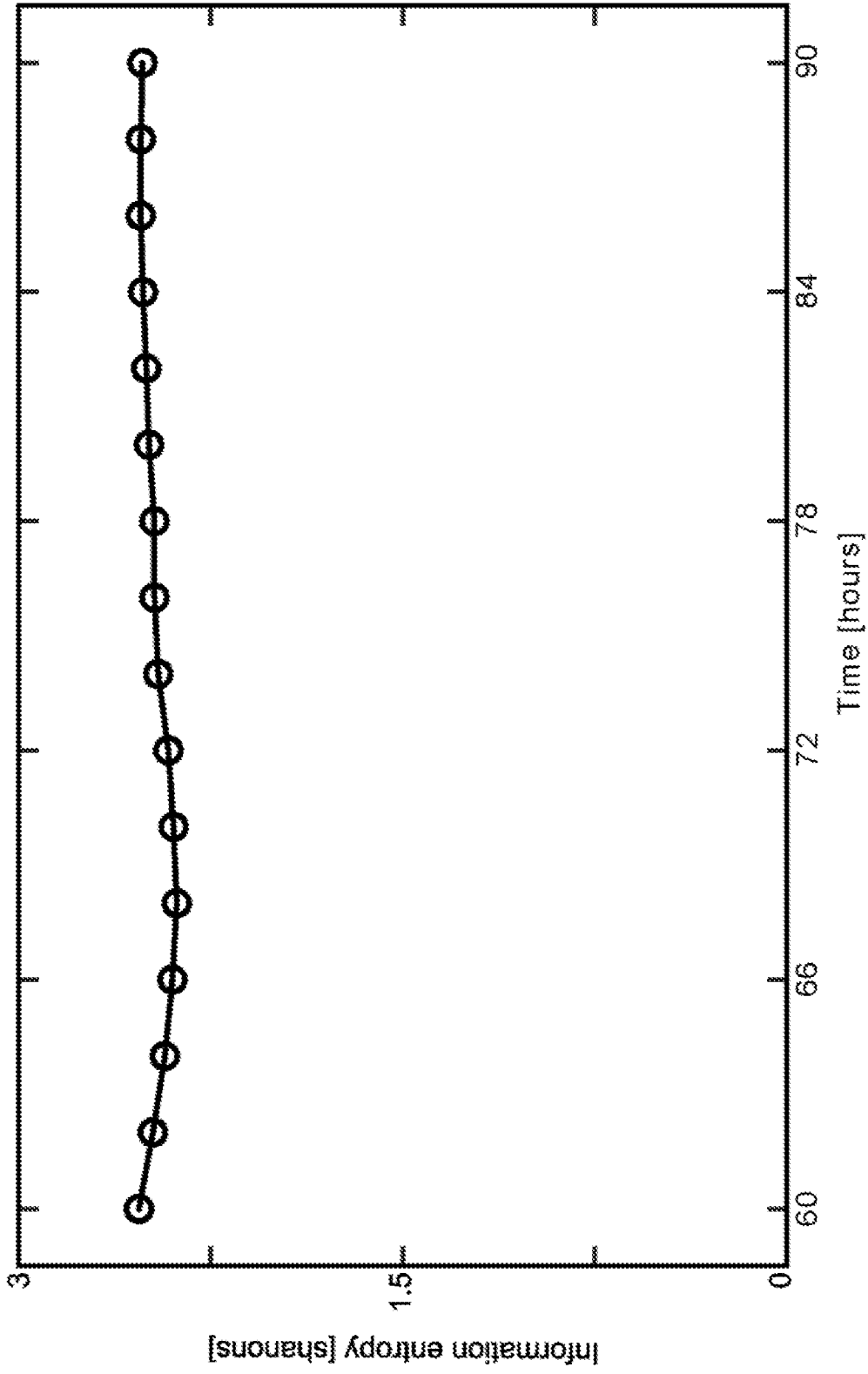
FIG. 25B
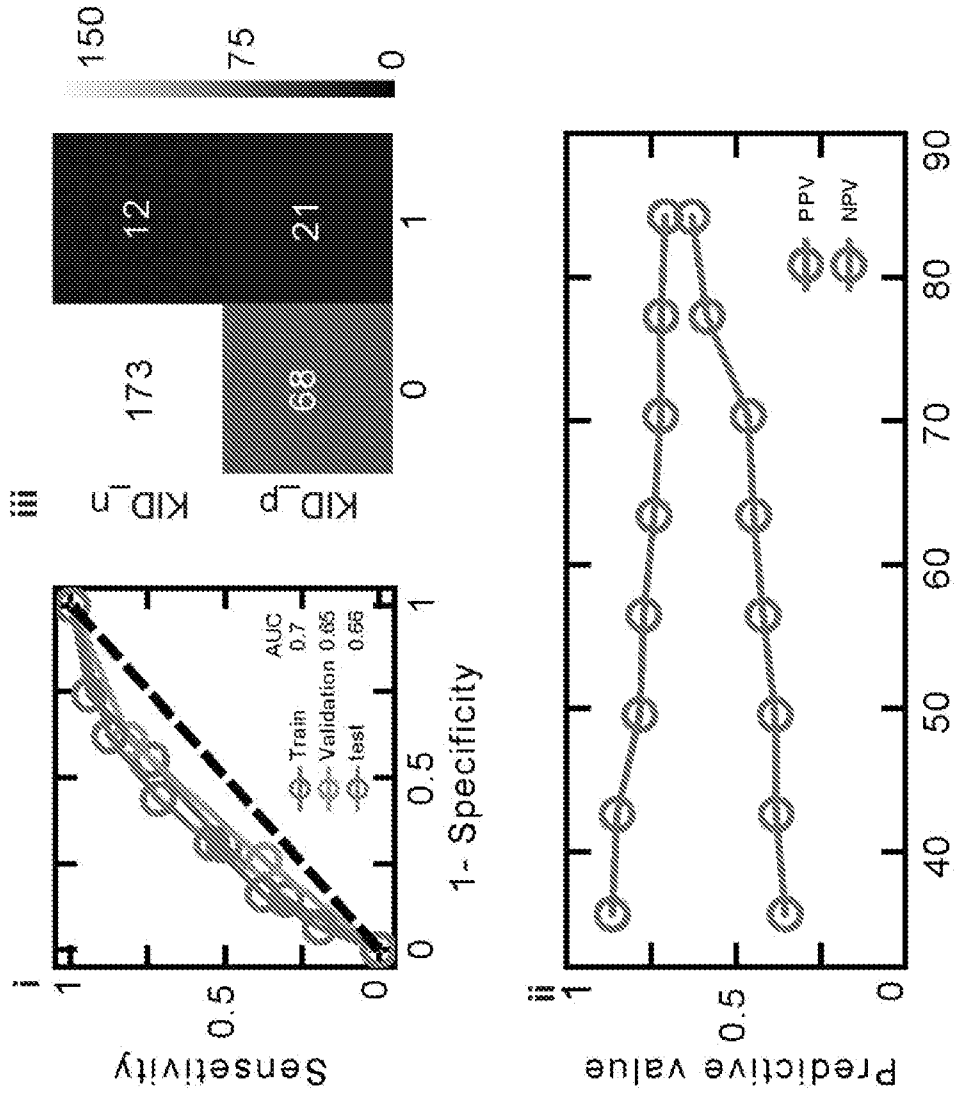
FIG. 26A
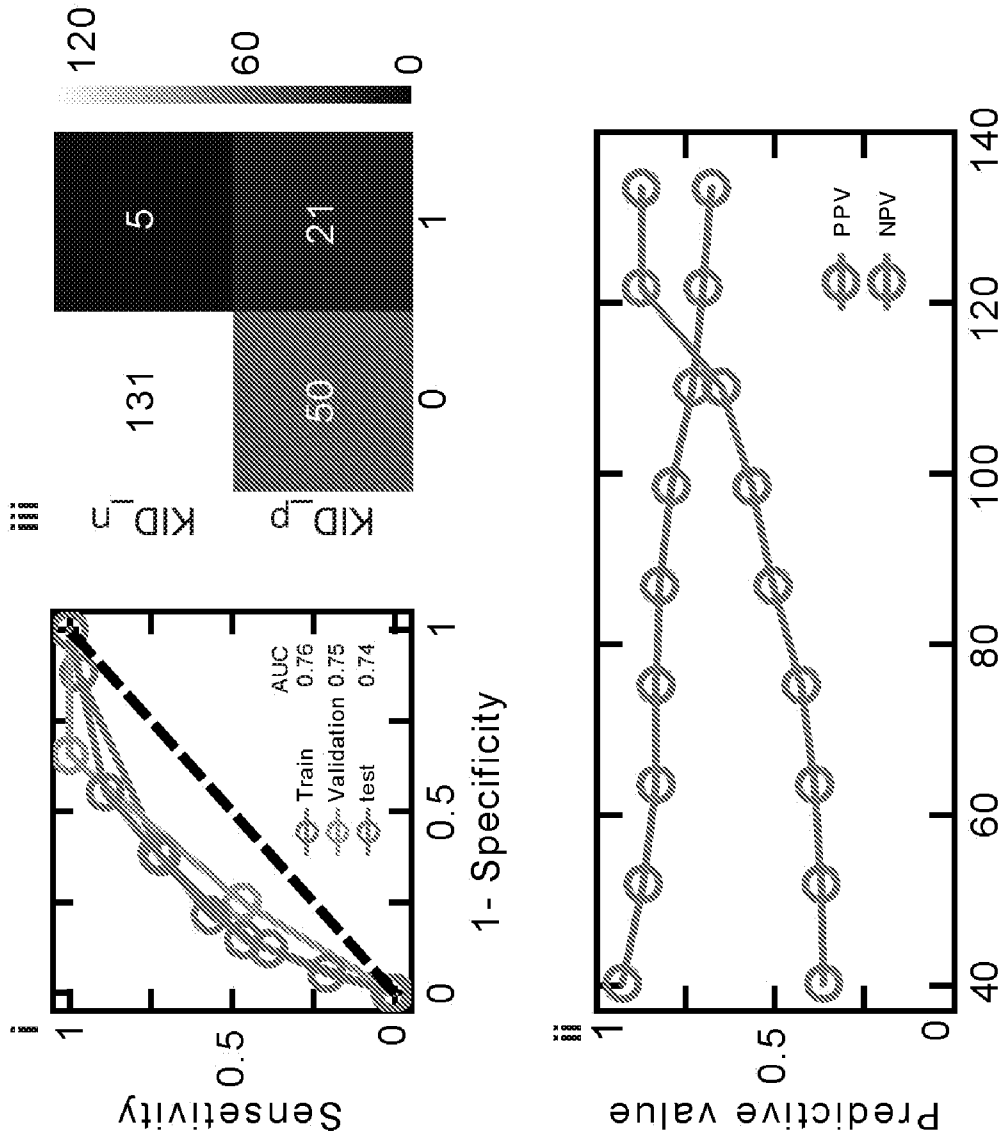
FIG. 26B

KIDScore	5	4	3	2	1
Sensitivity	0.85	0.89	0.92	0.97	0.99
	0	4	3	2	1
Specificity	0.39	0.25	0.17	0.06	0.04
	1				0

FIG. 27
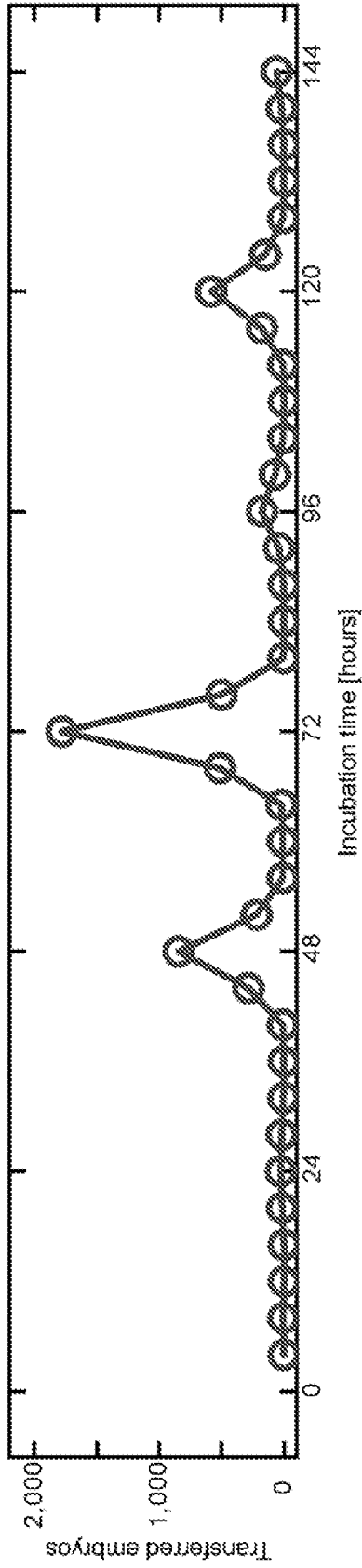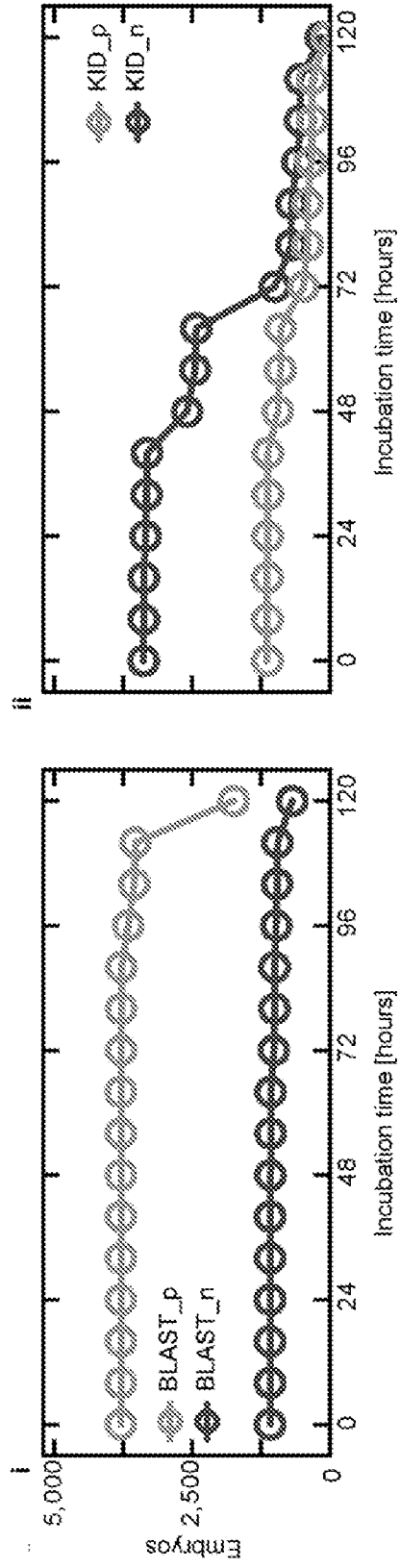
FIG. 28A

FIG. 28B
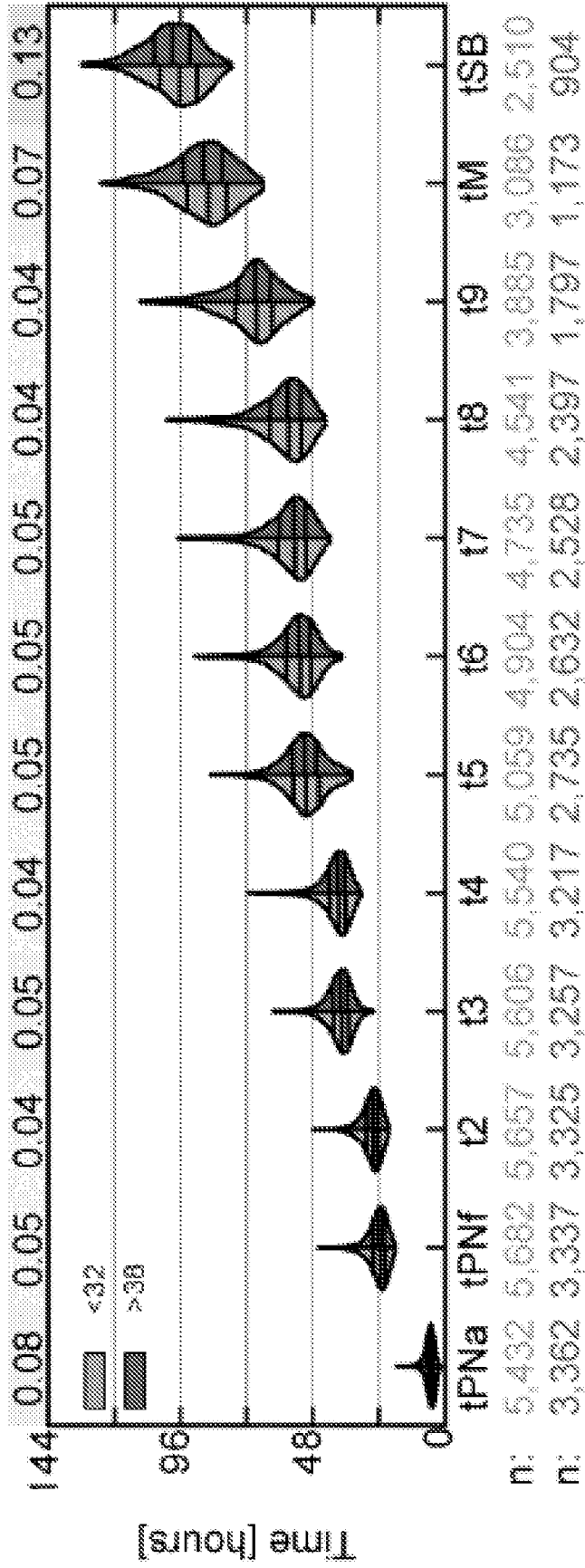
FIG. 29A
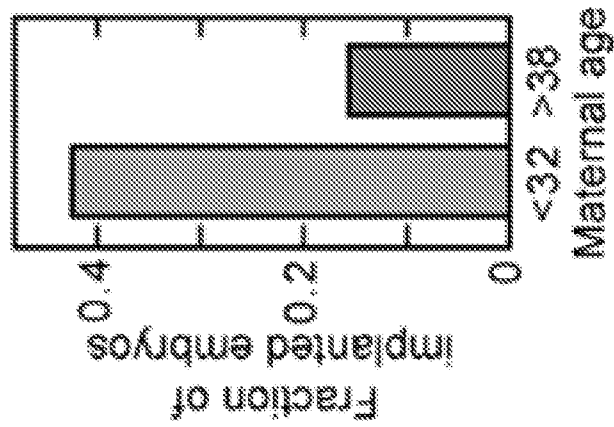
FIG. 29D
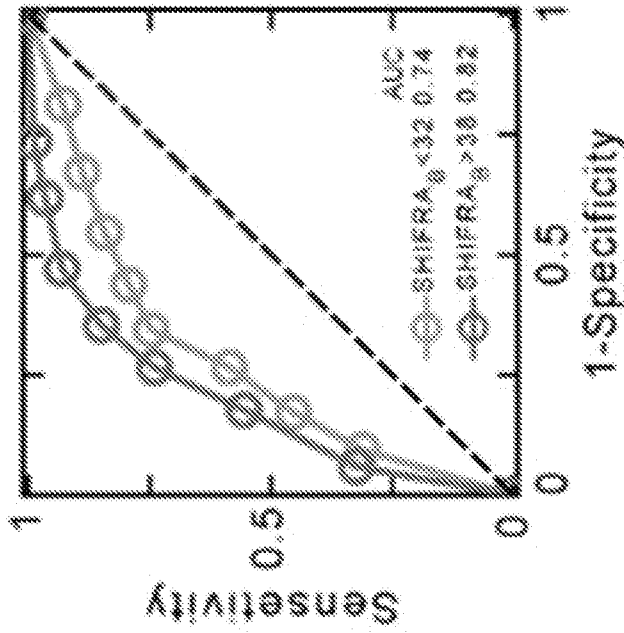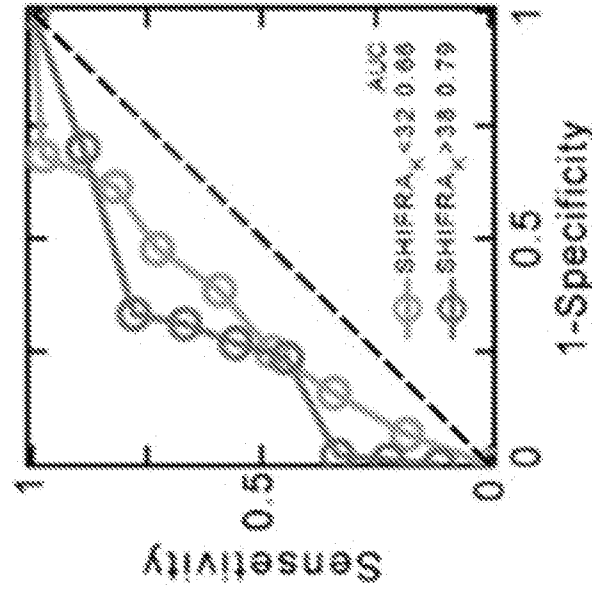
FIG. 29C

FIG. 29B

FIG. 30B

FIG. 30A

FIG. 30D

FIG. 30C

A.	CLASSIFICATION OF SUBJECT MATTER

IPC (20200101) G06N 20/00, G06K 9/00, G06T 7/00
CPC (20190101) G06N 20/00, G06K 9/00147, G06T 7/0014

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 (20200101) G06N 20/00, G06K 9/00, G06T 7/00
CPC (20190101) G06N 20/00, G06K 9/00147, G06T 7/0014

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

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

Databases consulted: Esp@cenet, Google Patents, Google Scholar, Orbit
Search terms used: embryo, development, machine learning, video segments

C.	DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 9710696 B2 Wang et al. 18 Jul 2017 (2017/07/18) the whole document	1-30
A	WO 2018179769 A1 Shinoda et al. 04 Oct 2018 (2018/10/04) the whole document	1-30
A	CN 109214437 A Lin Kai et al. 15 Jan 2019 (2019/01/15) the whole document	1-30

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

*	Special categories of cited documents:

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

"D" document cited by the applicant in the international application

"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
26 Apr 2020	26 Apr 2020

Name and mailing address of the ISA: Israel Patent Office Technology Park, Bldg.5, Malcha, Jerusalem, 9695101, Israel Email address: pctoffice@justice.gov.il	Authorized officer ROASH Yoaela Telephone No. 972-73-3927256

Form PCT/ISA/210 (second sheet) (revised January 2019)

Patent document cited search report	Publication date	Patent family member(s)	Publication Date
US 9710696 B2	18 Jul 2017	US 2016078275 A1	17 Mar 2016
		US 9710696 B2	18 Jul 2017
		CA 2901830 A1	04 Sep 2014
		CA 2901833 A1	04 Sep 2014
		CN 105407819 A	16 Mar 2016
		CN 105408746 A	16 Mar 2016
		EP 2961336 A1	06 Jan 2016
		EP 2961336 A4	02 Nov 2016
		EP 2962101 A1	06 Jan 2016
		EP 2962101 A4	19 Oct 2016
		IL 240855 D0	29 Oct 2015
		IL 240855 A	31 Jan 2019
		IL 240856 D0	29 Oct 2015
		IL 263963 D0	31 Jan 2019
		JP 2016509844 A	04 Apr 2016
		JP 2016509845 A	04 Apr 2016
		JP 2020036625 A	12 Mar 2020
		US 2014247972 A1	04 Sep 2014
		US 9177192 B2	03 Nov 2015
		US 2014247973 A1	04 Sep 2014
		US 9542591 B2	10 Jan 2017
		WO 2014134527 A1	04 Sep 2014
		WO 2014134550 A1	04 Sep 2014
WO 2018179769 A1	04 Oct 2018	WO 2018179769 A1	04 Oct 2018
		AU 2018245711 A1	29 Aug 2019
		CA 3056559 A1	04 Oct 2018
		CN 110447037 A	12 Nov 2019
		EP 3590068 A1	08 Jan 2020
		JP 2018171039 A	08 Nov 2018

Patent document cited search report	Publication date	Patent family member(s)		Publication Date
		US 2020110924 A1		09 Apr 2020
CN 109214437 A	15 Jan 2019	CN 109214437 A		15 Jan 2019