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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 December 2018 (27.12.2018)

WIPO | PCT

(10) International Publication Number
WO 2018/234549 A1

(51) International Patent Classification:
C12Q 1/6851 (2018.01) *C12Q 1/6883* (2018.01)
C12Q 1/6876 (2018.01)

(21) International Application Number:
PCT/EP2018/066771

(22) International Filing Date:
22 June 2018 (22.06.2018)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
17177735.2 23 June 2017 (23.06.2017) EP
17204951.2 01 December 2017 (01.12.2017) EP

(71) Applicant: **CHARITÉ UNIVERSITÄTSMEDIZIN BERLIN** [DE/DE]; Charitéplatz 1, 10117 Berlin (DE).

(72) Inventors: **KRAMER, Achim**; Goldenes Horn 62, 12107 Berlin (DE). **WITTENBRINK, Nicole**; Wichertstr. 39, 10439 Berlin (DE). **ANANTHASUBRAMANIAM, Bharath**; Ebelingstr. 4, 10249 Berlin (DE).

(74) Agent: **JUNGHANS, Claas**; Schulz Junghans Patentanwälte PartGmbB, Großbeerenstraße 71, 10963 Berlin (DE).

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

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

Declarations under Rule 4.17:
— *as to the identity of the inventor (Rule 4.17(i))*
— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

Published:
— *with international search report (Art. 21(3))*
— *with sequence listing part of description (Rule 5.2(a))*

WO 2018/234549 A1

(54) Title: BODYTIME - A NEW DIAGNOSTIC TOOL TO ASSESS THE INTERNAL CLOCK

(57) Abstract: The invention relates to a method of assessing a time-related physiological parameter selected from a person's circadian rhythm, a person's internal time and the robustness/strength of a person's circadian rhythm. The method comprises the steps of determining in a sample of blood monocytes obtained from said person an expression level of a plurality of circadian oscillatory genes, and assessing the time-related physiological parameter based on the determined expression level.

BodyTime – a new diagnostic tool to assess the internal clock

Description

The present invention relates to a method to characterize a person's chronobiological state.

Background

The circadian clock is a temporal, biological program found throughout nature including humans. It regulates physiology and behavior according to time of day. The last 25 years of chronobiology research has identified many important findings. The master oscillator that regulates behavior and hormone profiles (e.g. melatonin) resides in the suprachiasmatic nucleus (SCN) of the hypothalamus, whereas peripheral clocks are present in almost all cells of the body. The molecular makeup of the circadian oscillator is known. It is essentially identical in almost all cells of our body. This oscillator consists of interwoven sets of transcriptional-translational feedback loops and orchestrates a transcriptional program that drives the rhythmic transcription of ~10% of all expressed genes in a given tissue according to time of day.

The circadian clock is a pervasive, cell-based molecular program that is essential to health and well-being. Mis-aligned or disrupted circadian clocks are common in modern society and have been associated with numerous, highly prevalent, population diseases such as metabolic syndromes, cancer, psychiatric disturbance and cardiovascular pathologies. In addition, time-of-day adapted therapeutic intervention (chronotherapy) has been proven to be superior compared to standard therapy, e.g. in chemotherapy of cancer or rheumatoid arthritis. A new field, chronomedicine, is emerging, as evidenced by the appearance of specialized clinics.

However, an easy and convenient diagnostic tool to characterize the chronobiological state of humans, i.e. internal time and strength of rhythms, is still missing. Such a tool is urgently needed to boost chronomedicine as a field, e.g. (i) to personalize chronotherapy according to individual internal time (chronotype), (ii) to distribute shift-workers to day versus night shift according to their chronotype to minimize clock disruption; (iii) to develop treatment strategies (i.e. light or melatonin therapy) for patients with reduced internal rhythms (e.g. patients in ICUs, patients with neurodegenerative diseases, elderly people); (iv) to learn more about the interaction of the circadian system with population diseases; (v) as well as to stratify cohorts in clinical studies according to internal time (since pharmacokinetics and pharmacodynamics of many drugs have been shown to be time-of-day dependent).

Chronotyping an individual (i.e. assessing her/his individual internal time) is currently done either (i) by questionnaires, such as the Munich Chronotype Questionnaire or the Horne & Östberg morningness-eveningness questionnaire or (ii) by determining physiological or behavioral parameters using many repeated measurements (time series). Ad (i):

Questionnaires are intrinsically not objective since they depend on an individual's own declarations. Furthermore, they do not assess an acute chronobiological state but rather ask for an overall preference or description of sleep habits. Ad (ii) More objective alternatives measure rhythmic processes and determine a phase from this. Examples are actigraphy, in which an individual's activity profile is determined over several days with wearable devices. In addition, core body temperature rhythms are determined using appropriate temperature probes in a controlled clinical setting. The currently most frequently used method to objectively assess an individual's internal phase is to determine the so-called dim-light melatonin onset (DLMO) (Pandi-Perumal SR et al, Prog Neuropsychopharmacol Biol Psychiatry, 31, 1-11). For this, individuals need to stay in dim light for 6 hrs and deliver a saliva sample every 30 min, which is then analyzed for melatonin level – a tedious and complicated procedure, which cannot be used in a routine doctor's office. Together, in contrast to the new solution according to the present invention, none of the current standards is simple and feasible in a non-specialized setting.

Based on the above mentioned state of the art, the objective of the present invention is to provide means and methods for accurate, easy, fast determination of a person's internal time or the robustness/strength of his or her individual circadian rhythm. This objective is attained by the claims of the present specification.

Detailed description of the invention

The invention provides a method of assessing a time-related physiological parameter. The time-related physiological parameter is characteristic for a person's individual circadian rhythm. In certain embodiments, the time-related physiological parameter is a person's internal time. In certain embodiments, the time-related physiological parameter is the robustness or strength of a person's circadian rhythm.

The method comprises the steps of

a. determining, in a measurement step, an expression level of a plurality of circadian oscillatory genes in a sample of cells obtained from said person; and

b. assessing, in a calculation step, said time-related physiological parameter based on the expression level determined of said plurality of circadian oscillatory genes in said measurement step.

An expression level is determined for each gene comprised in the plurality of genes.

In certain embodiments, said cells are blood monocytes.

In certain embodiments, said cells are oral mucosa cells. In certain embodiments, said cells are skin fibroblasts.

In certain particular embodiments, said cells are blood monocytes.

In a preferred embodiment, said expression level of said plurality of circadian oscillatory genes is determined in blood monocytes.

In an alternative embodiment, said expression level of said plurality of circadian oscillatory genes is determined in skin fibroblasts. In an alternative embodiment, said expression level of said plurality of circadian oscillatory genes is determined in cells of the oral mucosa.

In certain embodiments, the plurality of circadian oscillatory genes comprises at least 2 of the genes of table 1. In certain embodiments, the plurality of circadian oscillatory genes comprises at least 4 of the genes of table 1. In certain embodiments, the plurality of circadian oscillatory genes comprises at least 8 of the genes of table 1. In certain embodiments, the plurality of circadian oscillatory genes comprises at least 10 of the genes of table 1. In certain embodiments, the plurality of circadian oscillatory genes comprises at least 16 of the genes of table 1. In certain embodiments, the plurality of circadian oscillatory genes comprises at least 20 of the genes of table 1. In certain embodiments, the plurality of circadian oscillatory genes comprises at least 32 of the genes of table 1.

In certain embodiments, the plurality of circadian oscillatory genes comprises at least 2 of the genes of table 1A. In certain embodiments, the plurality of circadian oscillatory genes comprises at least 4 of the genes of table 1A. In certain embodiments, the plurality of circadian oscillatory genes comprises at least 8 of the genes of table 1A. In certain embodiments, the plurality of circadian oscillatory genes comprises at least 10 of the genes of table 1A. In certain embodiments, the plurality of circadian oscillatory genes comprises at least 16 of the genes of table 1A. In certain embodiments, the plurality of circadian oscillatory genes comprises at least 20 of the genes of table 1A. In certain embodiments, the plurality of circadian oscillatory genes comprises at least 32 of the genes of table 1A.

In certain embodiments, the plurality of circadian oscillatory genes consists of at least 2 genes selected from the genes of table 1. In certain embodiments, the plurality of circadian oscillatory genes consists of at least 4 genes selected from the genes of table 1. In certain embodiments, the plurality of circadian oscillatory genes consists of at least 8 genes selected from the genes of table 1. In certain embodiments, the plurality of circadian oscillatory genes consists of at least 10 genes selected from the genes of table 1. In certain embodiments, the plurality of circadian oscillatory genes consists of at least 16 genes selected from the genes of table 1. In certain embodiments, the plurality of circadian oscillatory genes consists of at least 20 genes selected from the genes of table 1.

In certain embodiments, the plurality of circadian oscillatory genes consists of at least 2 genes selected from the genes of table 1A. In certain embodiments, the plurality of circadian

oscillatory genes consists of at least 4 genes selected from the genes of table 1A. In certain embodiments, the plurality of circadian oscillatory genes consists of at least 8 genes selected from the genes of table 1A. In certain embodiments, the plurality of circadian oscillatory genes consists of at least 10 genes selected from the genes of table 1A. In certain embodiments, the plurality of circadian oscillatory genes consists of at least 16 genes selected from the genes of table 1A. In certain embodiments, the plurality of circadian oscillatory genes consists of at least 20 genes selected from the genes of table 1A.

In certain embodiments, the plurality of circadian oscillatory genes is selected from the genes of table 1.

In certain embodiments, the plurality of circadian oscillatory genes is selected from the genes of table 1A.

In certain embodiments, the plurality of circadian oscillatory genes comprises one of PER2 or NR1D2. In certain embodiments, the plurality of circadian oscillatory genes comprises PER2 and NR1D2. In certain embodiments, the plurality of circadian oscillatory genes consists of PER2 and NR1D2.

In certain embodiments, the expression levels are determined in a single sample.

The sample is typically obtained by taking about 5-10 ml of peripheral blood in a standard setting and isolating blood monocytes from the blood sample. Monocytes are a subtype of white blood cells (leukocytes) present in the blood. They can be isolated from other blood cells. One convenient way of isolation is via magnetic activated cell sorting (MACS). This method uses superparamagnetic nanoparticles coated with antibodies against a particular cell surface antigen to tag the targeted cells. The particle-cell complex can be immobilized in a column that is placed between magnets. When the column is removed from the magnetic field, the complexes can be eluted.

In certain embodiments, the expression level of at least 2, at least 4, at least 8, at least 10, at least 16, at least 20, or at least 32 genes selected from the genes of table 2 is determined.

In certain embodiments, a single sample is obtained and the expression level of at least 2 genes of the genes of table 2 is determined. In certain embodiments, a single sample is obtained and the expression level of at least 4 genes of the genes of table 2 is determined. In certain embodiments, a single sample is obtained and the expression level of at least 8 genes of the genes of table 2 is determined. In certain embodiments, a single sample is obtained and the expression level of at least 10 genes of the genes of table 2 is determined. In certain embodiments, a single sample is obtained and the expression level of at least 16 genes of the genes of table 2 is determined. In certain embodiments, a single sample is obtained and the expression level of at least 20 genes of the genes of table 2 is determined. In certain

embodiments, a single sample is obtained and the expression level of at least 32 genes of the genes of table 2 is determined.

In certain embodiments, the expression levels are determined in each of two samples from the same patient. In certain embodiments, the two samples are obtained at a time-of-day (or night) that is 2 - 10 hours apart. In certain embodiments, the two samples are obtained at a time-of-day that is 4 - 8 hours apart, more particularly approximately 6 hours apart.

The term “time of day” herein refers to the hour given by the clock and might be, in certain embodiments, refer to measurements taken on different days.

Examples for two samples that are obtained at a time-of-day that is 6 hours apart are:

- sample 1 at 9 am (day 1), sample 2 at 3 pm (day 1);

- sample 1 at 3 pm (day 1), sample 2 at 9 am (day 2).

If sample 1 is obtained at 9 am on day 1, and sample 2 is obtained at 10 am on day 2, these samples are obtained at a time-of-day that is only 1 hour apart, although 25 hours have passed.

In certain embodiments, the expression level of at least 2 genes of the genes of table 3 is determined in each of two samples. In certain embodiments, the expression level of at least 4 genes of the genes of table 3 is determined in each of two samples. In certain embodiments, the expression level of at least 6 genes of the genes of table 3 is determined in each of two samples. In certain embodiments, the expression level of at least 8 genes of the genes of table 3 is determined in each of two samples. In certain embodiments, the expression level of at least 10 genes of the genes of table 3 is determined in each of two samples. In certain embodiments, the expression level of at least 16 genes of the genes of table 3 is determined in each of two samples. In certain embodiments, the expression level of at least 20 genes of the genes of table 3 is determined in each of two samples. In certain embodiments, the expression level of at least 32 genes of the genes of table 3 is determined in each of two samples.

In certain embodiments, the expression level of a plurality of circadian oscillatory genes selected from table 2 is analysed in one sample. In instances where the assessment of the time-dependent parameter is not satisfactory, the expression level of a plurality of circadian oscillatory genes selected from table 3 is analysed in two samples obtained from the same person. The two samples are obtained at a time-of-day that is at least 3, at least 6 or at least 12, particularly at least 6 hours apart.

In certain embodiments, the calculation step (1b) is effected by using the expression levels determined in the measurement step (1a) as input values, and applying an algorithm to the input values, thereby generating an output value. In certain embodiments, the algorithm is the

ZeitZeiger algorithm. In certain embodiments, the algorithm is the Molecular-timetable algorithm. In certain embodiments, the algorithm is partial least squares regression. In certain embodiments, the algorithm is LASSO. In certain embodiments, the algorithm is a modification of the ZeitZeiger algorithm or the Molecular-timetable algorithm or partial least squares regression or LASSO. In certain embodiments, a combination of the ZeitZeiger algorithm and/or the Molecular-timetable algorithm and/or partial least squares regression and/or LASSO is used. In certain embodiments, one or both of said algorithms are modified. A description of the ZeitZeiger algorithm and the Molecular-timetable algorithm is given below.

In certain embodiments, the output value corresponds to internal time expressed as hours past dim-light melatonin onset (DLMO).

In certain embodiments, the output value corresponds to strength of circadian rhythm. In certain particular embodiments, the output value corresponds to the amplitude of oscillation of the expression level of a circadian oscillatory gene. In certain particular embodiments, the output value corresponds to the amplitude of oscillation of the expression level of a set of circadian oscillatory genes. In certain particular embodiments, the output value corresponds to the amplitude of oscillation of one or several hormone levels (e.g. cortisol) determined in a body fluid. In certain particular embodiments, the output value corresponds to the amplitude of oscillation of the behavioural activity level as determined by an actigraphy.

Determining the rhythm strength is important for identifying individuals with disturbed or weak internal rhythms (e.g. patients in ICUs, patients with neurodegenerative diseases, elderly people). Such patients benefit from a treatment supportive for the circadian rhythm (i.e. light or melatonin therapy).

Determining the rhythm strength is also useful for diagnostic purposes, because disturbed or weak internal rhythms are associated with several pathologies.

In certain embodiments, the calculation step comprises

a. determining a first value SPC1 based on the expression level of a first subset of the plurality of circadian oscillatory genes;

b. determining a second value SPC2 based on the expression level of a second subset of the plurality of circadian oscillatory genes.

Each possible internal time has a unique combination of SPC1 and SPC2. The ZeitZeiger algorithm calculates for each sample a value corresponding to SPC1 and a value corresponding to SPC2 and identifies the internal time having the most similar combination. Thereby, the internal time of the sample is determined.

Fig. 2 (top row) illustrates the oscillation of SPC1 and SPC2 over a 24 hour period. When the values of SPC1 and SCP2 for a given sample are known, the internal time can be deduced from this information.

In certain embodiments, the first subset comprises PER2 and/or the second subset comprises NR1D2. In certain embodiments, the first subset comprises PER2 and the second subset comprises NR1D2.

In certain embodiments, a single sample is provided, the first subset comprises PER2 and/or the second subset comprises NR1D2. In certain embodiments, a single sample is provided, the first subset comprises PER2 and the second subset comprises NR1D2.

In certain embodiments, the first value is determined by multiplying a measured value, which is the relative expression level obtained in the measurement step and a constant value for each gene comprised in the first subset, thereby obtaining a product. Subsequently, the products obtained for each gene comprised in the first subset are added, thereby obtaining said first value. The second value is determined by multiplying a measured value, which is the relative expression level obtained in the measurement step and a constant value for each gene comprised in the second subset, thereby obtaining a product. Subsequently, the products obtained for each gene comprised in the second subset are added, thereby obtaining the second value. Said constant value corresponds to a loading coefficient specific for each gene.

Depending on the method that is used to determine the relative expression of the plurality of circadian oscillatory genes, the loading coefficient is different. In other words, when a platform is used for the first time for the determination of the relative expression of the plurality of circadian oscillatory genes, the loading coefficients specific for each gene on said platform have to be determined. In subsequent applications using the same platform, those loading coefficients specific for the platform can be used.

The way that the relative expression level of a gene is determined depends on the method used for quantification of the expression. If qPCR or NanoString technology is used to determine the expression level, the relative expression level is usually determined as expression of a gene relative to expression of a housekeeping gene or several housekeeping genes. A housekeeping gene is typically a constitutive gene that is required for the maintenance of basic cellular function, and is expressed in all cells of an organism under normal and patho-physiological conditions. Non-limiting examples of housekeeping genes are GAPDH, HPRT1, PSMB2 and PPIA.

The skilled person is aware of several methods suitable for the quantification of gene expression levels.

In certain embodiments, the expression levels of the plurality of circadian oscillatory genes are determined using quantitative PCR (qPCR). In certain embodiments, the expression levels of the plurality of circadian oscillatory genes are determined using a microarray.

In certain embodiments, the expression level is determined using the NanoString method. The NanoString technology is a sensitive method to determine gene expression levels. It uses probes that hybridize to mRNA molecules in solution. Each target-specific probe corresponding to a transcript of interest carries a so-called "molecular barcode" that allows its unambiguous identification and quantification by microscopic single-molecule imaging.

The principle of the invention is based on the fact that about 10% of all genes are rhythmically transcribed in nearly all human cells with phase of peak expression being gene-dependent. The relative levels of oscillating transcripts are therefore unique at any given time of day.

The inventors decided on blood monocytes as easily accessible cell type of humans, because (i) they are easy to purify using magnetic sorting of blood; (ii) they show robust circadian rhythms (in contrast to B and T cells); (iii) ~10% of their genes are rhythmically transcribed; (iv) they represent a systemic (not local) circadian oscillator of humans.

Likewise, cells of the oral mucosa or skin fibroblasts are particularly suitable cells for the method according to the invention.

Exemplary diagnostic pipeline for diagnosing an individual's internal time and rhythm strength:

(i) take about 5-10 ml blood in a standard setting,

(ii) magnetically sort the blood monocytes;

(iii) measure the expression level of up to 52 time-telling genes;

(iv) feed bioinformatics algorithms (modification and combination of published algorithms, e.g. "Molecular-timetable" and/or "ZeitZeiger" and/or partial least squares regression and/or LASSO).

In the past, scientists have proposed a single time point measurement to assess circadian phase and amplitude. These methods have been shown to be successful in animal models (in particular mouse) for predicting time from a single tissue sample. It is not straightforward to transfer these approaches to humans, due to several reasons: (i) laboratory animals are mostly isogenic, i.e. there is no genetic variability in sharp contrast to humans; (ii) in animal models (in particular Drosophila and mouse), many tissues have been analyzed regarding their circadian transcriptome. In contrast, time of day dependent transcriptome data for human cells are very rare. There exist a few transcriptome profiles from peripheral blood mononuclear cells (PBMCs), which is however a heterologous mixture of cells consisting of lymphocytes (T cells, B cells, NK cells) and monocytes; (iii) In most animal transcriptome studies, the internal time

is known, because mice are often kept in constant darkness during sampling (unmasking their internal clock). In human studies, this is often not the case – only external time is known, if no hormone profiles (e.g. melatonin) are assessed simultaneously.

Here, the inventors describe a method to measure internal time in humans from a single (or alternatively two) human blood samples. The major advances are the following: (i) Blood monocytes are selected as source for gene expression, because monocytes comprise a homogenous blood cell population and have been shown to possess a high-amplitude circadian clock in contrast to other PBMCs, such as B or T cells. (ii) The transcriptome of monocytes from 12 subjects kept in a constant routine protocol (see below) was analyzed. The data was used to extract a set of 48 plus 4 genes that - as a combination - are well suited to predict internal time. Properties of those genes are that they are rhythmic with similar phase, amplitude and magnitude across many subjects. (iii) Internal (rather than external) time was predicted with high accuracy, i.e. time relative to an internal phase marker that was simultaneously measured during blood collection (i.e. the dim-light melatonin onset, DLMO).

In certain embodiments, the sample is obtained by purification of blood monocytes from a provided peripheral blood sample by magnetic activated cell sorting (MACS).

Terms and definitions

In the context of the present specification, LASSO is an abbreviation for least absolute shrinkage and selection operator.

In the context of the present specification, the term "external time" relates to the time of day as indicated by a conventional 24 hour clock.

In the context of the present specification, a person's internal time or internal body time relates to "hours past dim-light melatonin onset (DLMO)" in said person. DLMO in the context of the present specification relates to the onset of a person's melatonin secretion under dim light conditions. "Dim light" is defined as a light of low intensity, in particular a light intensity below 50 lux. The onset of melatonin secretion is defined as the time point at which the salivary melatonin or the blood melatonin reaches a predetermined threshold (e.g. 4 pg/ml for salivary melatonin or 10 pg/ml for blood melatonin). DLMO is expressed relative to the external time (e.g. "DLMO at 9:30 pm"). In instances where a person's individual DLMO is at 10 pm external time, an internal time of "two hours past DLMO" would correspond to midnight.

In the context of the present specification, the term "circadian rhythm" relates to biological processes that display an endogenous oscillation of about 24 hours. The oscillation can be adjusted to the local environment by external cues including light and temperature.

In the context of the present specification, the term "circadian oscillatory gene" relates to genes that exhibit a rhythmic transcript abundance in a given tissue according to time of day. This

applies to approx. 10% of all expressed genes. Genes belonging to the core circadian oscillator (so-called clock genes) are rhythmically active in almost all tissues. The majority of circadian oscillatory genes (the so-called clock-controlled genes) exhibit however a rhythmic transcription that is specific for a certain cell type. When subjects are well-synchronized to the outside light-dark cycle, the phase of peripheral gene expression correlates with phase of hormonal rhythms controlled by the central clock in the suprachiasmatic nucleus (SCN).

In the context of the present specification, the term “strength of a person’s circadian rhythm” relates to the amplitude of a circadian oscillatory variable. It can be distinguished between the absolute amplitude, which is the difference between minimum and maximum value of a circadian oscillatory variable, and the relative amplitude, which is the ratio between minimum and maximum value of a circadian oscillatory variable.

Non-limiting examples of such circadian oscillatory variables are the expression level of a circadian oscillatory gene or the expression level of each of a set of circadian oscillatory genes (each determined relative to the expression level of a non-oscillatory gene, i.e. a so-called housekeeping gene), hormone levels (e.g. cortisol) or the behavioural activity level as determined by an actigraphy. In an actigraphy, an individual’s activity profile is determined over several days with wearable devices. A large amplitude of a circadian oscillatory variable is indicative for a strong individual circadian rhythm and a small amplitude is indicative for a weak circadian rhythm.

In the context of the present specification, the gene PER2 relates to human "period circadian clock 2" (Gene ID 8864, NG_012146 RefSeqGene, NC_000002.12 Reference GRCh38.p7 Primary Assembly, NC_018913.2 Alternate CHM1_1.1).

In the context of the present specification, the gene NR1D2 relates to human "nuclear receptor subfamily 1 group D member 2" (Gene ID 9975, NC_000003.12 Reference GRCh38.p7 Primary Assembly, NC_018914.2 Alternate CHM1_1.1).

The following description of the ZeitZeiger method is taken from Hughey et al., Nucleic Acids Res. 2016 May 5; 44(8):e80:

"ZeitZeiger” is a method to predict the value of a periodic variable, which we define as being continuous and bounded, where the maximum value is equivalent to the minimum value. For simplicity, we denote the periodic variable here as ‘time,’ but ZeitZeiger can be applied to any type of periodic measurement.

Similar to other supervised learning methods, training data should be a matrix $X \in \mathbb{R}^{n \times p}$ of measurements for n observations by p features and a vector $T \in \mathbb{R}^{n}$ of the corresponding time for each observation. ZeitZeiger assumes the density of each feature conditioned on time is Gaussian, so it is advisable to normalize the measurements accordingly. Time should be

scaled between 0 and 1. Training data can have missing measurements. Test data cannot have missing measurements for the features used in the predictor (typically a small subset). Time-points in the training data do not have to be evenly spaced and each time-point could have a different number of replicates.

The first step of training is to estimate the time-dependent density of each feature *j* (step 1). Due to the nature of periodic variables, if a feature goes up, it must eventually come back down. To capture this non-monotonic behavior in an unbiased way, ZeitZeiger estimates the time-dependent mean, denoted $f_j(t)$, by fitting a periodic smoothing spline to the training observations (using the bigsplines R package (Helwig et al., 2014, J. Comput. Graph. Stat., 24, 715–732)). Parameters of the spline, such as number of knots, can be adjusted as needed.

ZeitZeiger then estimates the variance of each feature, denoted s^2_j. Importantly, this is not simply the variance of the feature in the training observations, but the variance in the time-dependent density. By default, ZeitZeiger estimates the variance as the mean of the sum of squared residuals from the spline fit, i.e. $s^2_j = \frac{RSSj}{n}$, so s_j is the estimated standard deviation about the mean curve. This assumes the variance of each feature about the mean is constant across time, which is simpler and more robust than trying to estimate a time-dependent variance (and seems to yield slightly more accurate predictions).

Next, ZeitZeiger identifies the major patterns that describe how the features change over time (steps 2 and 3). To do this, ZeitZeiger first constructs a matrix $Z \in \mathbb{R}^{m \times p}$ of time-points by features, in which the time-dependent mean of each feature is discretized into a number of time-points and scaled by that feature's standard deviation about the mean curve (step 2). The time-points are evenly spaced from 0 to 1, and the number of time-points *m* is adjustable. The value of m will be the maximum number of sparse principal components (SPCs) that can be used for prediction. If τ_i is the corresponding time-point for the *i*th row in *Z*, then

$$z_{ij} = \frac{f_j(\tau_i) - \bar{f}_j}{s_j},$$

where f_j is the mean of feature *j* over the selected timepoints, calculated as:

$$\bar{f}_j = \frac{1}{m}\sum_{i=1}^{m} f_j(\tau_i).$$

Dividing by s_j ensures that each feature is expressed in terms of signal to noise.

ZeitZeiger then subjects *Z* to a penalized matrix decomposition (Witten et al., 2009, Biostatistics, 10, 515–534) (PMD; step 3). By performing the PMD on *Z* and not on *X*, we are explicitly capturing the variation in the features associated with time (making ZeitZeiger

conceptually similar to supervised principal components (Bair et al., 2006, J. Am. Stat. Assoc., 101, 119–137)). The right singular vectors from the PMD are the SPCs, which are linear combinations of a tunably small number of features. The SPCs are the source of ZeitZeiger's L_1 regularization, the strength of which is controlled by the parameter *sumabsv*. By default, ZeitZeiger performs the PMD such that the left singular vectors are orthogonal to each other, which discourages the SPCs from being highly correlated with each other. We denote the matrix of m SPCs, each of length p, as $V \in \mathbb{R}^{p \times m}$. ZeitZeiger then uses the SPCs to project the training data from high-dimensional featurespace to low-dimensional SPC-space (step 4), producing a new matrix $\tilde{X} \in \mathbb{R}^{n \times m}$ calculated as $\tilde{X} = XV$.

In the last step of training, ZeitZeiger uses $\tilde{X}$ to estimate the time-dependent density of each SPC in exactly the same way as was done for each individual feature (step 5). Although the time-dependent means of the SPCs could be extracted from the left singular vectors of the PMD, calculating the variances requires $\tilde{X}$. We denote the time-dependent mean of the kth SPC as $\tilde{f}_k(t)$ and the variance as $\tilde{s}^2_k$.

Once the predictor is trained, making a prediction for a test observation $w \in \mathbb{R}^p$ requires only two steps. First, ZeitZeiger projects the test observation from feature-space to SPC-space: $\tilde{w} = wV$ (step 6). Second, given the SPC values of the test observation and the estimated timedependent densities of those SPCs from the training data, ZeitZeiger uses maximum-likelihood to predict the time of the test observation (step 7). Because we assume each SPC is Gaussian at any given time, the likelihood of time t given $\tilde{w}_k$ is,

$$\ell_k(t \mid \tilde{w}_k) = \frac{1}{\tilde{s}_k\sqrt{2\pi}} e^{-(\tilde{w}_k - \tilde{f}_k(t))^2 / 2\tilde{s}_k^2}$$

.

The final parameter of ZeitZeiger is nSPC, the number of SPCs used to calculate the likelihood, where $nSPC \leq m$. Only features that contribute to at least one of the first $nSPC$ SPCs will contribute to the prediction. If we treat the SPCs as if they were independent (which is not valid, but empirically works well), then the likelihood as a function of time is,

$$\ell(t \mid \tilde{w}) = \prod_{k=1}^{nSPC} \ell_k(t \mid \tilde{w}_k)$$

and the log-likelihood is,

$$L(t \mid \tilde{w}) = \sum_{k=1}^{nSPC} L_k(t \mid \tilde{w}_k)$$

.

The predicted time $\hat{t}$ for test observation w is,

$$\hat{t} = \underset{t \in [0,1)}{\arg\max}\, L(t \mid \tilde{w})$$

To solve for $\hat{t}$, which is a bound-contrained optimization problem, ZeitZeiger uses the bbmle R package. For each test observation, ZeitZeiger provides the predicted time and the corresponding log-likelihood."

ZeitZeiger is available as an R package at https://github.com/jakejh/zeitzeiger. All code, data and results for the ZeitZeiger study (Hughey et al., 2016) are available at http://dx.doi.org/10.5061/dryad.hn8gp.

Further developments of the ZeitZeiger method are described in Hughey et al., Genome Med. 2017 Feb 28;9(1):19.

The following description of the Molecular-timetable method is taken from Ueda et al., Proc Natl Acad Sci U S A. 2004 Aug 3; 101(31):11227-32:

"We first normalized an expression level, *Xi*, of the time-indicating gene *i* (*i* = 1 . . . *N*) using its average μ_i and SD σ_i in the molecular timetable. The normalized expression level Y_i is described as follows: $Yi = (X_i - \mu_i) / \sigma_i$. We created an expression profile $\{t_i, Y_i\}$ (*i* = 1 . . . *N*) composed of the molecular peak time t_i of gene *i* and its normalized expression level Y_i. To estimate the BT of an expression profile, we calculated the correlation over genes (*i* = 1 . . . *N*) between $(\sqrt{1/24 \int_0^{24} Cos^2(2\pi(t-b)/24)\, dt} = 1/\sqrt{2})$ an expression profile $\{t_i, Y_i\}$ and a 24-h cosine curve $\{t_i, \sqrt{2}\ Cos(2\pi(t_i - b)_24)\}$ with a certain phase *b* ($0 \leq b < 24$). The amplitude of the cosine curve is set to _2, so that the SD of the normalized expression level *Yi* matches the SD of a continuous cosine waveform.

We prepared 24-h cosine curves with a phase *b* from 0 to 24 h in increments of 10 min. We then selected the best-fitted cosine curve that gave the best correlation value *c*. We noted that the best correlation value *c* is always positive, because we calculated the maximum value of correlation between an expression profile and 144 test cosine curves. We also noted that the phase of the best-fitted cosine curve (b_c) indicates an estimated BT."

The Molecular-timetable method is also described in patent application US2006078883, which is incorporated herein by reference.

A comprehensive review of Circadian Clocks is available in Kramer and Merrow, Handbook of Experiment al Pharmacology, Springer, Heidelberg,2013, ISBN 978-3-642-25950-0.

A comprehensive review of chronomedicine is available in Roenneberg and Merrow, 2016, Curr Biol. 26, R432-43.

Wherever alternatives for single separable features are laid out herein as “embodiments”, it is to be understood that such alternatives may be combined freely to form discrete embodiments of the invention disclosed herein.

Item 1: A method of assessing a time-related physiological parameter selected from

- a person's circadian rhythm,
- a person's internal time and
- the strength of a person's circadian rhythm,

said method comprising the steps of

a. determining, in a measurement step, an expression level of each of a plurality of circadian oscillatory genes in a sample of cells obtained from said person;
b. assessing, in a calculation step, said time-related physiological parameter based on the expression level determined of each of said plurality of circadian oscillatory genes in said measurement step.

Item 1A: A method of assessing a time-related physiological parameter according to item 1, characterized in that said cells are blood monocytes.

Item 1B: A method of assessing a time-related physiological parameter according to item 1, characterized in that said cells are skin fibroblasts.

Item 1C: A method of assessing a time-related physiological parameter according to item 1, characterized in that said cells are oral mucosal cells.

Item 2: The method according to item 1, 1A, 1B or 1C, wherein said plurality of circadian oscillatory genes comprises at least 2, at least 4, at least 8, at least 10, at least 16, at least 20, or at least 32 of the genes of table 1.

Item 3: The method according to any one of the above items, wherein said plurality of circadian oscillatory genes consists of at least 2, at least 4, at least 8, at least 10, at least 16, at least 20, or at least 32 genes selected from the genes of table 1.

Item 4: The method according to any one of the above items, wherein said plurality of circadian oscillatory genes comprises one of PER2 or NR1D2, or both.

Item 5: The method according to any one of the above items, wherein said plurality of circadian oscillatory genes consists of PER2 and NR1D2.

Item 6: The method according to any one of the above items, wherein said expression levels are determined in a single sample.

Item 7: The method according to item 6, wherein an expression level of at least 2, at least 4, at least 8, at least 10, at least 16, at least 20, or at least 32 of the genes of table 2 is determined.

Item 8: The method according to any one of the above items, wherein said expression level is determined in each of two samples, wherein said two samples are obtained at a time of day 2 - 10 hours apart, particularly at least 4 - 8 hours apart, more particularly approximately 6 hours apart.

Item 9: The method according to item 8, wherein an expression level of at least 2, at least 4, at least 8, at least 10, at least 16, at least 20, or at least 32 of the genes of table 3 is determined.

Item 10: The method according to any one of the above items, wherein said calculation step is effected by using said expression levels determined in said measurement step as input values, and applying an algorithm to said input values, thereby generating an output value, wherein in particular said algorithm is selected from the ZeitZeiger algorithm, the Molecular-timetable algorithm, partial least squares regression, LASSO, and a combination of any of the aforementioned four methods.

Item 11: The method according to item 10, wherein said output value corresponds to

a. internal time or
b. strength of circadian rhythm.

Item 12: The method according to any one of the above items, wherein said calculation step comprises

a. determining a first value SPC1 based on the expression level of a first subset of said plurality of circadian oscillatory genes;
b. determining a second value SPC2 based on the expression level of a second subset of said plurality of circadian oscillatory genes.

Item 13: The method according to item 12, wherein said first subset comprises PER2 and said second subset comprises NR1D2.

Item 14: The method according to item 12 to 13, wherein

said first value SPC1 is determined by multiplying a measured value and a constant value for each gene comprised in said first subset, thereby obtaining a product, and subsequently adding the products obtained for each gene, thereby obtaining said first value; and

said second value SPC2 is determined by multiplying a measured value and a constant value for each gene comprised in said second subset, thereby obtaining a product, and subsequently adding the products obtained for each gene, thereby obtaining said second value,

wherein said measured value corresponds to said relative expression level obtained in said measurement step and said constant value corresponds to a loading coefficient specific for each gene comprised in said plurality of genes.

Item 15: The method according to any one of the above items, wherein said expression level is determined using a method selected from quantitative PCR (qPCR), NanoString and microarray, in particular using the NanoString method.

The invention is further illustrated by the following examples and figures, from which further embodiments and advantages can be drawn. These examples are meant to illustrate the invention but not to limit its scope.

Brief description of the figures

Fig. 1 illustrates the accuracy of internal body time prediction. Scatter plots illustrate the correlation of observed and predicted body time for the single sample (top panel) or two sample approach (bottom panel). Each dot represents one sample (one sample approach) or ratio of two samples taken 6 h apart (two sample approach) color-coded by subject. Pearson correlation coefficients (r) are indicated. Prediction was performed using leave-one-subject-out cross-validation; prediction accuracy (right panel) is rated by median absolute error of observed and predicted internal time and its inter-quartile range (IQR).

Fig. 2 illustrates the extraction of internal time-telling gene sets. To identify the best internal time-telling gene sets each one model was trained on all samples (one sample approach, left panel) or ratios of two samples taken 6 h apart (two sample approach, right panel). Supervised sparse principal component analysis (PCA) was used to reduce the 9,115 genes of the blood monocyte transcriptome to two sparse principal components (SPC1, SPC2). The course of SPC1 and SPC2 over time is illustrated in the top panel. Each sparse principal component is a linear combination of the relative expression of 13 to 18 genes and their loadings (middle panel); e.g. SPC1=0.605 * expression(FKBP5) + 0.398 * expression(PER1) + … - 0.015 * expression(ADRB2). The time-telling gene set of the one- and two-sample approach include 30 or 33 genes in total; 15 genes are included in both (bottom panel).

Fig. 3 is a flow-chart illustrating time-telling gene identification and usage for predicting internal time.

Fig. 4 shows the purity of monocyte preparations prepared by MACS and analyzed using FACS. Given are median purities.

Fig. 5 shows the performance of internal time prediction for various numbers of time-telling genes using ZeitZeiger (ZZ) or Molecular-timetable (MTT).

Fig. 6 Extraction of candidate biomarkers and migration to the NanoString platform. (A) Cumulative frequency distributions of the absolute prediction errors for four types of ZeitZeiger internal cross-validation predictors of either internal or external time by 1- or 2-sample mRNA abundance profiles. Each type of predictor was built for nine combinations of the ZeitZeiger parameters sumabsv={1, 2, 3} and nSPC={1, 2, 3}. Insets show the average number of genes in the internal cross-validation predictors as a function of sumabsv and nSPC. (B) Global gene sets of the best-performing internal cross-validation predictors shown in panel A. Each column depicts a predictor defined by the type of the predictor variable (internal or external time), the format of the data input (1-sample or 2-sample) and the ZeitZeiger parameters (sumabsv, nSPC). Each predictor includes ten leave-one-subject-out cross-validation runs, i.e. ten gene sets. The ordering (from top to bottom) and the colors indicate how often a gene was identified as time-telling and assigned to SPC1, SPC2 or both among those ten gene sets. 34 genes that showed a high frequency of identification among cross-validation runs and were consistently identified across the best-performing predictors were chosen as a candidate biomarker set for internal time and migrated to the NanoString platform (highlighted in bold font). (C) Impact of platform migration on the performance of the candidate biomarkers for internal time. Given are cumulative frequency distributions of the absolute prediction errors of ZeitZeiger internal cross-validation models built on either RNA-Seq (blue) or NanoString data (red) obtained from the same RNA preparations. Platform comparison was performed for four types of predictors of either internal or external time by 1- or 2-sample mRNA abundance profiles.

Fig. 7 Composition and properties of the final NanoString BodyTime predictors. (A) 1-sample and 2-sample predictors trained on the NanoString data of the BOTI study for sum-absv={1, 2} and nSPC=2. Genes assigned to SPC1 or SPC2 as well as their loadings are shown. (B) NanoString expression profiles of the BOTI study's samples (n=154) in the SPC space of the 1-sample 12-gene predictor. Colors indicate bins of the internal time. (C) Time course of expression of the genes building the 1-sample 12-gene predictor. Colors indicate the individual subjects of the BOTI study. Each time course starts with the internal time of the first sample of a subject (M1, day 1) and ends with its last (M14, day 2).

Fig. 8 External validation and performance of the NanoString BodyTime predictors. (A) Cumulative frequency distributions of the absolute prediction errors of the 1-sample and 2-sample NanoString BodyTime predictors when they were applied to the VALI study data set. In case of the 1-sample assay, the internal time

stamps of all morning (M1) or afternoon (M2) samples were predicted; in case of the 2-sample assay the time stamp of the sample ratio was predicted (M1/M2). Proportion refers to the number of predictions with an absolute error that is less or equal to the specified value divided by the total number of predictions (1-sample, M1: n=28, 1-sample, M2: n=28, 2-sample, M1/M2: n=28). (B) Correlation of DLMO estimated by the BodyTime predictors and DLMO determined by the saliva melatonin RIA kit (gold standard); circular Pearson correlation coefficients (r) and p-values are indicated. (C) Bland-Altman analysis of the bias between saliva melatonin RIA and BodyTime estimations. The dashed horizontal line indicates the mean of the differences (bias), dotted lines represent the upper and lower limits (mean of the differences ± 2 standard deviations) with their 95% confidence intervals being shaded light gray. The morning sample of one subject was excluded from A-C because its 12-gene predictor maximum likelihood curve was ambiguous.

Fig. 9 Rhythmicity analysis of monocyte transcripts during constant routine. (**A**) The number of circadian transcripts identified at different false discovery rate thresholds for each subject. The circle marks the number of circadian transcripts at the chosen threshold of 0.05. (**B**) A different representation of the circadian transcripts identified at the 0.05 FDR threshold showing the ones rhythmic across different subjects. (**C**) The list of circadian transcripts shared between 6, 7, 8, 9 and 10 subjects (at the 0.05 FDR threshold). (**D, E**) Distributions of relative amplitudes and phases of the circadian transcripts. The boxplots (D) show the median, lower quartile and upper quartile of relative amplitude distribution of circadian genes for each subject. The phases of the circadian transcripts (E) in each subject are counted in 1 h bins over the time of day (external time).

Fig. 10 Platform comparison correlation analysis. The correlations of the prediction errors estimated based on NanoString and RNA-Seq internal cross-validation models are shown. Spearman r correlation coefficients are indicated.

Fig. 11 Platform comparison Bland-Altman analysis. Bland-Altman analysis of the bias between the prediction errors estimated based on NanoString and RNA-Seq internal cross-validation models. The dashed horizontal line indicates the mean of the differences (bias), dotted lines represent the upper and lower limits (mean of the differences ± 2 standard deviations) with their 95% confidence intervals being shaded light gray.

Fig. 12 Properties of the final BodyTime predictors. (**A-C**). NanoString expression profiles of the BOTI study's samples (n=154) in the SPC space of the 1-sample 2-gene, 2-sample 13-gene or 2-sample 2-gene predictor. (**D**) NanoString expression profiles of the BOTI study's samples (n=154) in the SPC space of the 1-sample 12-gene predictor faceted by subject. Colors indicate bins of the internal time.

Fig. 13 Comparison of DLMO estimated by the 2-gene BodyTime predictors to DLMO determined by the saliva melatonin RIA (gold standard). (**A**) Circular correlation analysis. Circular Pearson correlation coefficients (r) and p-values are indicated. (**B**) Bland-Altman analysis. The dashed horizontal line indicates the mean of the differences (bias), dotted lines represent the upper and lower limits (mean of the differences ± 2 standard deviations) with their 95% confidence intervals being shaded light gray.

Fig. 14 The prediction error of the BodyTime predictor is independent of the chronotype of the subject. (**A**) Correlation plot of the absolute prediction error and DLMO determined from saliva melatonin concentrations measured by RIA (gold standard) of subjects. The Pearson correlation coefficient and its significance are indicated in the top-left corner. (**B**) Boxplot of the absolute prediction error for early and late chronotypes (DLMO > 21:30 h). The p-values of Mann-Whitney U-tests are indicated.

Fig. 15 External validation of predictors in the independent VALI study using LASSO or partial least squares (PLS). Cumulative frequency distributions of the absolute prediction errors of the 1-sample and 2-sample NanoString predictors when they were applied to the VALI study data set using LASSO or PLS. In case of the 1-sample assay, the internal time stamps of all morning (M1) or afternoon (M2) samples were predicted; in case of the 2-sample assay the time stamp of the sample difference was predicted (M1 - M2). Proportion refers to the number of predictions with an absolute error that is less or equal to the specified value divided by the total number of predictions (1-sample, M1: n=28, 1-sample, M2: n=28, 2-sample, M1-M2: n=28).

Examples

Material and Methods

Screening study design. To identify genes that are rhythmically expressed in human blood monocytes, the inventors performed a so-called constant routine protocol (Czeisler et al., Science, 1989 Jun 16; 244(4910):1328-33) with 12 male subjects (19 - 30 years), from whom

peripheral blood samples (~10 ml) were taken every three hours. For each subject, the inventors also determined the dim-light melatonin onset (DLMO) from hourly saliva samples as a reference for internal time. The constant routine protocol is designed to minimize unwanted effects of sleep, activity or meals on circadian gene expression. Therefore, subjects were kept in a semi-recumbent posture in bed under dim light, constant temperature and humidity during a period of 42 hours of sleep deprivation. They received isocaloric snacks every hour. Hourly saliva samples were taken for determination of melatonin rhythms.

Validation study design. Fourteen individuals of early type (“larks,” 5 males and 9 females) and 14 of late type (“owls,” 6 males and 9 females) were chosen for participation based upon responses to the Horne–Ostberg Morningness-Eveningness Questionnaire and the Munich Chronotype Questionnaire. Participants were asked to maintain a regular sleep-wake rhythm– with approximately 8 hours of sleep, within ± 30 min of self-selected target times, based on their habitually chosen bed times. Compliance was controlled by a wrist worn activity monitor (Daqtix®, Oetzen-Süttorf, Germany) and sleep logs. On the study day the first (second) blood sampling was scheduled to begin approximately two (eight) hours after the assigned habitual subjective wake-time. Approximately 10 hours after habitual subjective wake-time, saliva samples were collected every 30 min to determine DLMO.

Blood monocyte isolation. Blood monocytes were isolated as described (Spies et al., Clin Exp Rheumatol, 2015 Jan-Feb; 33(1):34-43). Briefly, heparinized whole blood was immediately placed on ice after sampling. $CD14^+$ blood monocytes were collected by MACS sort using whole blood CD14-microbeads (Miltenyi Biotec) and the Auto-MACS-Pro device. All steps were performed at 4 °C according to the manufacturer’s instructions. CD14+ cell fraction was pelleted by centrifugation and frozen at -80 °C. Purity of monocyte preparations was analyzed using FACS and resulted in a median purity of 89% of $CD15^-$ $CD11b^+$ cells.

RNA preparation and sequencing. Total RNA was isolated from each monocyte sample using the TRIzol reagent (Thermo Fisher) according to the manufacturer’s instructions. All samples were quantitated and quality controlled with the NanoDrop 2000 Spectrophotometer (Thermo Scientific) and the Qubit 2.0 Fluormeter (Thermo Scientific). Before 3´-end RNA-sequencing, all samples were DNAse-treated. For 3´-end RNA-sequencing, the inventors used the protocol described in Shishkin et al. (Shishkin et al., Nat Methods, 2015 Apr; 12(4):323-5) with the exception of $polyA^+$ selecting (using Dynabeads™ $Oligo(dT)_{25}$, Thermo Fisher, according to the manufacturer’s instructions) after pooling the samples into a single tube, instead of rRNA depletion step.

Gene expression analysis using NanoString. The isolated RNA was subjected to NanoString nCounter™ Gene Expression analysis (NanoString Technologies, Seattle, WA, USA) using nCounter® Elements™ TagSets, according to the manufacturer’s protocol. Probes

A and B used for NanoString analysis are shown in Table 4. Probes for the time-telling genes and the housekeeping genes GAPDH, HPRT1, PSMB2 and PPIA were designed and by NanoString Technologies and synthesized by Integrated DNA Technologies (Leuven, Belgium). Raw data were analysed using nSolver™ software (NanoString Technologies) using standard settings and was normalized against the housekeeping genes.

Data analysis. The sequenced data was aligned to the human genome (hg19) using STAR short-sequence aligner (Dobin et al. 2013, Bioinformatics, 2013 Jan 1; 29(1):15-21) with standard settings such that only the uniquely mapped reads were retained. These uniquely mapped reads were assigned to RefSeq genes using the ESAT tool that was designed for quantification of 3'-end RNA-Seq (Derr et al., Genome Res, 2016 Oct; 26(10):1397-1410). Only libraries with at least 2 million total mapped and assigned 3'-end reads were used for further analysis. Libraries were normalized using the total number of mapped reads and corrected for weighted trimmed mean of M-values (Robinson and Oshlack, 2010, Genome Biology 11, R25). The resulting count data for each library was analyzed for circadian genes by linear regression of sinusoids using edgeR (Robinson et al., Bioinformatics, 2010 Jan 1; 26(1):139-40). For supervised sparse PCA analysis, the count data was converted to normalized counts-per-million data using edgeR.

Extraction of internal time-telling gene sets. To identify internal time-telling gene sets we applied the ZeitZeiger method of Hughey et al. (2016). Central to the method is a supervised sparse PCA that reduces the variation associated with the periodic variable in the training set (in our case: internal time) to a low-dimensional subspace (two sparse principal components: SPC1, SPC2) of the monocyte blood transcriptome (9,115 genes). The prediction of internal time of test observations is based on maximum-likelihood estimates. Leave-one-subject out cross-validation is used to assess prediction performance. Prediction accuracy is rated by the median absolute error of observed and predicted internal time and its inter-quartile range (IQR). The statistical significance of prediction is assessed by label-permutation testing, running cross-validation on 1,000 independent random shuffles of the internal time label vector. The proportion of outcomes where the median absolute error equals or betters the median absolute error of the unshuffled data gives an estimate of the true p-value; i.e. the p-value reflects the probability to achieve a given prediction accuracy by chance. For prediction, two different approaches were used: (i) one-sample method based on a single sample expression values; (ii) two-sample method based on the ratio of gene expression from samples taken approximately 6 hours apart.

Results

To establish a validated test, the inventors followed a three-stage biomarker development strategy: (i) unbiased discovery of biomarkers (i.e., time-telling genes) using a large-content

platform (RNA-Seq). Time-telling genes are likely to be identified among those genes, whose expression is robustly time-of-day dependent across many individuals with similar phase, amplitude and expression level; (ii) migration to a targeted gene expression profiling platform (RNA-Seq -> NanoString); and (iii) independent external biomarker/assay validation. The inventors chose to use peripheral blood monocytes as source material for gene expression profiling, because (i) blood is an easily accessible source of human cells and (ii) monocytes comprise a homogenous blood cell population and have been shown to possess a high-amplitude circadian clock in contrast to other peripheral blood mononuclear cells (PBMCs), such as B or T cells. For the final multiplex gene expression profiling, the inventors chose the NanoString Technologies nCounter platform, since it offers key advantages (sensitivity, reproducibility, technical robustness, etc.) over more traditional methods such as microarrays and quantitative RT-PCR. Moreover, NanoString-based diagnostic tests with Food and Drug Administration (FDA) clearance are already on the market. To identify genes in blood monocytes that are rhythmically expressed across many individuals, the inventors sorted monocytes from peripheral blood taken every three hours from 12 young, male volunteers during 42 hours of constant routine. For each subject, 14 blood samples were taken at regular intervals of 3 hours over a period of 40 hours in a constant routine protocol to minimize unwanted effects of sleep, activity or meals on circadian gene expression. Subjects remained in a semi-recumbent posture in bed under dim light, constant temperature and humidity during a period of 40 hours of sleep deprivation. They received isocaloric snacks every hour and water. Hourly saliva samples were taken for determination of melatonin secretion profiles. Genome-wide circadian gene expression was analyzed using RNA sequencing and read mapping to the human reference genome. For each subject, the inventors also determined the dim-light melatonin onset (DLMO) from hourly saliva samples as a reference for internal time. After quality assessment, the final RNA-Seq time series included mRNA abundance profiles (9115 genes) from 10 subjects. Among the various bioinformatics methods proposed to obtain predictors that map gene signatures to time, the inventors applied ZeitZeiger to the RNA-Seq data set, because ZeitZeiger achieves good prediction performance with fewer genes compared to other methods. The inventors devised a 4-predictors approach, i.e., they tested four types of predictors that differ with respect to the predicted variable (external time or internal time) and the format of the RNA-Seq data input (1-sample or 2-samples 6 hours apart). The difference between the two formats is the mRNA abundance profile assigned to each measurement (M_i) in the time series (1-sample: single profile recorded at M_i; 2-sample: ratio of two profiles recorded 6 h apart, M_i/M_{i+2}). The idea of the 4-predictor approach is that genes with high and robust time-telling properties should be less dependent on the type of predictor and thus should be frequently extracted by ZeitZeiger. The inventors identified

138 genes that had consistent profiles (amplitude, magnitude and phase) of gene expression across the 10 subjects at FDR<0.05.

Gene sets (i.e. combination of genes) suitable for internal time prediction (i.e. relative to DLMO) from one or two samples were extracted using supervised sparse PCA analysis. Accuracy of predication was estimated by leave-one subject-out cross validation (LOOCV) (Figure 1, Figure 6). The predictors were trained with data from all subjects except one, and internal/external times of the samples from this left-out subject were subsequently predicted. This was repeated for all subjects. Moreover, cross-validation was always done for nine combinations of ZeitZeiger's two main parameters sumabsv and nSPC. Briefly, sumabsv controls how many genes form each sparse principal components (SPC) and nSPC determines how many SPC are used for prediction, i.e. the higher sumabsv or nSPC, the more genes are needed for prediction. All four types of predictors performed comparably (Figure 6A, Table 5) in terms of optimal parameter combination (nSPC=2, sumabsv={2, 3}) and mean 183 number of genes used for prediction (12 to 15 genes for sumabsv=2, 30 to 35 genes for sum-184 absv=3). As a measure of accuracy, the median absolute difference between the predicted and the observed (internal or external time) time stamps and its interquartile range (IQR) is used. Hereafter, for simplicity, median absolute difference is abbreviated to MdAE. The accuracy achieved in predicting internal time is similar for the 1-sample (MdAE=1.6 h, IQR=2.4-3.2 h) and the 2-sample method (MdAE=1.4-1.7 h, IQR=1.9 h) (Figure 6A, Table 5); differences between the two methods are not statistically significant (Benjamini-Hochberg adjusted p-value of pairwise Wilcoxon test > 0.05). With 53.7-59.6% of the predictions showing an error of ≤2 h, the presented predictors' performances are comparable to those of previously published internal cross-validation predictors. The MdAEs show no significant variation over the entire constant routine (Kruskal-Wallis test on 3 h time bins, p-value > 0.05) for any of the predictors, indicating that the cumulative sleep deprivation experienced by the subjects does not affect prediction accuracy. Likewise, there is no statistically significant difference in terms of MdAE between samples obtained during day 1 [0-24 h] and day 2 [24-40 h] of the constant routine (Mann-Whitney U test, p-value > 0.05). The accuracies of the external time of day predictors (MdAE=1.2-1.6 h, IQR=1.8-2.4 h; Figure 6A, Table 5) are similar to those of the internal time predictors. Likewise, the 1-sample (MdAE=1.4-1.6 h, IQR=1.8-2.0 h) and 2-sample methods (MdAE=1.2-1.4 h, IQR=2.3-2.4 h) perform comparably. The global gene sets of all predictors are shown in Figure 6B. The global gene set of a predictor aggregates all genes identified by ZeitZeiger for prediction (biomarkers) across the internal cross-validation runs. For reasons of comprehensibility, the composition of the global gene sets will be specified in the context of migration to the NanoString platform. For internal time prediction based on a single blood sample, the prediction accuracy was high with a median absolute prediction error of 1.57 hours

(interquartile range: 2.40 hrs). For two-sample based prediction, accuracy was even higher (median absolute error: 1.37 hrs; interquartile range: 1.89 hrs). Statistical significance was assessed by label shuffling for both approaches (1000 runs, $p<0.001$).

Subsequently, the inventors used available data from all 10 subjects to train single models in order to obtain the best sets of time-telling genes for both the one-sample and the two sample approach (Figure 2). For both models, prediction is based on two sparse principle components: SPC1, a linear combination of 18 (16 for two sample model) genes (red) and SPC2, a linear combination of 13 (18) genes (blue). Example: SPC1 = loading (gene 1) * rel. expression (gene 1) + …. + loading (gene 18) * rel. expression (gene 18), thus each gene is essential. The extracted gene sets strongly overlapped with those being identified during LOOCV.

It is crucial which and how many time-telling genes are selected for migration to the NanoString platform. Therefore, the inventors predefined the following selection criteria: (i) the final number of genes should be as small as possible, and (ii) the genes selected for migration should have robust time-telling properties. The selection of a candidate set of time-telling genes is based on the global gene sets of the best-performing predictors (Table 5, Figure 6B). The global gene set of a predictor aggregates the genes extracted by ZeitZeiger for prediction across the internal cross-validation runs. In case of leave-one-subject-out cross-validation, as performed here, the number of cross-validation runs equals the number of subjects (N=10). Since the predictors use two sparse principal components for prediction, the maximal times a gene can be identified is N x nSPC=20. Figure 6B summarizes the global gene sets of the best-performing predictors in terms of which genes were identified and how often. A total of 119 genes were extracted at least once across all predictors. Gene ontology (GO) functional enrichment analysis revealed that they are significantly associated (p-value < 10E-5) with biological processes related to the immune system (e.g. immune response, defense response, cell activation; for a full list see Table 7). Furthermore, four genes have previously been related to the circadian clock (DBP, NR1D2, PER1 and PER2).Out of the 119 genes, 34 were selected as candidates for migration to the NanoString platform because they were frequently identified during cross-validation of the individual predictors and across predictors: *ABHD5, AGFG1, C7orf50, CD99, CLEC4E, CRISPLD2, CX3CR1, 234 CYP51A1, DBP, ELMO2, FASN, FKBP4, FKBP5, FUS, HNRNPDL, HSPA1A, HSPH1, IRAK3, IRS2, 235 LGALS3, LILRA5, MLKL, NID1, NR1D2, PER1, PER2, PHC2, RBM3, RSRP1, SERPINB9, SMAP2, TSC22D3, TSPAN4* and *UBE2J1*. Due to the size of the candidate set, the inventors opted for a 48-plex configuration for the NanoString assay. In addition to the 34 candidate genes the inventors chose to further include 4 housekeeping genes (GAPDH, HPRT1, PPIA, PSMB2), 3 clock (or clock associated) genes that were not part of the RNA-Seq data set because of low expression (KLF9, NR1D1, PER3), 2 clock genes that were not identified in the biomarker extraction process (CRY1, CRY2) as well as

5 genes that showed high frequencies of detection when the RNA-Seq data set was still incomplete (CPED1, DHRS9, HGSNAT, ODC1, PLAC-8).

To assess the impact of migration of our candidate set to the NanoString platform on prediction accuracy, the inventors performed platform comparison. To this end, gene expression profiles using the 48-plex NanoString gene set were acquired for all blood monocyte RNA-preparations obtained during the first study. Afterwards, internal cross-validation predictors for both the RNA-Seq and the NanoString data sets were built only using the genes shared by both assays (n=41) and all samples that passed quality control for both assays (n=136). The inventors again followed the 4-predictor leave-one-subject-out approach; predictor performances are shown in Figure 6C. Most strikingly, all predictors' performances significantly increase after migration to the NanoString platform (internal time 1-sample: $p < 0.0008$, internal time 2-sample: $p < 0.005$, external time 1-sample: $p < 0.0001$, external time 2-sample: $p < 0.03$; Benjamini-Hochberg adjusted p-value of pairwise Wilcoxon test). The MdAEs improved by about 1 hour ranging between 0.6 and 0.9 hours for the different predictors. Moreover, the spread of MdAEs narrowed by about a factor of 2, now reaching values in the range of 1 hour. Consequently, the frequencies of predictions with an error ≤1 h or ≤2 h increased up to >50% and >70%, respectively. Of particular note is that the NanoString predictors performed equally well for different values of sumabsv; that is, a predictor using just 2 genes for prediction achieved the same accuracy as a predictor using ≥12 genes. Spearman's correlation and Bland-Altman analyses further revealed that, in terms of prediction error, there was no relevant bias between the two platforms ($p \leq 0.33$, Figure 10; Bland-Altman mean difference [-0.2 to 0.5 h], Figure 11). Taken together, the migration of the candidate set to the NanoString platform was not only successful, but in fact significantly improved its performance.

After migration of the candidate set of time-telling genes to the NanoString platform, the inventors sought to establish and validate a final assay to assess internal circadian time. By means of internal cross-validation (leave-one-subject-out approach), they first identified the optimal parameter values for ZeitZeiger-based prediction of internal time in the NanoString data set of the first study. This analysis differed slightly from the one performed for platform comparison, in that here all genes (m=44) and all samples that passed quality control were considered (n=154 in 11 subjects) and not only those that are also present in the RNA-Seq data set. Regarding accuracy and optimal parameters (nSPC=2, sumabsv={1, 2}) the results of the final 1-sample and 2-sample assays' internal validation were essentially identical to those described above for platform comparison (for a detailed view on performance measures and a description of the global gene sets see Table 5).

Using the identified optimal parameter combinations, the inventors next trained two final internal time predictors on all samples in the 1-sample and 2-sample NanoString data set of

the first study (Figure 7A). Both the 1-sample and the 2-sample predictor trained with sumabsv=1 are formed by just two genes (NR1D2, PER2). The predictors trained with sumabsv=2 comprise 12 genes (SPC1: *NR1D2, PER3, NR1D1, LGALS3, PER2, ELMO2, FKBP4, HSPH1, CRY1; SPC2: 282 CRY1, PER2, CRISPLD2, KLF9, PER1*) and 13 different genes (SPC1: *NR1D2, PER3, NR1D1, 283 LGALS3, TSPAN4, FKBP4, ELMO2, CRY1; SPC2: CRY1, PER2, CRISPLD2, CPED1, CX3CR1, KLF9, 284 ELMO2, PER3*), respectively. More importantly, the 1-sample and 2-sample predictors show a strong overlap in terms of genes that form SPC1 and SPC2, emphasizing the high time-telling capacity of these very small sets of genes. In line with this, the expression levels of the identified genes show consistent circadian oscillations in each subject of the first study (Figure 7C). Interestingly, the graphical illustration of the course of SPC1 and SPC2 over internal time resembles a clock (12-gene 1-sample predictor, Figure 7B; for the other predictors see Figure 12B-D). When plotted in the two-dimensional SPC-space, the first study's samples describe a virtually perfect circle with the progression of internal time following an anti-clockwise trajectory. Such behavior is observed for each individual subject (Figure 12A) and might form the basis for future personalized approaches to detect perturbations of the circadian clock in humans. Taken together, the inventors established not one but four final BodyTime predictors that all show comparably high accuracy and use a very small set of genes.

To test the ability of the identified gene set to predict internal time, a second Validation Study was performed. To this end, blood samples from 29 newly recruited subjects with putatively extreme chronotypes (according to questionnaires) were taken at two different time points during the day (~6 hours apart), monocytes were sorted and expression of time-telling genes was quantified via NanoString technology. From same subjects, DLMO was simultaneously determined. Now the inventors decided on a setting better reflecting real life conditions, i.e., subjects of the Validation study were allowed to sleep, to eat meals and to be exposed to light at their habitual times.

Using the previously constructed prediction models, internal time for each sample was predicted and then used to estimate DLMO based on when the blood was actually drawn. This estimated DLMO was then compared to observed DLMO values. Example: when blood sample was drawn at 10:30am and internal time was predicted to be 12.5 [hours after DLMO], the DLMO estimate would be 10:00pm.

The 1-sample 2-gene and 12-gene BodyTime predictors applied to the Validation study morning samples achieved the best prediction accuracy of internal time (Figure 8A, Table 6). With MdAEs of 0.54 h (IQR=0.82 h) or 0.69 h (IQR=0.87 h) and 100% or 96.3% of samples showing an error of ≤2 h, both predictors performed as well as they can, considering that the error of the gold-standard reference method (DLMO determined from saliva melatonin

concentrations measured by radioimmunoassay, RIA) itself lies between 0.5 and 1 h. The agreement of our BodyTime predictors with the current gold-standard is further emphasized by the high (circular Pearson r≥0.9) and significant (p<0.0001) correlation of DLMO estimated by the predictors with DLMO determined from saliva melatonin concentrations (Figure 8B). Moreover, Bland-Altman analyses showed that there is no systematic difference between the two methods (Bland-Altman mean difference [-1.02 to 0.24h]; Figure 8C and Figure 13). Application of the 1-sample BodyTime predictors to the afternoon samples resulted in a small but statistically insignificant loss of accuracy in terms of MdAE (2-gene: 0.99 h, 12-gene: 0.80 h). Likewise, the accuracies of the two 2-sample predictors (2-gene: 0.75 h, 13-gene: 0.74 h) were somewhat better than those of the 1-sample predictors with the difference not being statistically significant. Furthermore, the accuracy of prediction does not depend on the phase of entrainment (Supplemental Figure 14).

All the prediction models perform very well with a median absolute prediction error between 0.5 and 1 hour demonstrating the predictive power of the time-telling gene set (Figure 5). With 44 genes, 29 genes, 14 genes and two genes, median prediction error were 0.94, 0.83, 0.75 and 0.70 hours.

Discussion

Following a three-stage biomarker development strategy, the inventors established and validated a novel assay (BodyTime) for assessing internal circadian time using monocyte NanoString-based gene expression profiles. The BodyTime assay achieves similar accuracy (MdAE: 0.54-0.69 h) as the current gold standard DLMO assay (error: 0.5-1 h (16, 17)) while eliminating its shortcomings. It requires only one blood sample taken at any time during the day and there is no need for a dim light environment during sampling. Moreover, it depends on just a handful of genes (≤12), is of low complexity and, thus low cost (<100$ per test). In its current form, the assay is readily useable in a research or clinical context, since the requirements regarding equipment (MACS sorting device) and staff are easily met. If not available in-house, NanoString gene expression analysis can be out-sourced to one of many external service providers.

What distinguishes the approach from all previous attempts to establish an assay for internal time is that we focused on biomarker development strategy and decision-making instead of bioinformatics tool development. While bioinformatics is certainly crucial for biomarker discovery, the approach demonstrates that, on its own, it is not sufficient. In fact, the success of the BodyTime predictors can be attributed in large part to the decision to migrate the assay to the NanoString platform. Just the migration increased the accuracy of the predictors significantly. The inventors used ZeitZeiger to identify the time-telling genes for use in the multiplex assay for internal time. ZeitZeiger has the advantage that it performs both model

construction for internal time estimation and feature extraction to find the best multiplex for use in the assay. Other established methods (molecular time table or partial least squares) use theoretically sub-optimal approaches to find the best combination of time-telling genes. The inventors used the maximum likelihood scheme to estimate internal time from a subject sample. However, other machine learning methods (LASSO and partial least squares) resulted in very similar internal time estimation performances (Figure 15, Table 8), further emphasizing that biomarker discovery and the analysis platform (NanoString) rather than a bioinformatics approach had a greater influence on the success of the approach. The choice of monocytes as a source material was driven by a desire to reduce the cellular noise in the data, at the cost of a slightly more expensive and complex assay. In contrast to whole PBMCs or other PBMC subsets, monocytes have been shown to comprise a homogenous blood cell population with a high-amplitude circadian clock. Whether this decision significantly contributes to the success in establishing the BodyTime assay is difficult to assess, because all previous attempts (while being PBMC-based) did not go beyond stage 1 of the biomarker development pipeline. That is, direct comparison cannot be made. However, some observations during initial biomarker extraction suggest that the choice of monocytes had a favorable impact on the performance of the assay: (i) independent of the predicted variable (internal or external time) and the data input format (1-sample or 2-sample), the inventors observed a strong overlap in the genes extracted by ZeitZeiger (Figure 7A), (ii) compared to other PBMC-based internal-cross validation approaches with similar performance the iventors' predictors need a much smaller set of genes (12-15 versus ~100). The inventors believe that these observations reflect a combination of reduced cellular noise of monocytes in terms of gene expression compared to PBMCs and an improved clinically proven gene expression measurement platform (NanoString) leaving the possibility that the final assay would perform just as well on PBMCs.

The accuracy of the BodyTime assay is the same for all chronotypes (Figure 14), even the extreme ones as revealed in our external validation study. It is thus not affected by where the sleep period of a person is located during the 24-hour window of a day as long as sleep occurs at the person's normal time (i.e. similar phase angle with respect to the melatonin rhythm and assuming no internal desynchronization). Whether BodyTime works similarly well in patients or cohorts with known lower circadian amplitudes (e.g. the elderly, shift workers) and/or in the presence of internal circadian desynchronization remains to be elucidated. It is important to note that neither BodyTime nor DLMO determined from saliva melatonin concentrations report the circadian phase of the SCN; rather BodyTime estimates DLMO, and the classical melatonin RIA directly measures the time of onset of melatonin secretion. Onset of melatonin secretion is a characteristic of circadian rhythms in the pineal gland,

although in clinical circadian and sleep research it serves as the gold standard proxy for SCN phase.

Among the BodyTime predictors were several core clock genes or genes associated with the circadian clock, which is in contrast to biomarker sets previously suggested based on analysis of data acquired on high-content discovery platforms such as microarrays. Since the BodyTime predictors are small in size (2 to 15 genes) the inventors cannot deduce any further biological meaning. However, although they did not make it into the final BodyTime predictors, the inventors identified many time-telling candidate genes that are associated with biological processes related to the immune system. This is in line with the previously suggested biomarker sets for internal time and hints at modulation of immune processes by the circadian clock. The assay provides a simple yet effective tool for chronotyping individuals from a single blood sample. This assay should therefore become a useful tool for example during drug development to assess whether the effectiveness and side effects of a drug are influenced by the administration schedule. Beyond that, it opens up the opportunity to investigate the dynamics of the internal circadian phase in response to environmental changes and interventions (in large study cohorts), e.g. shift work, jet lag, season (photoperiod), light level, daylight saving time as well as with age. In addition, the BodyTime predictor genes may have the diagnostic potential to assess molecular perturbations of the circadian system. When plotted in the SPC-space defined by the BodyTime predictor genes, the time series data of the first study describe a circle with the progression of internal time following an anti-clockwise trajectory (Figure 7B).

In summary, the BodyTime assay is a new precision medicine tool that allows the personalization of diagnostics and therapy according to an individual's circadian phase – an under-recognized yet important physiological parameter. BodyTime is simple (it requires only a single blood sample taken anytime during the day) and highly accurate (as good as the current gold standard) and therefore should foster the spread of chronomedicine. The main advantage of the method of the present invention over the prior art's methods is that it requires only one sample of monocytes easily obtainable from a blood sample and still achieves high accuracy. The inventors' selection of circadian oscillatory genes comprises genes oscillating with a comparable phase, amplitude and mean of expression to be able to assess a person's circadian rhythm from only one gene expression sample.

Previous studies have investigated oscillatory circadian genes in mice. Such studies, however, do not provide evidence which genes will facilitate a robust determination of internal time for a large group of patients that were not previously investigated. Such determination requires an analysis of the phase, amplitude and mean of expression, and their respective variance between individuals. By principle, such analysis cannot be performed in inbred laboratory mice,

which are isogenic. Thus, experiments with mice do not take into account that phase, amplitude and mean of expression of certain circadian oscillatory genes differ with genotype.

Table1. Total gene set

1	ADRB2
2	AGFG1
3	AHNAK
4	ALDH2
5	BTN3A2
6	C7orf50
7	CD99
8	CLEC4E
9	CRISPLD2
10	CRY1
11	CX3CR1
12	CXCR4
13	DBP
14	DPYSL2
15	ERGIC1
16	FAM129A
17	FAM26F
18	FASN
19	FCGR3A
20	FKBP5
21	FUS
22	HIST1H2BK
23	HSPA1A
24	HSPH1
25	IRAK3
26	IRS2
27	KLF9
28	LGALS3
29	LILRA5
30	MDM2
31	MLKL
32	NAGA
33	NDUFA7
34	NID1
35	NR1D1
36	NR1D2
37	PER1
38	PER2
39	PER3
40	PHC2
41	RBM3
42	RNASE6
43	RNF144B

44	RSRP1
45	SERPINB9
46	SMAP2
47	SPIDR
48	STK17B
49	TSC22D3
50	TSPAN4
51	UBE2J1
52	ZNF703

Table 1A: 25 best-performing genes

1	NR1D2
2	PER2
3	CRY1
4	PER3
5	CRISPLD2
6	NR1D1
7	KLF9
8	PER1
9	LGALS3
10	ELMO2
11	FKBP4
12	HSPH1
13	ELMO2
14	C7orf50
15	UBE2J1
16	ODC1
17	CPED1
18	FUS
19	CYP51A1
20	FASN
21	TSC22D3
22	RSRP1
23	FKBP5
24	CD99
25	TSPAN4

Table 2. One-sample method gene set

1	ADRB2
2	AGFG1
3	C7orf50
4	CLEC4E
5	CRISPLD2
6	CRY1

7	CXCR4
8	DBP
9	DPYSL2
10	ERGIC1
11	FKBP5
12	FUS
13	HSPA1A
14	HSPH1
15	IRAK3
16	IRS2
17	KLF9
18	LGALS3
19	MDM2
20	MLKL
21	NDUFA7
22	NR1D1
23	NR1D2
24	PER1
25	PER2
26	PER3
27	PHC2
28	RBM3
29	RSRP1
30	SMAP2
31	SPIDR
32	STK17B
33	TSC22D3
34	ZNF703

Table 3. Two-sample method gene set

1	AHNAK
2	ALDH2
3	BTN3A2
4	CD99
5	CLEC4E
6	CRISPLD2
7	CRY1
8	CX3CR1
9	CXCR4
10	DPYSL2
11	FAM129A
12	FAM26F
13	FASN
14	FCGR3A
15	FKBP5
16	FUS
17	HIST1H2BK
18	HSPA1A

19	IRAK3
20	IRS2
21	KLF9
22	LGALS3
23	LILRA5
24	NAGA
25	NID1
26	NR1D1
27	PER1
28	PER3
29	PHC2
30	RBM3
31	RNASE6
32	RNF144B
33	SERPINB9
34	SMAP2
35	TSC22D3
36	TSPAN4
37	UBE2J1

Table 4 NanoString probes

Gene	Sequence Probe A
ABCG1	(SEQ ID NO 001)
C7orf50	(SEQ ID NO 002)
CD99	(SEQ ID NO 003)
CLEC4E	(SEQ ID NO 004)
CRISPLD2	(SEQ ID NO 005)
CRY1	(SEQ ID NO 006)
CX3CR1	(SEQ ID NO 007)
DBP	(SEQ ID NO 008)
FAM129A	(SEQ ID NO 009)
FASN	(SEQ ID NO 010)
FCGR3A	(SEQ ID NO 011)
FKBP5	(SEQ ID NO 012)
FUS	(SEQ ID NO 013)
HSPA1A	(SEQ ID NO 014)
HSPH1	(SEQ ID NO 015)
IRAK3	(SEQ ID NO 016)
IRS2	(SEQ ID NO 017)
KLF9	(SEQ ID NO 018)
LGALS3	(SEQ ID NO 019)
LILRA5	(SEQ ID NO 020)
MLKL	(SEQ ID NO 021)
NID1	(SEQ ID NO 022)

NR1D1	(SEQ ID NO 023)
NR1D2	(SEQ ID NO 024)
PER1	(SEQ ID NO 025)
PER2	(SEQ ID NO 026)
PER3	(SEQ ID NO 027)
PHC2	(SEQ ID NO 028)
RBM3	(SEQ ID NO 029)
RSRP1	(SEQ ID NO 030)
SERPINB9	(SEQ ID NO 031)
SMAP2	(SEQ ID NO 032)
TSC22D3	(SEQ ID NO 033)
TSPAN4	(SEQ ID NO 034)
UBE2J1	(SEQ ID NO 035)

Gene	**Sequence Probe B**
ABCG1	(SEQ ID NO 036)
C7orf50	(SEQ ID NO 037)
CD99	(SEQ ID NO 038)
CLEC4E	(SEQ ID NO 039)
CRISPLD2	(SEQ ID NO 040)
CRY1	(SEQ ID NO 041)
CX3CR1	(SEQ ID NO 042)
DBP	(SEQ ID NO 043)
FAM129A	(SEQ ID NO 044)
FASN	(SEQ ID NO 045)
FCGR3A	(SEQ ID NO 046)
FKBP5	(SEQ ID NO 047)
FUS	(SEQ ID NO 048)
HSPA1A	(SEQ ID NO 049)
HSPH1	(SEQ ID NO 050)
IRAK3	(SEQ ID NO 051)
IRS2	(SEQ ID NO 052)
KLF9	(SEQ ID NO 053)
LGALS3	(SEQ ID NO 054)
LILRA5	(SEQ ID NO 055)
MLKL	(SEQ ID NO 056)
NID1	(SEQ ID NO 057)
NR1D1	(SEQ ID NO 058)
NR1D2	(SEQ ID NO 059)
PER1	(SEQ ID NO 060)
PER2	(SEQ ID NO 061)
PER3	(SEQ ID NO 062)

PHC2	(SEQ ID NO 063)
RBM3	(SEQ ID NO 064)
RSRP1	(SEQ ID NO 065)
SERPINB9	(SEQ ID NO 066)
SMAP2	(SEQ ID NO 067)
TSC22D3	(SEQ ID NO 068)
TSPAN4	(SEQ ID NO 069)
UBE2J1	(SEQ ID NO 070)

Table 5. Best-performance internal cross-validation predictors built on the BOTI RNA-Seq or NanoString data sets

Predictor (time)	Assay	Data source	Parameters (sumabsv, nSPC)	Number of genes (mean±sd)	MdAE [IQR]	AE ≤ 1 h [% of samples]	AE ≤ 2h [% of samples]
internal	1-sample	RNA-Seq	2, 2	15 ± 3	1.6 [3.2]	39.0	53.7
internal	1-sample	RNA-Seq	3, 2	32 ± 2	1.6 [2.4]	38.2	59.6
internal	2-sample	RNA-Seq	2, 2	12 ± 2	1.7 [1.9]	35.7	59.1
internal	2-sample	RNA-Seq	3, 2	30 ± 3	1.4 [1.9]	40.9	58.7
external	1-sample	RNA-Seq	2, 2	14 ± 2	1.6 [2.0]*	33.8	59.6
external	1-sample	RNA-Seq	3, 2	35 ± 3	1.4 [1.8]	34.6	64.0
external	2-sample	RNA-Seq	2, 2	12± 1	1.2 [2.3]*	43.5	66.1
external	2-sample	RNA-Seq	3, 2	31± 4	1.4 [2.4]*	40.9	64.3
internal	1-sample	NanoString	1, 2	2 ± 0	0.9 [1.2]	55.2	79.2
internal	1-sample	NanoString	2, 2	14 ± 1	0.8 [1.1]	59.7	88.3
internal	1-sample	NanoString	3, 2	30 ± 1	0.8 [1.1]	59.1	87.0
internal	2-sample	NanoString	1, 2	2 ± 0	0.8 [1.1]*	61.4	85.6
internal	2-sample	NanoString	2, 2	13 ± 2	0.7 [1.0]	64.4	90.9
internal	2-sample	NanoString	3, 2	29 ± 1	0.7 [1.0]*	61.4	90.9

Table 6: External validation of the BodyTime predictors in the independent VALI study.

Predictor	Type of validation sample	Absolute prediction error median [IQR]	Absolute prediction error ≤ 1 h [% of samples]	Absolute prediction error ≤ 2 h [% of samples]
1-sample, 12-gene	morning	0.69 [0.82]	63.0	96.3
	afternoon	0.80 [0.62]	60.7	92.9
1-sample, 2-gene	morning	0.54 [0.87]	71.4	100.0
	afternoon	0.99 [1.00]	50.0	78.6
2-sample, 13-gene	morning/ afternoon	0.74 [1.10]	60.7	92.9
2-sample, 2-gene	morning/ afternoon	0.75 [0.72]	64.3	92.9

5

Table 7: Gene ontology (GO) functional enrichment analysis.

GO.ID	Term	Annotated	Significant	Expected	p-value
GO:0002376	immune system process	1508	54	21.85	6.00E-12
GO:0009605	response to external stimulus	923	38	13.37	7.80E-10
GO:0042221	response to chemical	1836	55	26.6	4.30E-09
GO:0043207	response to external biotic stimulus	433	23	6.27	4.20E-08
GO:0051707	response to other organism	433	23	6.27	4.20E-08
GO:0050896	response to stimulus	3805	83	55.13	8.00E-08
GO:0006955	immune response	1092	38	15.82	8.60E-08
GO:0009607	response to biotic stimulus	454	23	6.58	1.00E-07
GO:0051704	multi-organism process	1209	40	17.52	1.40E-07
GO:0006952	defense response	750	30	10.87	1.60E-07
GO:0051239	regulation multicell. organism. process	1135	38	16.45	2.40E-07
GO:0048518	positive regulation of biological process	2685	65	38.9	3.90E-07
GO:0032501	multicellular organismal process	2622	64	37.99	3.90E-07
GO:0048856	anatomical structure development	2200	57	31.88	4.40E-07
GO:0010033	response to organic substance	1480	44	21.44	5.50E-07
GO:0044707	single-multicellular organism process	2460	61	35.64	5.80E-07
GO:0001775	cell activation	807	30	11.69	8.00E-07
GO:0070887	cellular response to chemical stimulus	1396	42	20.23	8.40E-07
GO:0032502	developmental process	2429	60	35.19	9.30E-07
GO:0080144	regulation of response to stress	773	29	11.2	1.10E-06

GO analysis performed on the 119 genes extracted by ZeitZeiger at least once by RNA-Seq data based internal cross-validation (see Figure 6B). As background set, 9115 genes identified as expressed in monocytes across all subjects were used.

Table 8: External validation of the BodyTime predictors in the independent VALI study using LASSO or PLS.

Predictor	Type of validation sample	Absolute prediction error median [IQR]	Absolute prediction error ≤ 1 h [% of samples]	Absolute prediction error ≤ 2 h [% of samples]
1-sample, LASSO	morning	0.59 [0.76]	72.4	100.0
	afternoon	0.82 [0.94]	69.0	96.6
1-sample, PLS	morning	0.66 [0.93]	65.5	93.1
	afternoon	0.94 [0.89]	51.7	93.1
2-sample, LASSO	morning/afternoon	0.62 [0.72]	72.4	100.0
2-sample, PLS	morning/afternoon	0.68 [0.75]	72.4	100.0

Claims

1. A method of assessing a time-related physiological parameter selected from

   - a person's circadian rhythm,
   - a person's internal time and
   - the strength of a person's circadian rhythm,

   said method comprising the steps of

   a. determining, in a measurement step, an expression level of each of a plurality of circadian oscillatory genes in a sample of cells obtained from said person;
   b. assessing, in a calculation step, said time-related physiological parameter based on the expression level determined of each of said plurality of circadian oscillatory genes in said measurement step;

   characterized in that said cells are blood monocytes and wherein said plurality of circadian oscillatory genes comprises at least 2, at least 4, at least 8, at least 10, at least 16, at least 20, or at least 32 of the genes of table 1,

   particularly wherein said plurality of circadian oscillatory genes comprises at least 2, at least 4, at least 8, at least 10, at least 16 or at least 20, of the genes of table 1A.

2. The method according to claim 1, wherein said plurality of circadian oscillatory genes consists of at least 2, at least 4, at least 8, at least 10, at least 16, at least 20, or at least 32 genes selected from the genes of table 1, particularly wherein said plurality of circadian oscillatory genes consists of at least 2, at least 4, at least 8, at least 10, at least 16 or at least 20 selected from the genes of table 1A.

3. The method according to any one of the above claims, wherein said plurality of circadian oscillatory genes comprises one of PER2 or NR1D2, or both.

4. The method according to any one of the above claims, wherein said plurality of circadian oscillatory genes consists of PER2 and NR1D2.

5. The method according to any one of the above claims, wherein said expression levels are determined in a single sample.

6. The method according to claim 5, wherein an expression level of at least 2, at least 4, at least 8, at least 10, at least 16, at least 20, or at least 32 of the genes of table 2 is determined.

7. The method according to any one of the above claims, wherein said expression level is determined in each of two samples, wherein said two samples are obtained at a time of day 2 - 10 hours apart, particularly at least 4 - 8 hours apart, more particularly approximately 6 hours apart.

8. The method according to claim 7, wherein an expression level of at least 2, at least 4, at least 8, at least 10, at least 16, at least 20, or at least 32 of the genes of table 3 is determined.

9. The method according to any one of the above claims, wherein said calculation step is effected by using said expression levels determined in said measurement step as input values, and applying an algorithm to said input values, thereby generating an output value, wherein in particular said algorithm is selected from

    a. the ZeitZeiger algorithm;

    b. the Molecular-timetable algorithm;

    c. partial least squares regression

    and/or

    d. LASSO.

10. The method according to claim 9, wherein said output value corresponds to

    a. internal time or
    b. strength of circadian rhythm.

11. The method according to any one of the above claims, wherein said calculation step comprises

    a. determining a first value SPC1 based on the expression level of a first subset of said plurality of circadian oscillatory genes;
    b. determining a second value SPC2 based on the expression level of a second subset of said plurality of circadian oscillatory genes.

12. The method according to claim 11, wherein said first subset comprises PER2 and said second subset comprises NR1D2.

13. The method according to claim 11 to 12, wherein

    said first value SPC1 is determined by multiplying a measured value and a constant value for each gene comprised in said first subset, thereby obtaining a product, and subsequently adding the products obtained for each gene, thereby obtaining said first value; and

said second value SPC2 is determined by multiplying a measured value and a constant value for each gene comprised in said second subset, thereby obtaining a product, and subsequently adding the products obtained for each gene, thereby obtaining said second value,

wherein said measured value corresponds to said relative expression level obtained in said measurement step and said constant value corresponds to a loading coefficient specific for each gene comprised in said plurality of genes.

14. The method according to any one of the above claims, wherein said expression level is determined using a method selected from quantitative PCR (qPCR), NanoString and microarray, in particular using the NanoString method.

Fig. 1

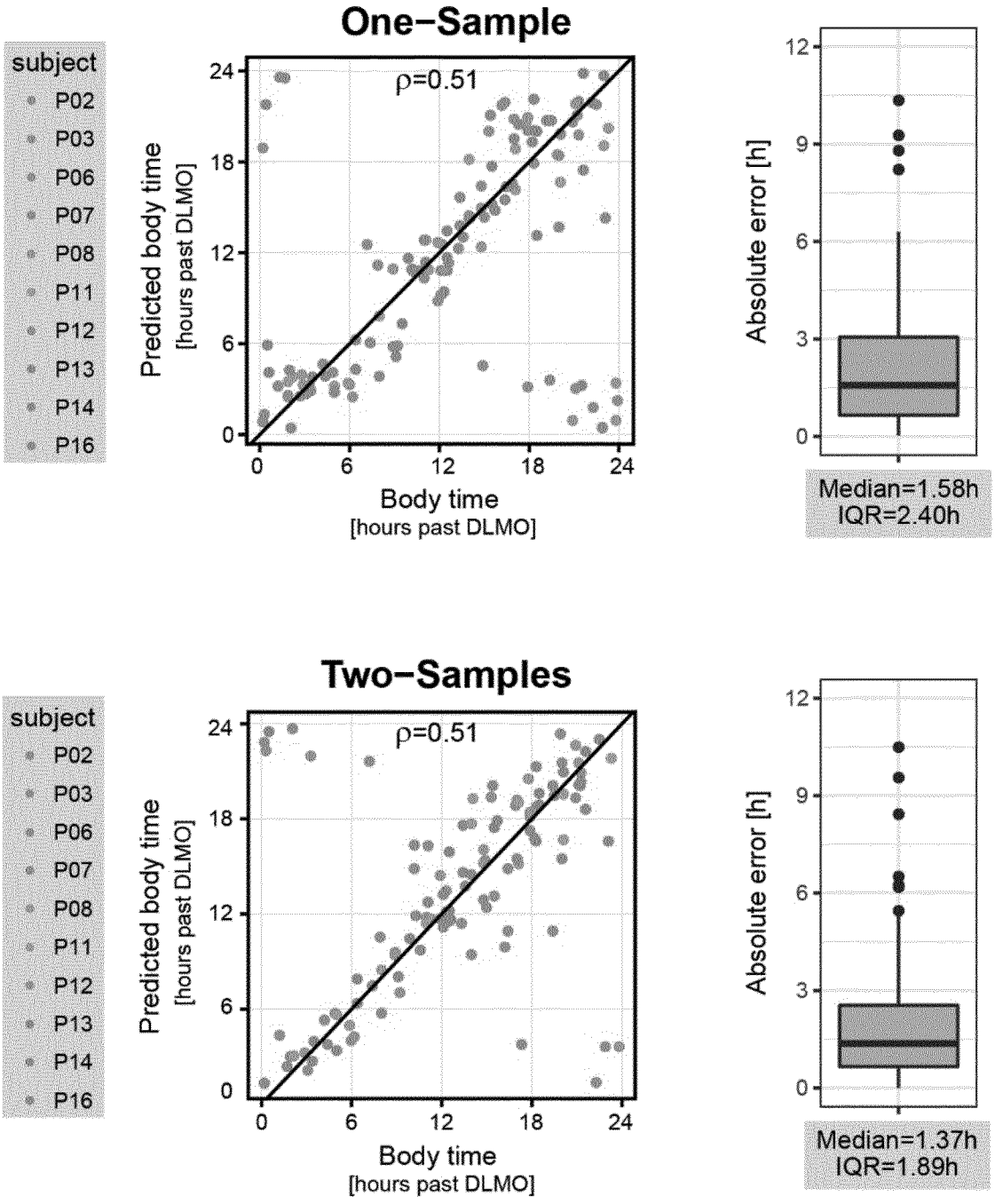

One−Sample
subject
P02
P03
P06
P07
P08
P11
P12
P13
P14
P16
ρ=0.51
Predicted body time
[hours past DLMO]
Body time
[hours past DLMO]
Absolute error [h]
Median=1.58h
IQR=2.40h
Two−Samples
subject
P02
P03
P06
P07
P08
P11
P12
P13
P14
P16
ρ=0.51
Predicted body time
[hours past DLMO]
Body time
[hours past DLMO]
Absolute error [h]
Median=1.37h
IQR=1.89h

Fig. 2

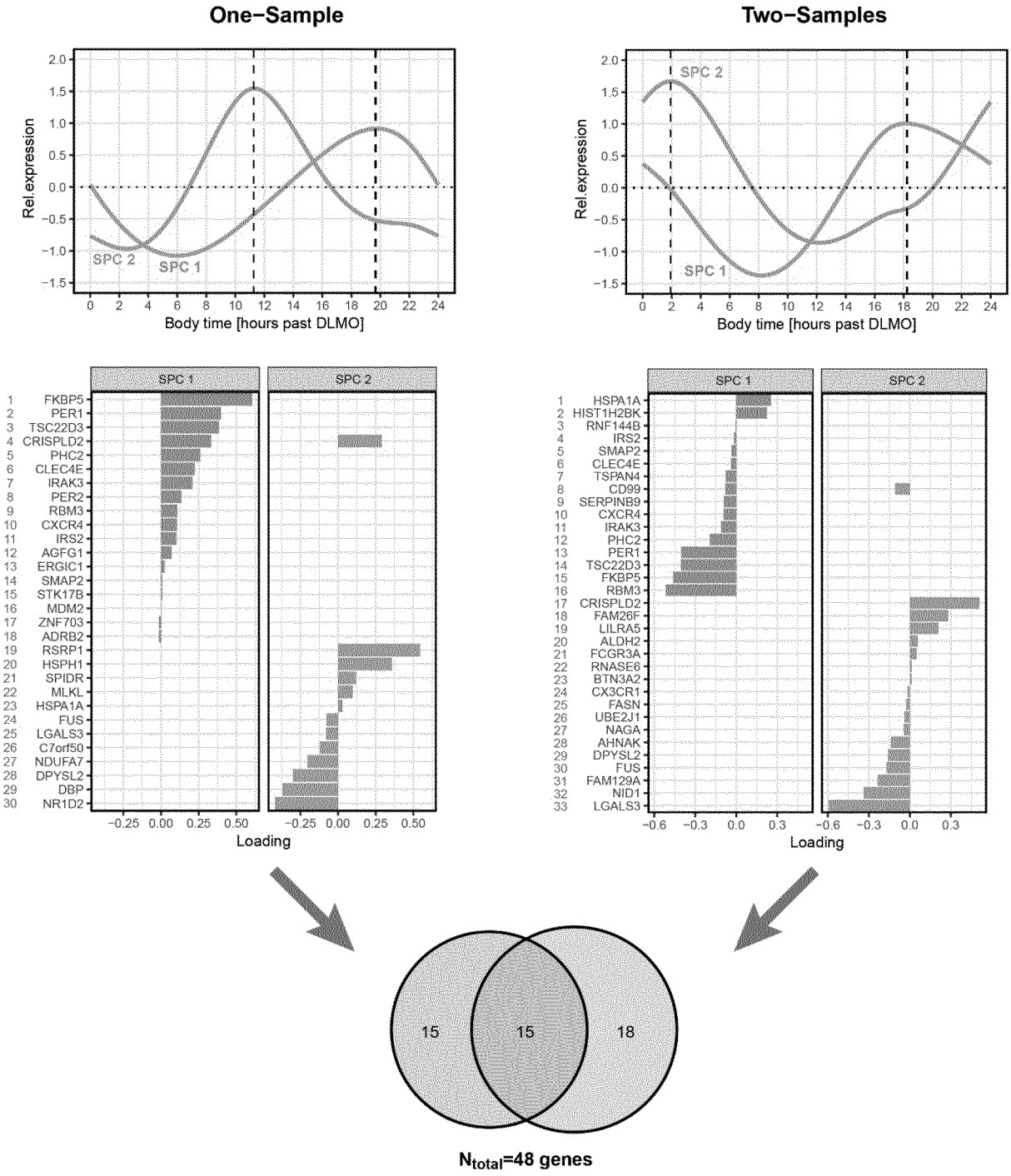

One−Sample
Two−Samples
Rel.expression
Body time [hours past DLMO]
SPC 1
SPC 2
Loading
1 FKBP5
2 PER1
3 TSC22D3
4 CRISPLD2
5 PHC2
6 CLEC4E
7 IRAK3
8 PER2
9 RBM3
10 CXCR4
11 IRS2
12 AGFG1
13 ERGIC1
14 SMAP2
15 STK17B
16 MDM2
17 ZNF703
18 ADRB2
19 RSRP1
20 HSPH1
21 SPIDR
22 MLKL
23 HSPA1A
24 FUS
25 LGALS3
26 C7orf50
27 NDUFA7
28 DPYSL2
29 DBP
30 NR1D2
1 HSPA1A
2 HIST1H2BK
3 RNF144B
4 IRS2
5 SMAP2
6 CLEC4E
7 TSPAN4
8 CD99
9 SERPINB9
10 CXCR4
11 IRAK3
12 PHC2
13 PER1
14 TSC22D3
15 FKBP5
16 RBM3
17 CRISPLD2
18 FAM26F
19 LILRA5
20 ALDH2
21 FCGR3A
22 RNASE6
23 BTN3A2
24 CX3CR1
25 FASN
26 UBE2J1
27 NAGA
28 AHNAK
29 DPYSL2
30 FUS
31 FAM129A
32 NID1
33 LGALS3
15
15
18
N_{total}=48 genes

Fig. 3

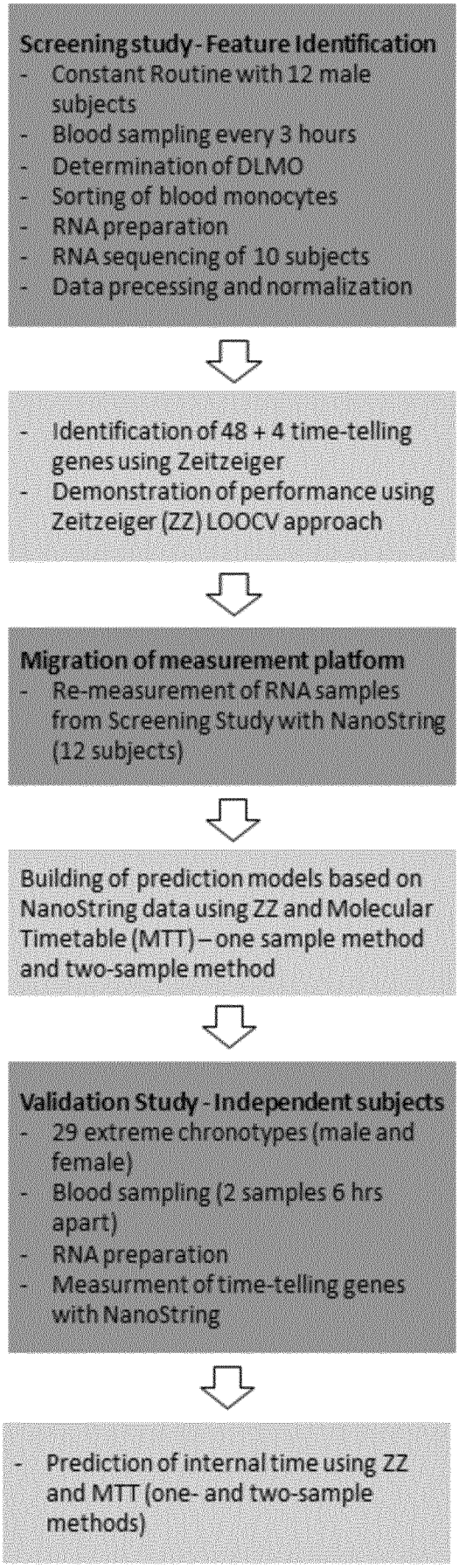

Screening study - Feature Identification
- Constant Routine with 12 male subjects
- Blood sampling every 3 hours
- Determination of DLMO
- Sorting of blood monocytes
- RNA preparation
- RNA sequencing of 10 subjects
- Data precessing and normalization
- Identification of 48 + 4 time-telling genes using Zeitzeiger
- Demonstration of performance using Zeitzeiger (ZZ) LOOCV approach
Migration of measurement platform
- Re-measurement of RNA samples from Screening Study with NanoString (12 subjects)
Building of prediction models based on NanoString data using ZZ and Molecular Timetable (MTT) – one sample method and two-sample method
Validation Study - Independent subjects
- 29 extreme chronotypes (male and female)
- Blood sampling (2 samples 6 hrs apart)
- RNA preparation
- Measurment of time-telling genes with NanoString
- Prediction of internal time using ZZ and MTT (one- and two-sample methods)

Fig. 4

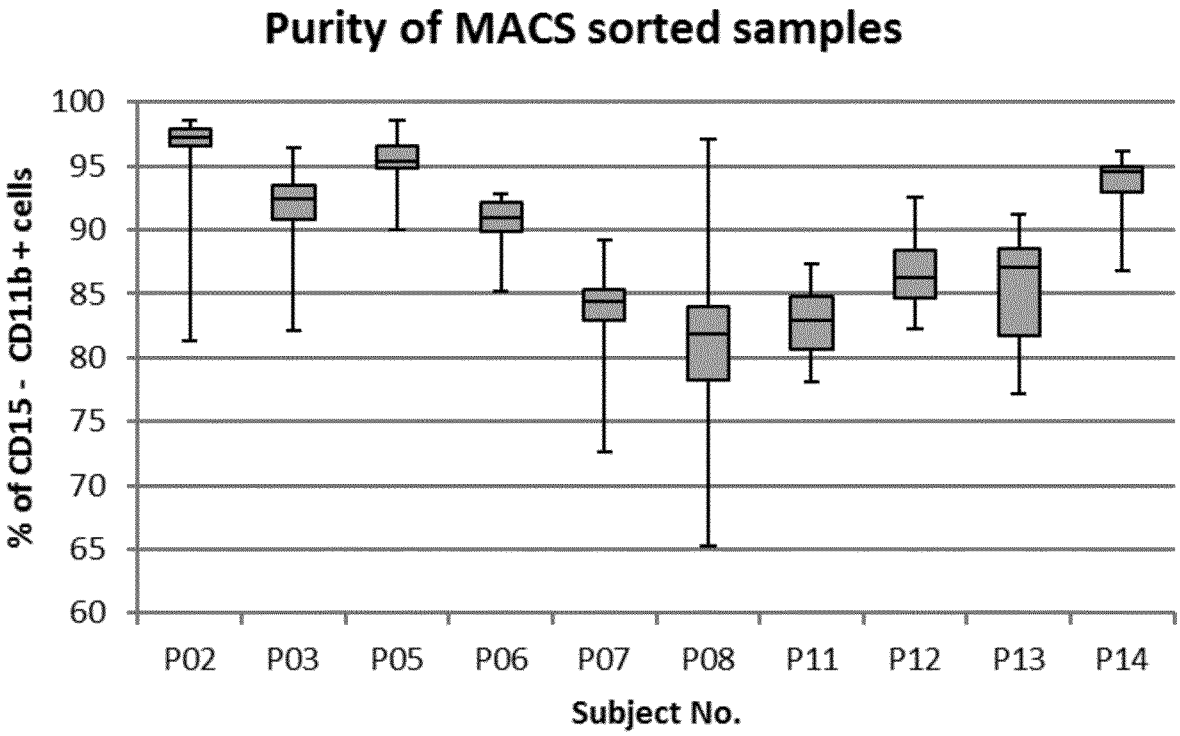

Purity of MACS sorted samples
% of CD15 - CD11b + cells
100
95
90
85
80
75
70
65
60
P02
P03
P05
P06
P07
P08
P11
P12
P13
P14
Subject No.

Fig. 5

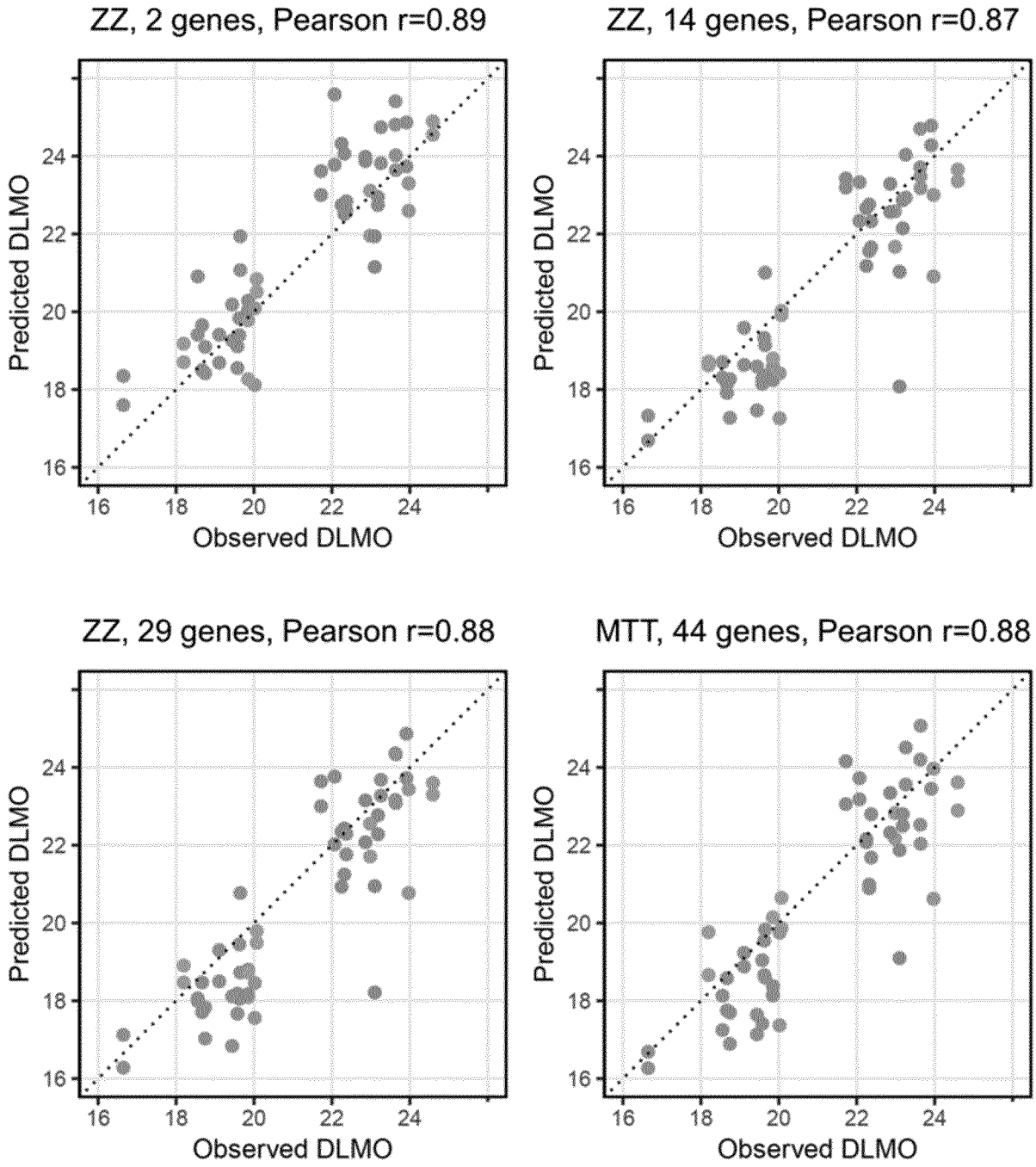

ZZ, 2 genes, Pearson r=0.89
ZZ, 14 genes, Pearson r=0.87
ZZ, 29 genes, Pearson r=0.88
MTT, 44 genes, Pearson r=0.88
Predicted DLMO
Observed DLMO
16
18
20
22
24

Fig. 6

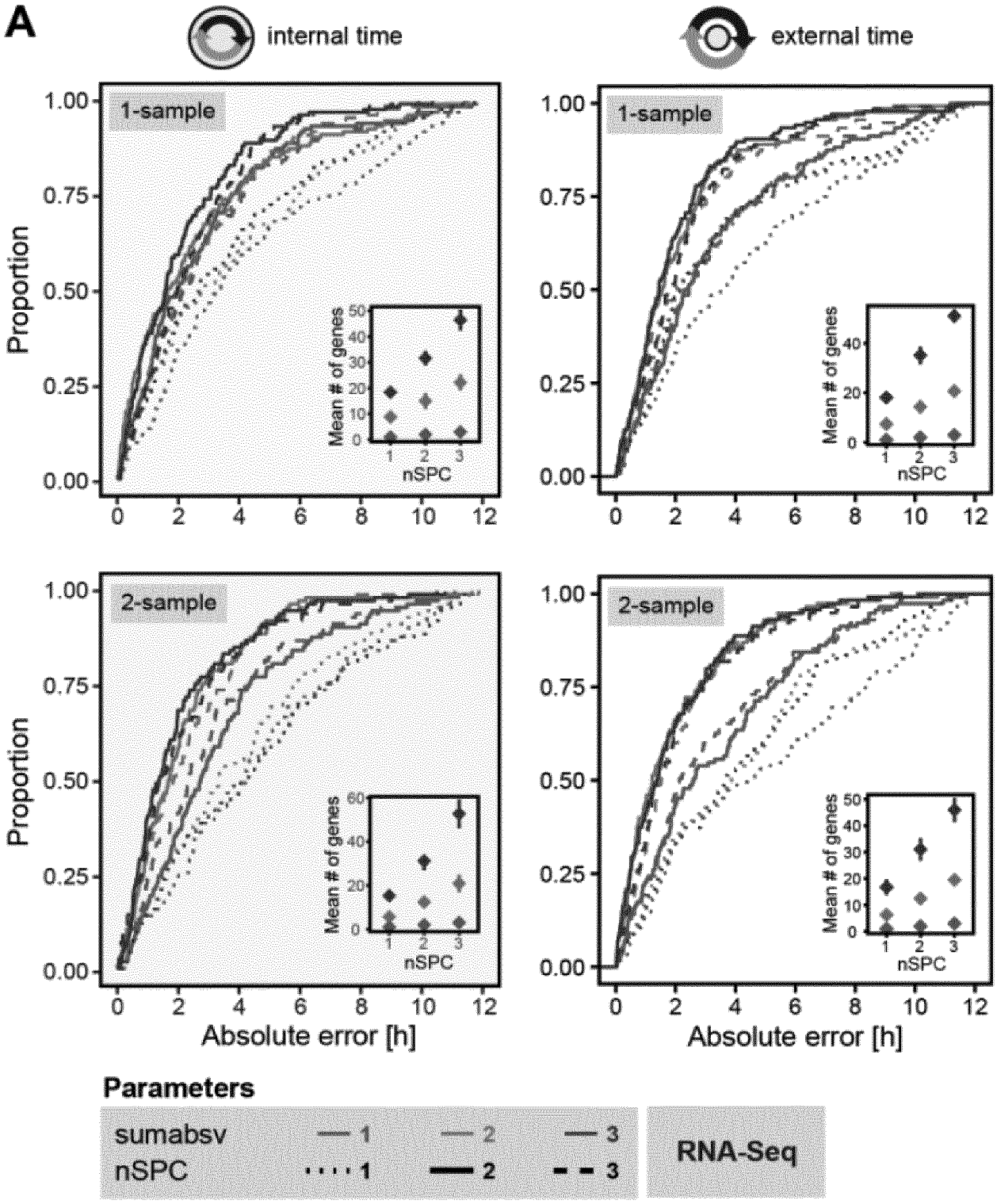

A
internal time
external time
1-sample
2-sample
Proportion
Absolute error [h]
Mean # of genes
nSPC
Parameters
sumabsv 1 2 3
nSPC 1 2 3
RNA-Seq

Fig. 6 (continued)

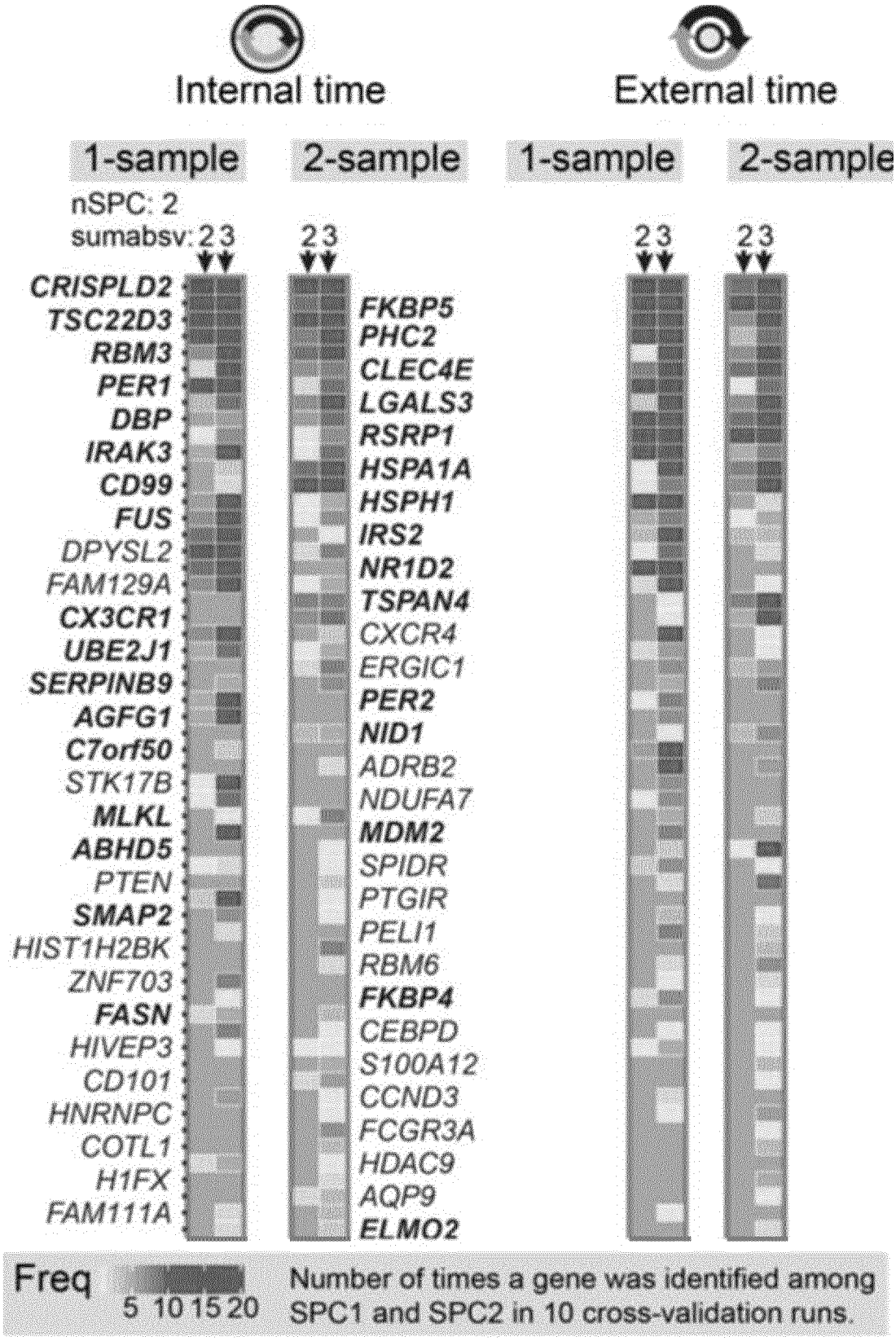


Fig. 6 (continued)

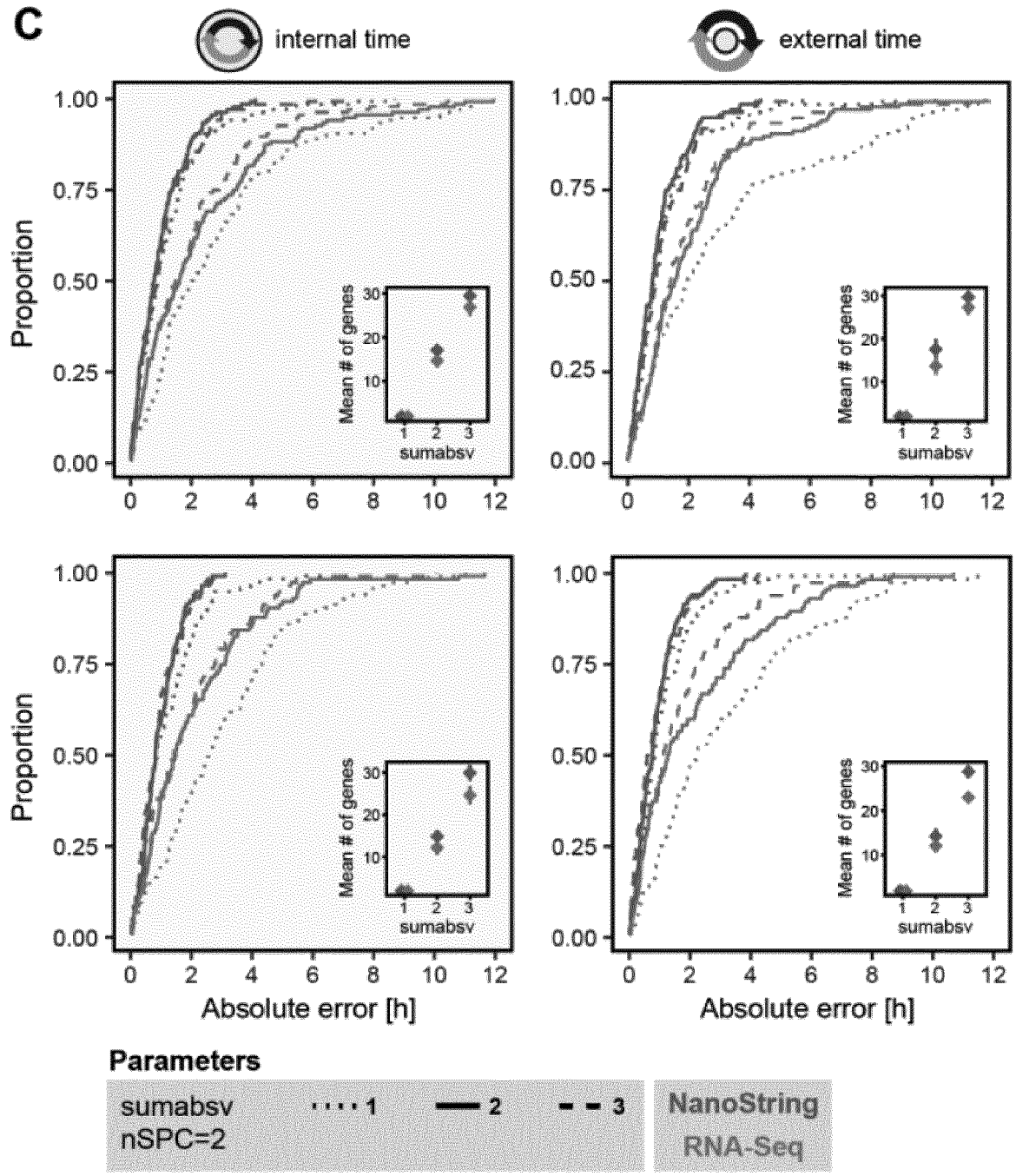


Fig. 7

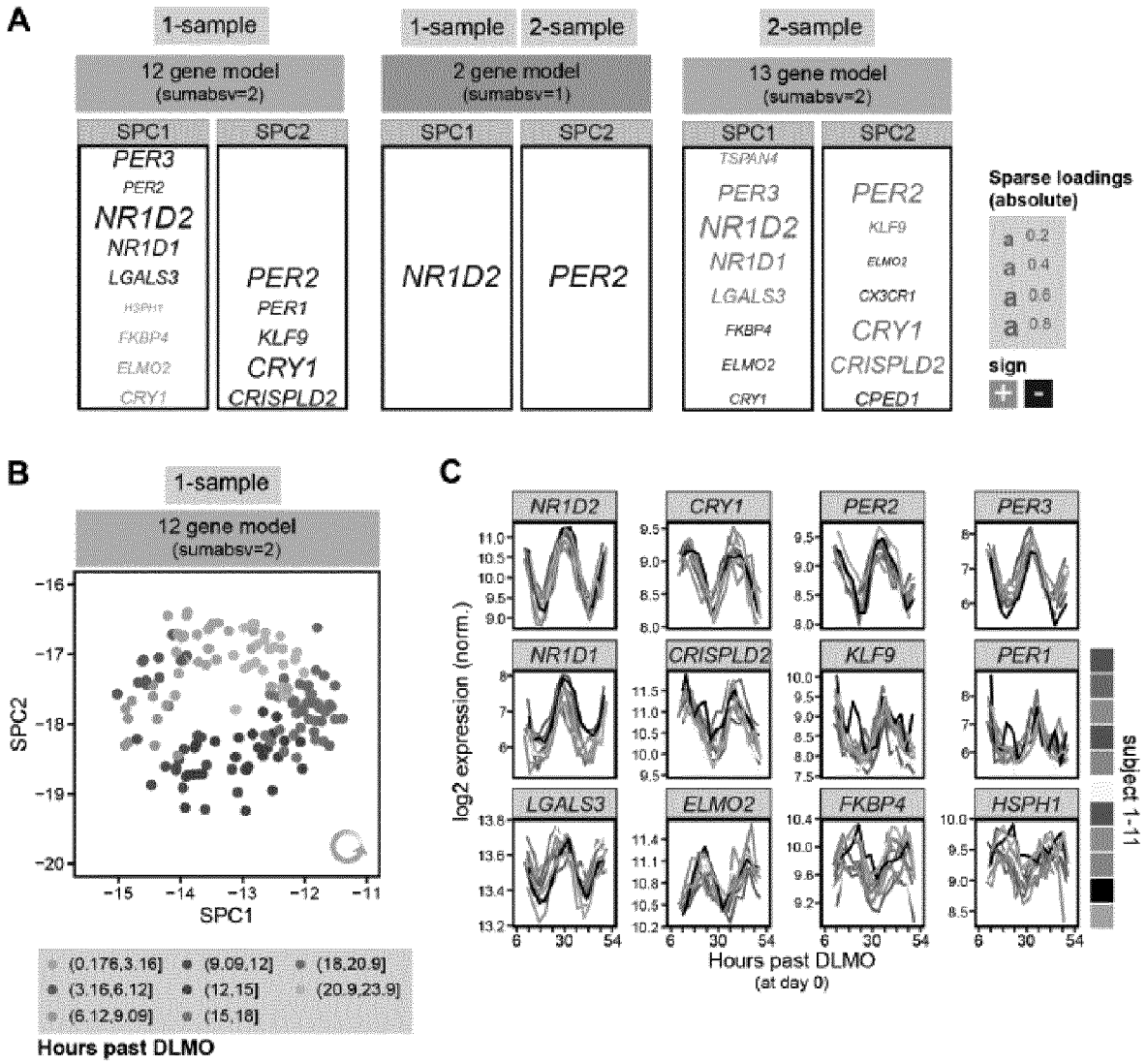

A
1-sample
12 gene model
(sumabsv=2)
SPC1
PER3
PER2
NR1D2
NR1D1
LGALS3
HSPH1
FKBP4
ELMO2
CRY1
SPC2
PER2
PER1
KLF9
CRY1
CRISPLD2
1-sample
2-sample
2 gene model
(sumabsv=1)
NR1D2
PER2
2-sample
13 gene model
(sumabsv=2)
TSPAN4
PER3
NR1D2
NR1D1
LGALS3
FKBP4
ELMO2
CRY1
PER2
KLF9
ELMO2
CX3CR1
CRY1
CRISPLD2
CPED1
Sparse loadings
(absolute)
sign
B
1-sample
12 gene model
(sumabsv=2)
SPC2
SPC1
(0.176,3.16]
(3.16,6.12]
(6.12,9.09]
(9.09,12]
(12,15]
(15,18]
(18,20.9]
(20.9,23.9]
Hours past DLMO
C
NR1D2
CRY1
PER2
PER3
NR1D1
CRISPLD2
KLF9
PER1
LGALS3
ELMO2
FKBP4
HSPH1
log2 expression (norm.)
Hours past DLMO
(at day 0)
subject 1-11

Fig. 8

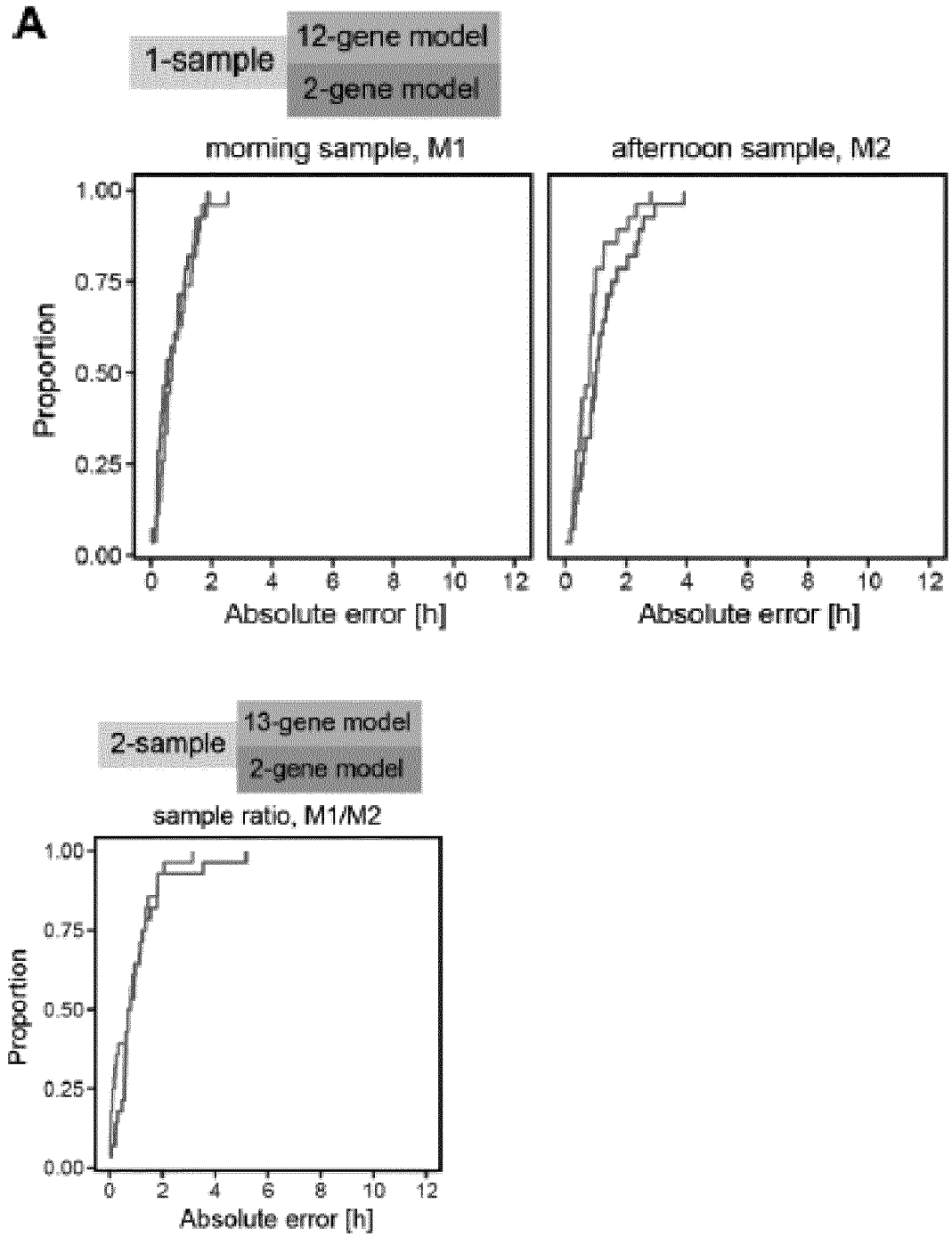

A
1-sample
12-gene model
2-gene model
morning sample, M1
afternoon sample, M2
Proportion
1.00
0.75
0.50
0.25
0.00
0
2
4
6
8
10
12
Absolute error [h]
2-sample
13-gene model
2-gene model
sample ratio, M1/M2

Fig. 8 (continued)

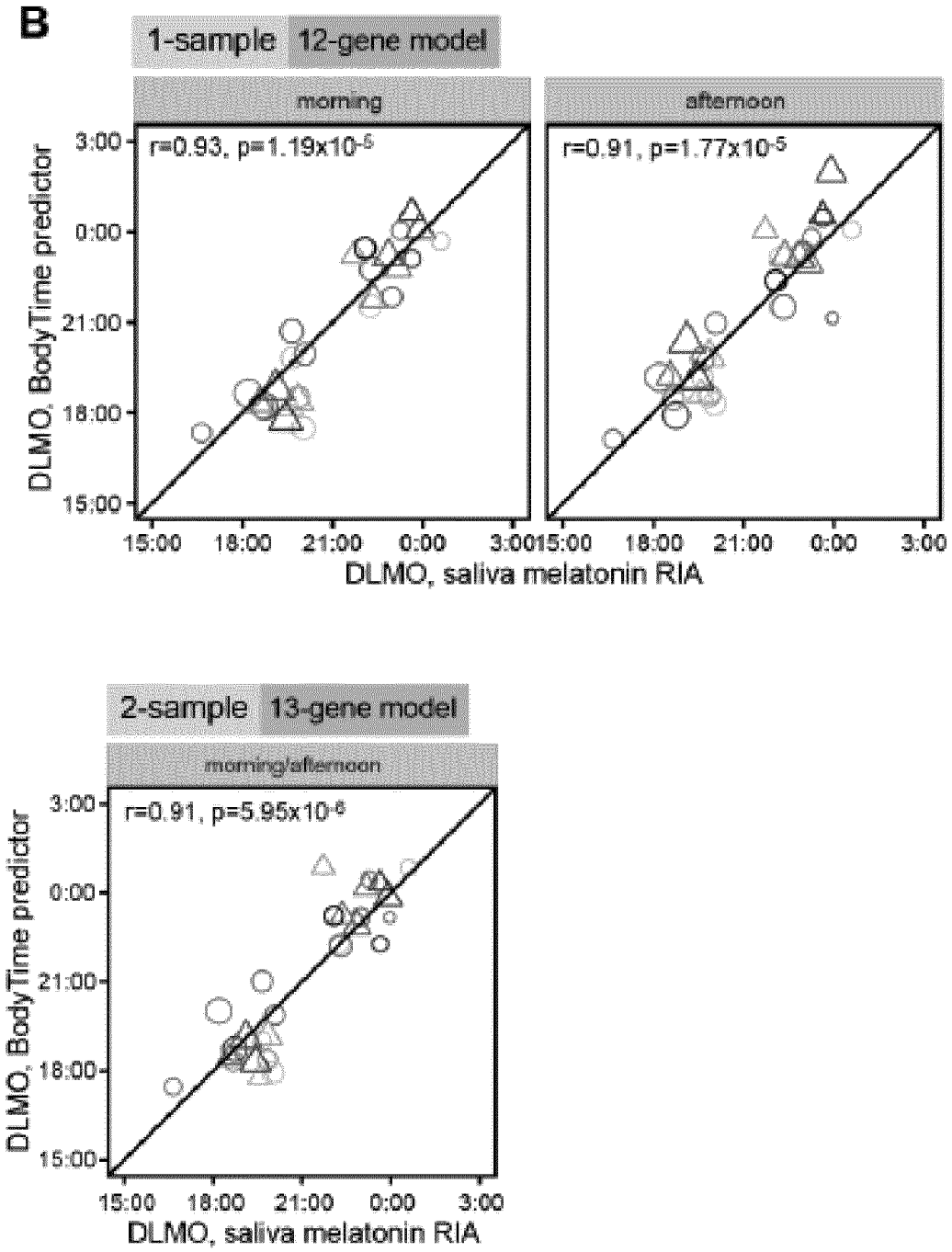


Fig. 8 (continued)

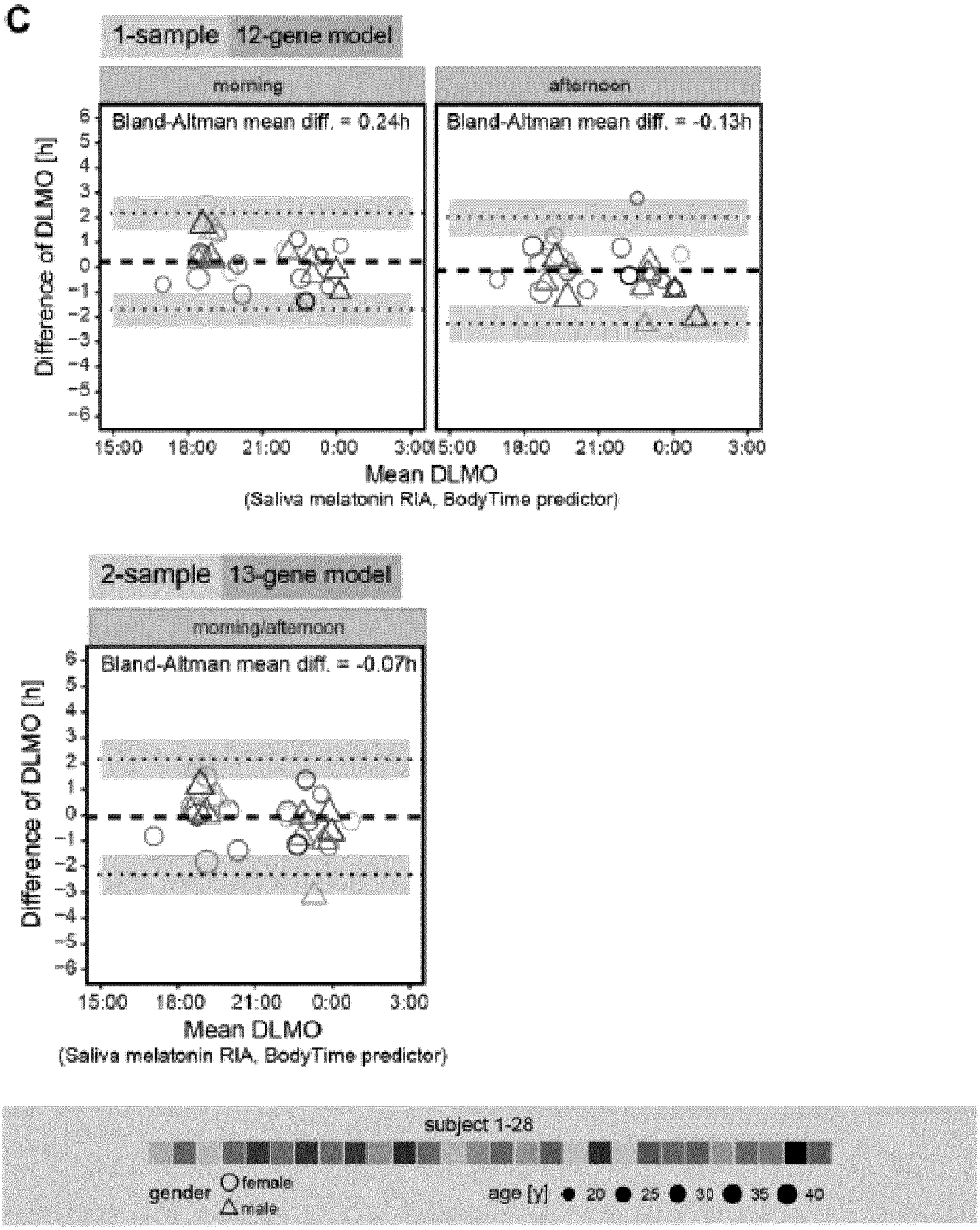


Fig. 9

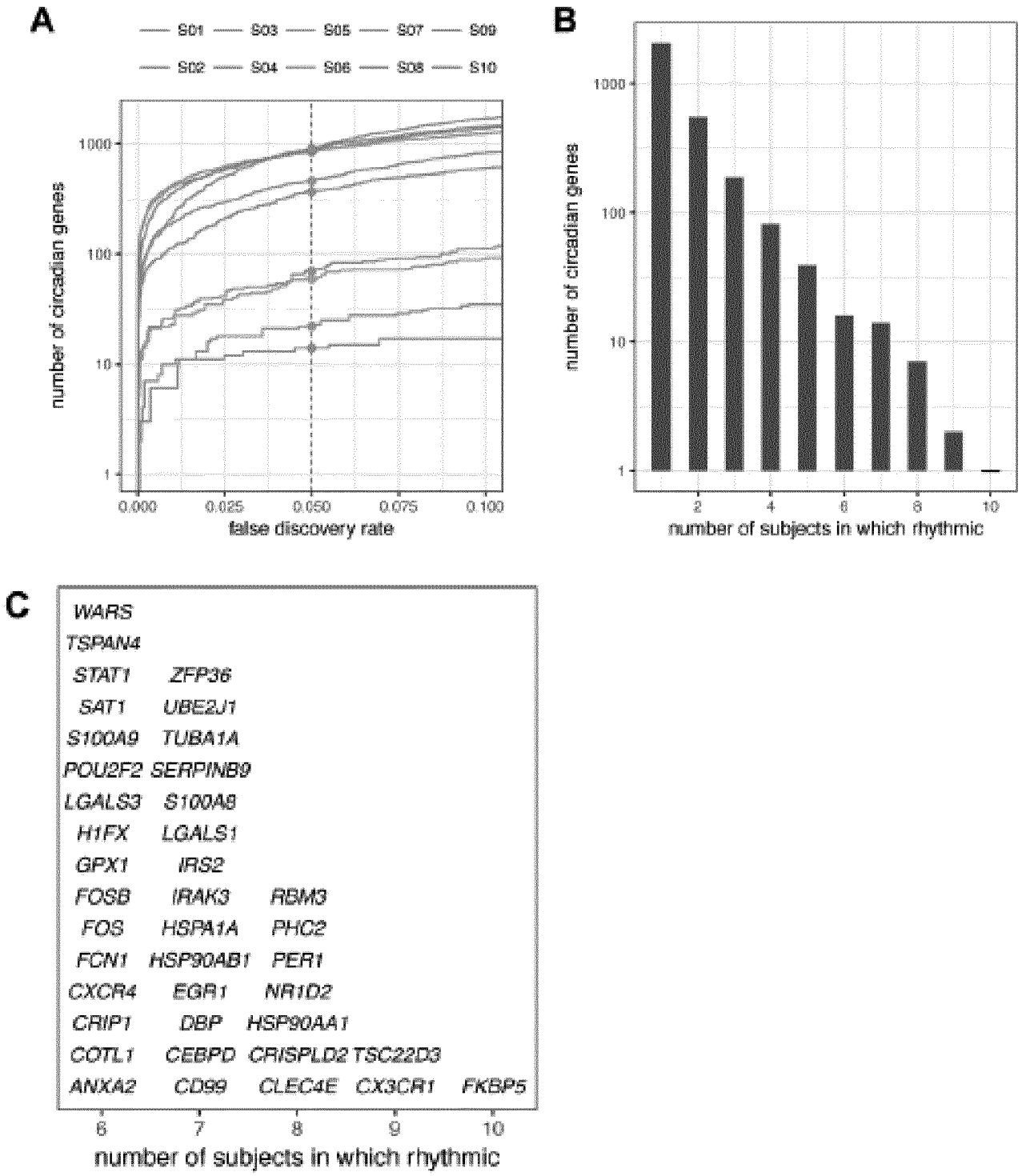
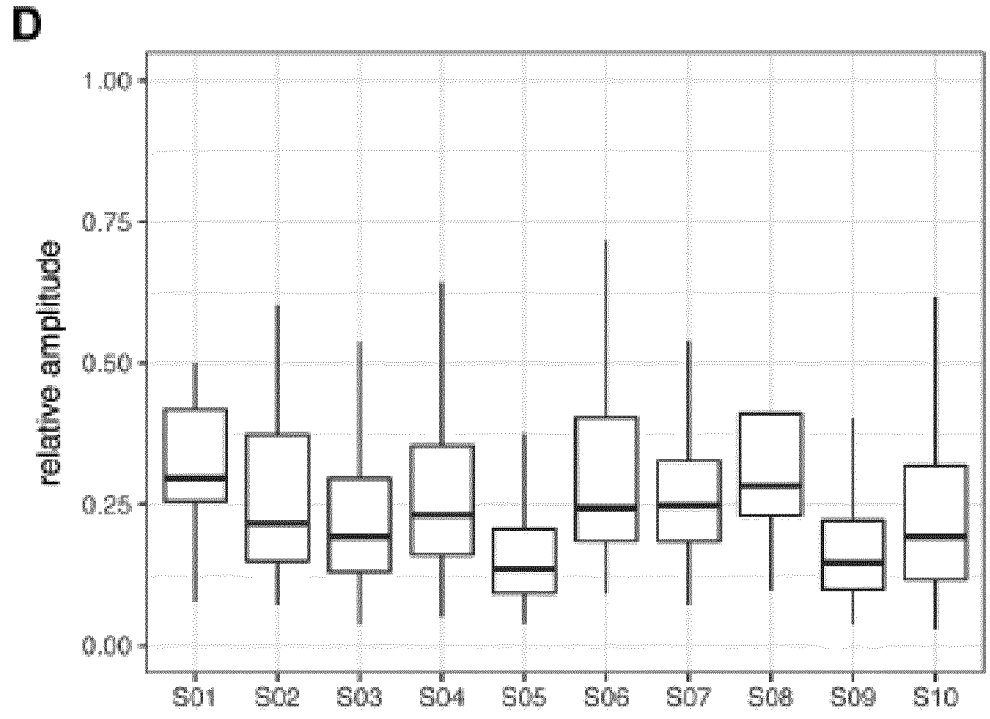

A
S01 S03 S05 S07 S09
S02 S04 S06 S08 S10
number of circadian genes
1000
100
10
1
0.000 0.025 0.050 0.075 0.100
false discovery rate
B
number of circadian genes
1000
100
10
1
2 4 6 8 10
number of subjects in which rhythmic
C
WARS
TSPAN4
STAT1 ZFP36
SAT1 UBE2J1
S100A9 TUBA1A
POU2F2 SERPINB9
LGALS3 S100A8
H1FX LGALS1
GPX1 IRS2
FOSB IRAK3 RBM3
FOS HSPA1A PHC2
FCN1 HSP90AB1 PER1
CXCR4 EGR1 NR1D2
CRIP1 DBP HSP90AA1
COTL1 CEBPD CRISPLD2 TSC22D3
ANXA2 CD99 CLEC4E CX3CR1 FKBP5
6 7 8 9 10
number of subjects in which rhythmic

Fig. 9 (continued)



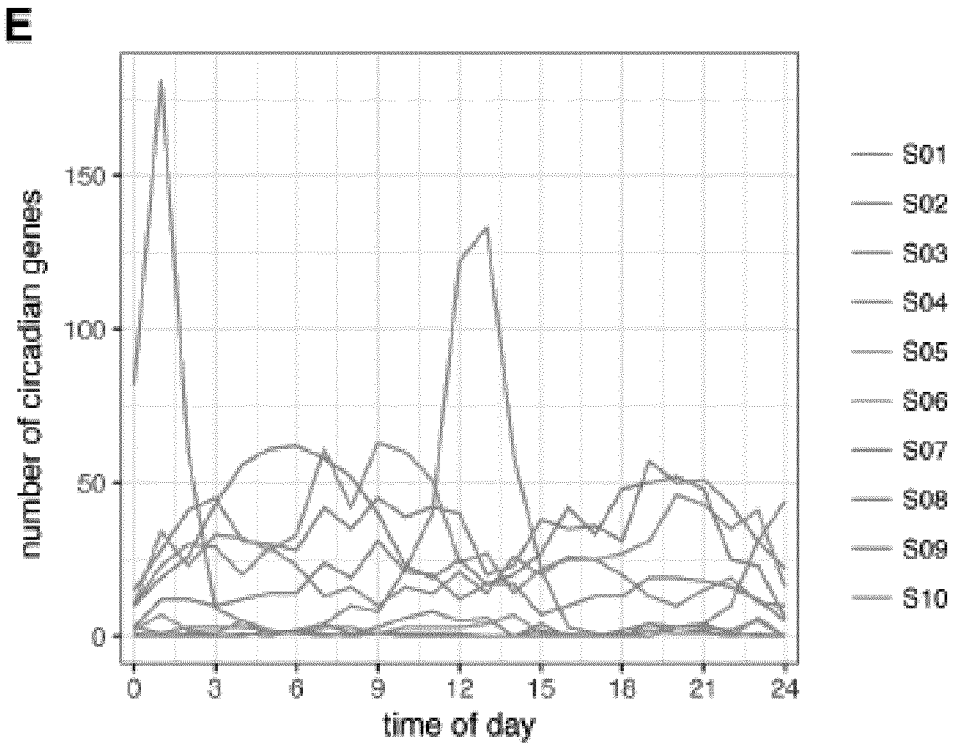


Fig. 10

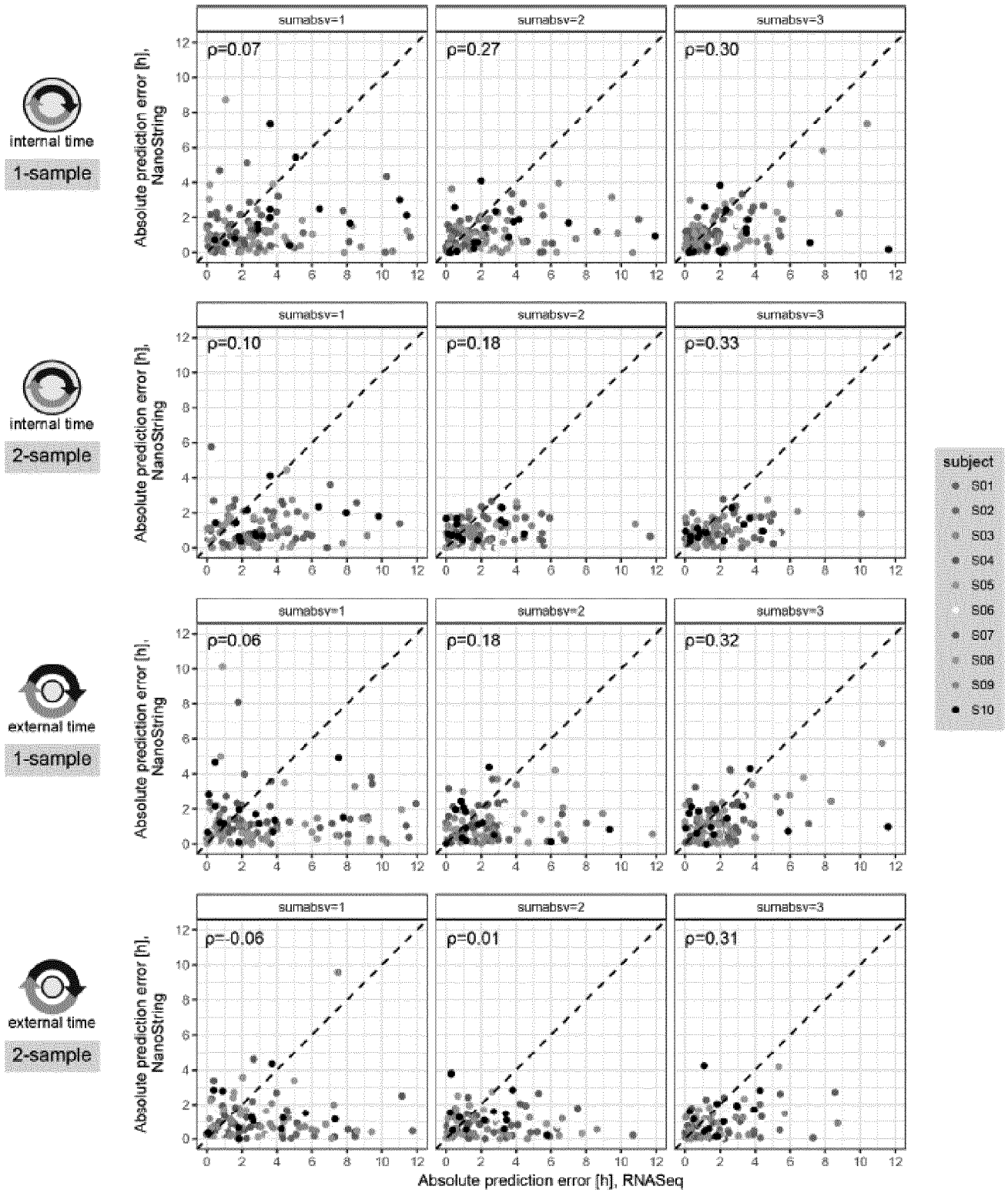

internal time
1-sample
internal time
2-sample
external time
1-sample
external time
2-sample
sumabsv=1
sumabsv=2
sumabsv=3
ρ=0.07
ρ=0.27
ρ=0.30
ρ=0.10
ρ=0.18
ρ=0.33
ρ=0.06
ρ=0.18
ρ=0.32
ρ=-0.06
ρ=0.01
ρ=0.31
Absolute prediction error [h], NanoString
Absolute prediction error [h], RNASeq
subject
S01
S02
S03
S04
S05
S06
S07
S08
S09
S10

Fig. 11

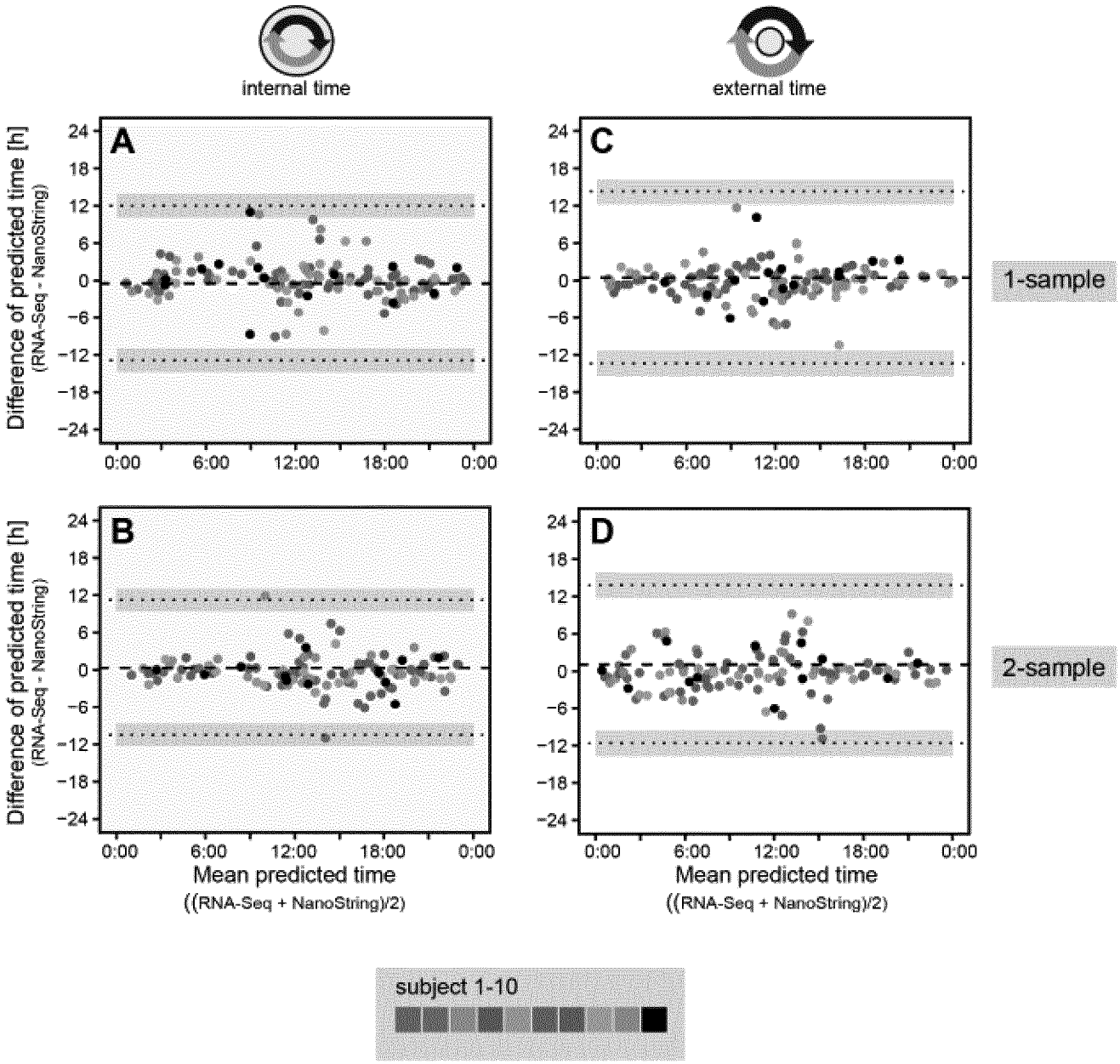
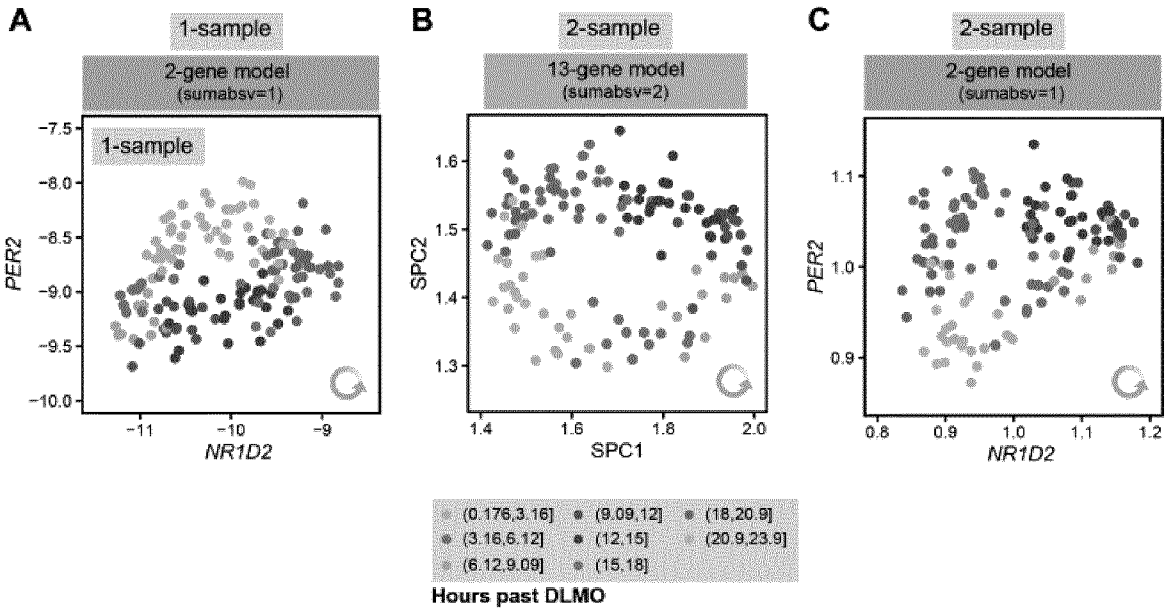

internal time
external time
A
B
C
D
Difference of predicted time [h]
(RNA-Seq - NanoString)
24
18
12
6
0
−6
−12
−18
−24
0:00
6:00
12:00
18:00
0:00
Mean predicted time
((RNA-Seq + NanoString)/2)
1-sample
2-sample
subject 1-10

Fig. 12


A
1-sample
2-gene model
(sumabsv=1)
1-sample
PER2
NR1D2
B
2-sample
13-gene model
(sumabsv=2)
SPC2
SPC1
C
2-sample
2-gene model
(sumabsv=1)
PER2
NR1D2
(0.176,3.16]
(3.16,6.12]
(6.12,9.09]
(9.09,12]
(12,15]
(15,18]
(18,20.9]
(20.9,23.9]
Hours past DLMO

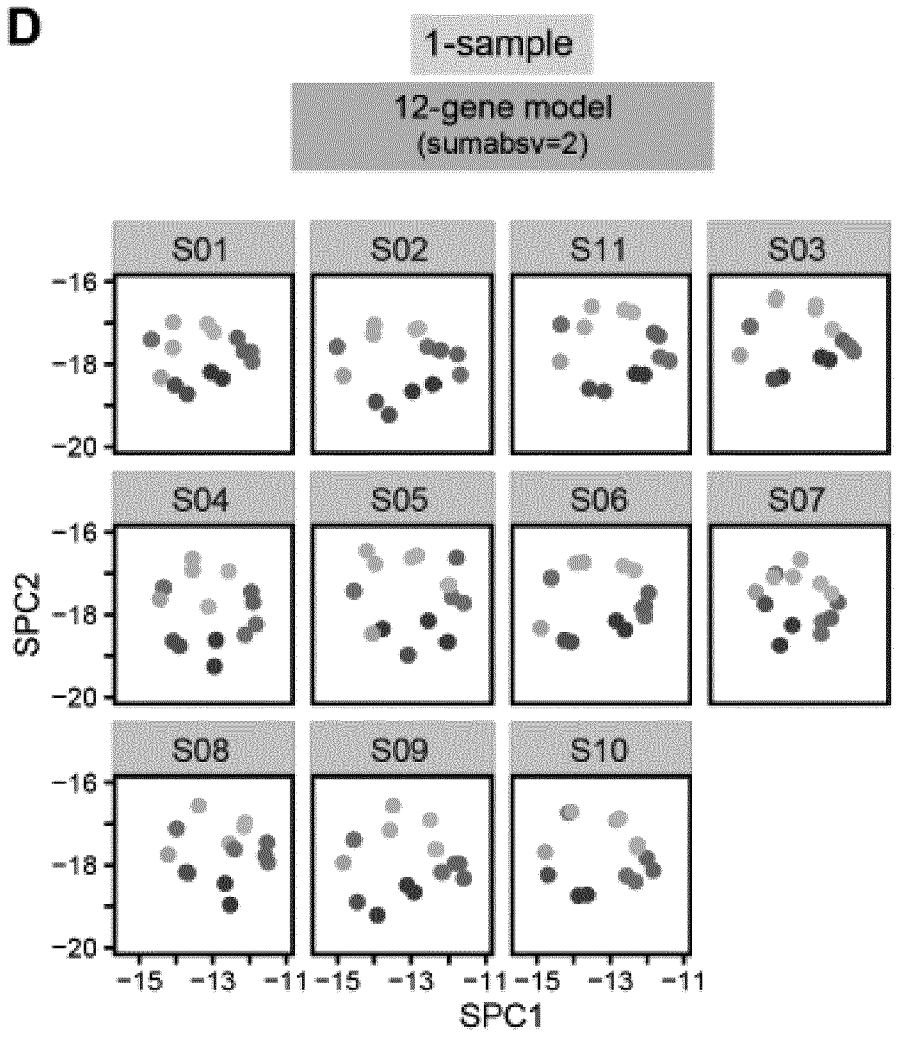

D
1-sample
12-gene model
(sumabsv=2)
S01
S02
S11
S03
S04
S05
S06
S07
S08
S09
S10
SPC2
SPC1

Fig. 13

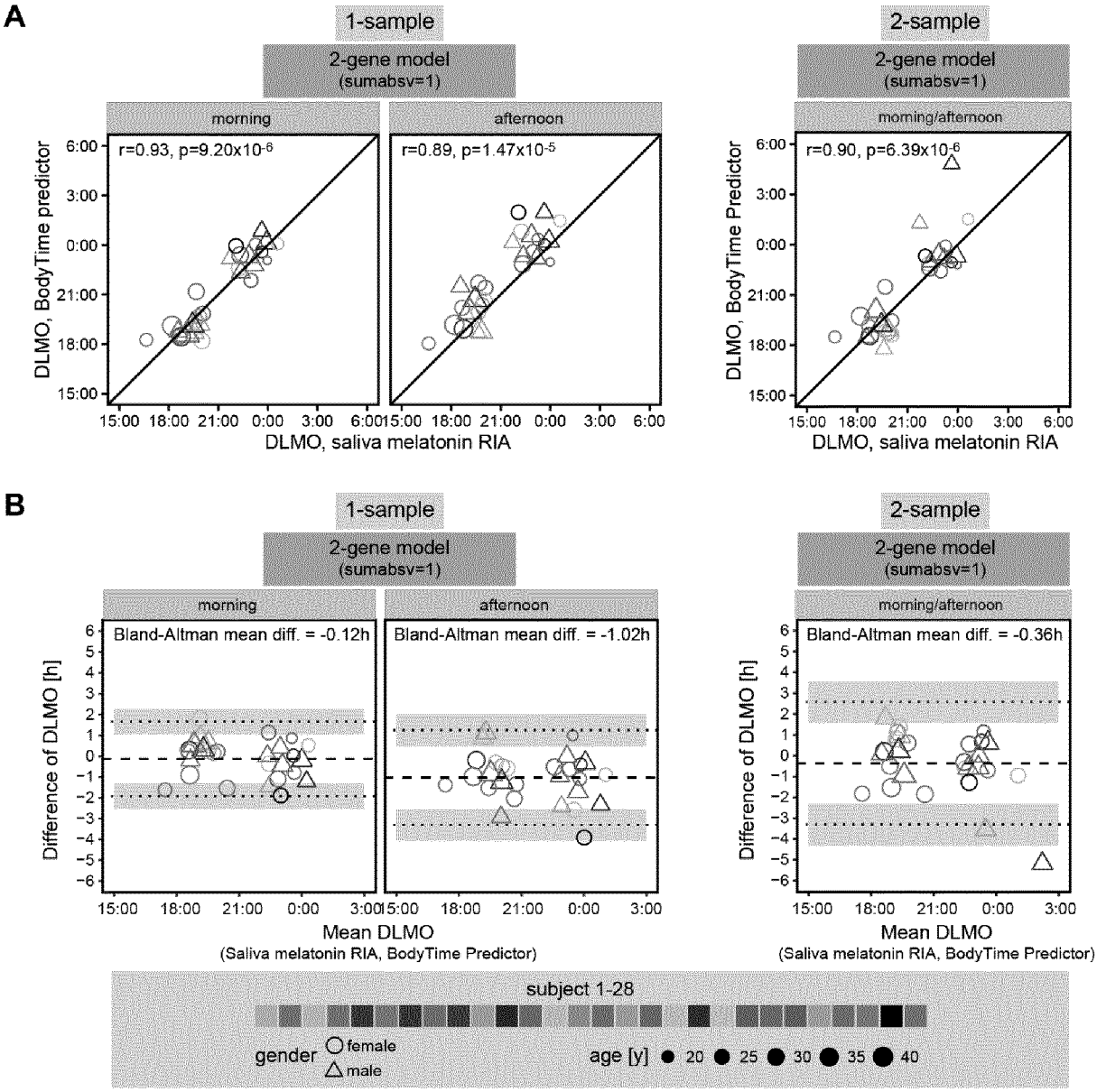

A
1-sample
2-gene model
(sumabsv=1)
morning
afternoon
r=0.93, p=9.20x10-6
r=0.89, p=1.47x10-5
DLMO, BodyTime predictor
DLMO, saliva melatonin RIA
2-sample
2-gene model
(sumabsv=1)
morning/afternoon
r=0.90, p=6.39x10-6
DLMO, BodyTime Predictor
DLMO, saliva melatonin RIA
B
1-sample
2-gene model
(sumabsv=1)
morning
afternoon
Bland-Altman mean diff. = -0.12h
Bland-Altman mean diff. = -1.02h
Difference of DLMO [h]
Mean DLMO
(Saliva melatonin RIA, BodyTime Predictor)
2-sample
2-gene model
(sumabsv=1)
morning/afternoon
Bland-Altman mean diff. = -0.36h
Difference of DLMO [h]
Mean DLMO
(Saliva melatonin RIA, BodyTime Predictor)
subject 1-28
gender female male
age [y] 20 25 30 35 40

Fig. 14

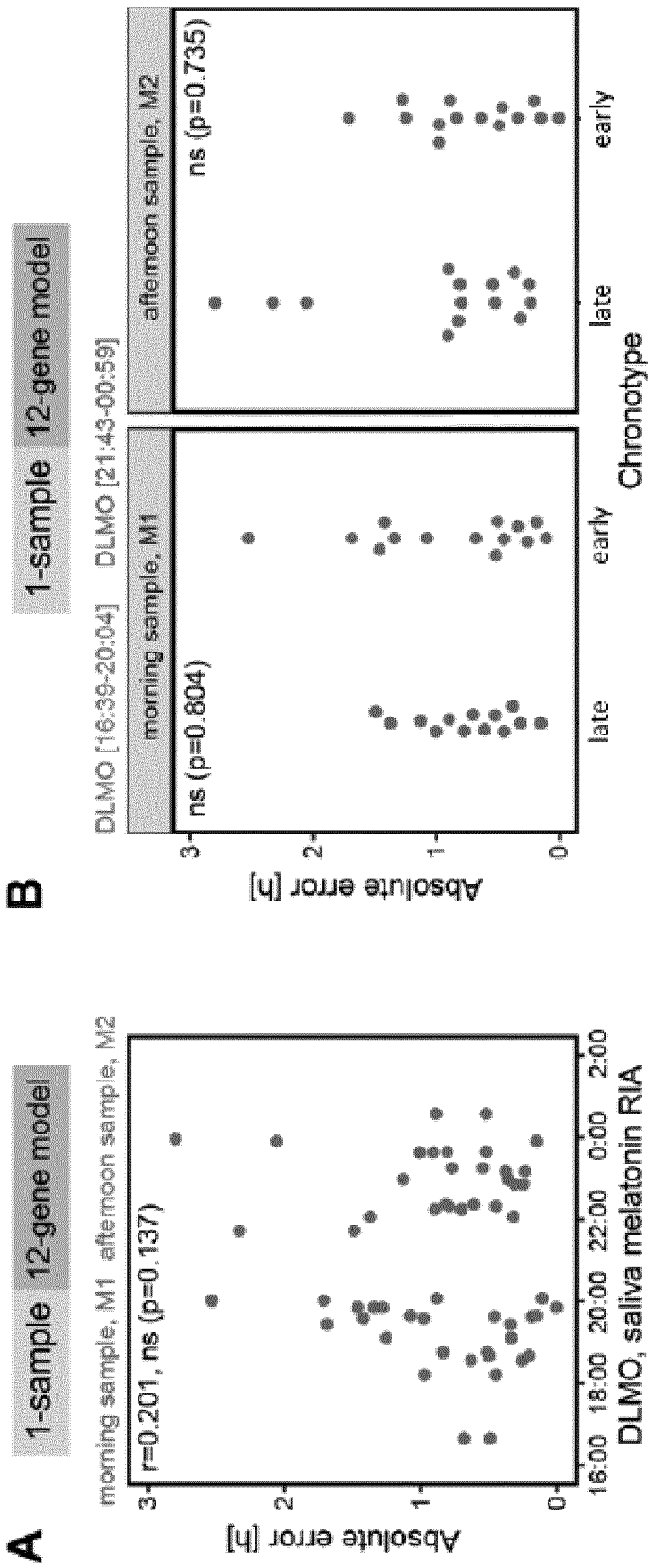
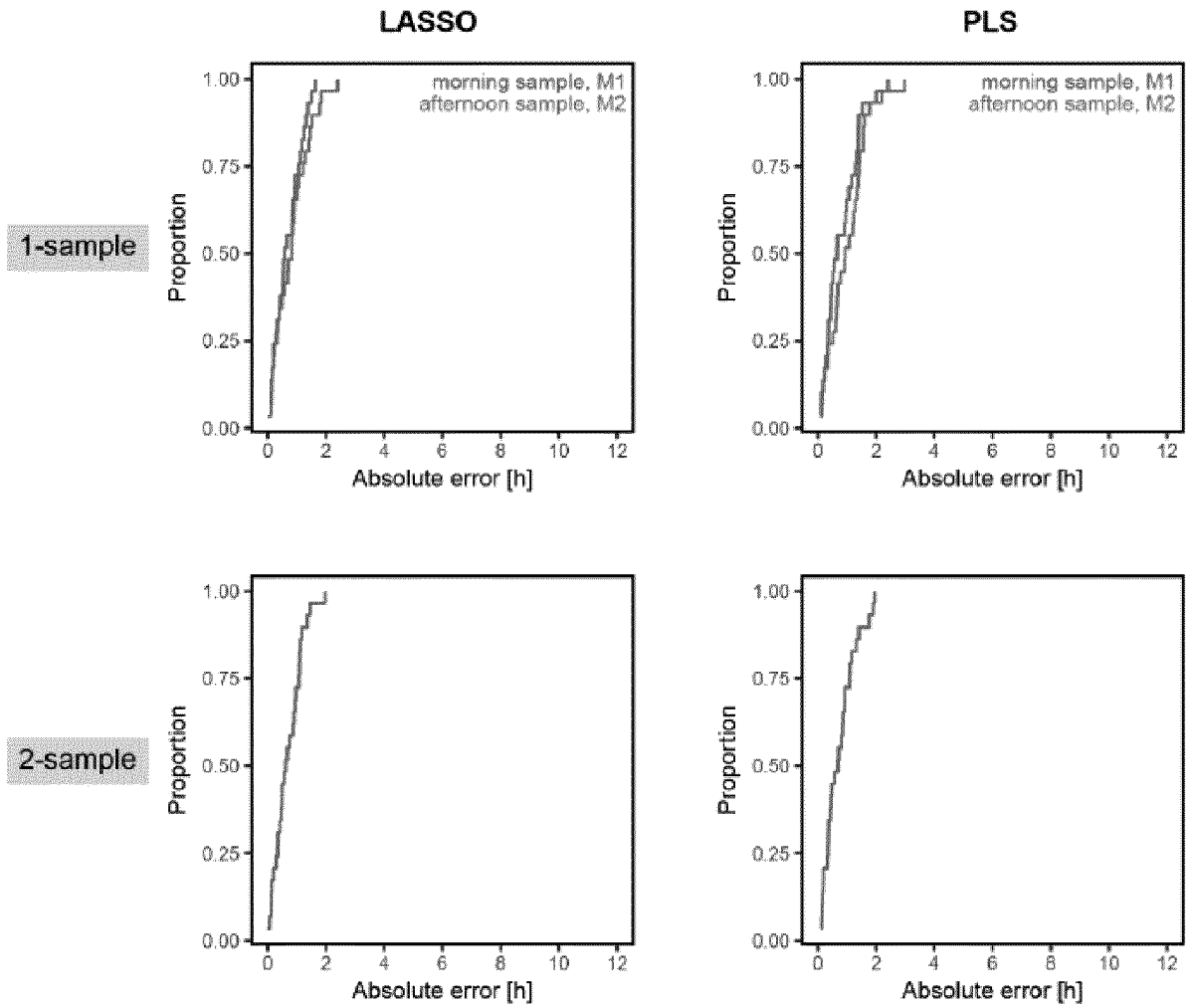

A
1-sample
12-gene model
morning sample, M1
afternoon sample, M2
r=0.201, ns (p=0.137)
Absolute error [h]
DLMO, saliva melatonin RIA
16:00
18:00
20:00
22:00
0:00
2:00
B
1-sample
12-gene model
DLMO [16:39-20:04]
DLMO [21:43-00:59]
morning sample, M1
afternoon sample, M2
ns (p=0.804)
ns (p=0.735)
Absolute error [h]
late
early
Chronotype

Fig. 15


LASSO
PLS
1-sample
2-sample
morning sample, M1
afternoon sample, M2
Proportion
Absolute error [h]
0.00
0.25
0.50
0.75
1.00
0
2
4
6
8
10
12

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2018/066771

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12Q1/6851 C12Q1/6876 C12Q1/6883
ADD.

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12Q

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	PANDI-PERUMAL ET AL: "Dim light melatonin onset (DLMO): A tool for the analysis of circadian phase in human sleep and chronobiological disorders", PROGRESS IN NEURO-PSYCHOPHARMACOLOGY & BIOLOGICAL PSYCHIATRY, ELSEVIER, GB, vol. 31, no. 1, 22 December 2006 (2006-12-22), pages 1-11, XP005813391, ISSN: 0278-5846, DOI: 10.1016/J.PNPBP.2006.06.020 title and abstract -----	1-14
A	WO 2016/038107 A1 (CECCATELLI SANDRA [SE]; SPULBER DAN STEFAN [SE]) 17 March 2016 (2016-03-17) claims 1, 6, Fig. 4 and p. 12 last para -----	1-14

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
19 July 2018	30/07/2018

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Lapopin, Laurence

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	EP 2 175 035 A1 (SONY CORP [JP]) 14 April 2010 (2010-04-14) para. 20-23, Fig. 4 and para. 43	1
A	H. R. UEDA ET AL: "Molecular-timetable methods for detection of body time and rhythm disorders from single-time-point genome-wide expression profiles", PROCEEDINGS NATIONAL ACADEMY OF SCIENCES PNAS, vol. 101, no. 31, 3 August 2004 (2004-08-03), pages 11227-11232, XP055445343, US ISSN: 0027-8424, DOI: 10.1073/pnas.0401882101 abstract and whole document	1-14
Y	JACOB J. HUGHEY ET AL: "ZeitZeiger: supervised learning for high-dimensional data from an oscillatory system", NUCLEIC ACIDS RESEARCH, vol. 44, no. 8, 5 May 2016 (2016-05-05), pages e80-e80, XP055445340, GB ISSN: 0305-1048, DOI: 10.1093/nar/gkw030 abstract and whole document	9,10
X	US 2006/078883 A1 (UEDA HIROKI [JP] ET AL) 13 April 2006 (2006-04-13) cited in the application	2-8, 11-14
Y	para. 12, 13, 50-59, claims 1-8, and Fig. 6	1,9,10
A	RAY ZHANG ET AL: "A circadian gene expression atlas in mammals: Implications for biology and medicine", PROCEEDINGS NATIONAL ACADEMY OF SCIENCES PNAS, vol. 111, no. 45, 27 October 2014 (2014-10-27), pages 16219-16224, XP055446290, US ISSN: 0027-8424, DOI: 10.1073/pnas.1408886111 tab. 1	1-14

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2018/066771

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2016038107 A1	17-03-2016	NONE	
EP 2175035 A1	14-04-2010	CN 101755055 A	23-06-2010
		EP 2175035 A1	14-04-2010
		JP 4858344 B2	18-01-2012
		JP 2009027952 A	12-02-2009
		US 2010196901 A1	05-08-2010
		WO 2009014036 A1	29-01-2009
US 2006078883 A1	13-04-2006	AU 2003252284 A1	16-02-2004
		CA 2494360 A1	05-02-2004
		CN 1672162 A	21-09-2005
		EP 1542149 A1	15-06-2005
		JP WO2004012128 A1	24-11-2005
		KR 20050026016 A	14-03-2005
		US 2006078883 A1	13-04-2006
		WO 2004012128 A1	05-02-2004