Chapter

How Morphology of the Human Pluripotent Stem Cells Determines the Selection of the Best Clone

Vitaly Gursky, Olga Krasnova, Julia Sopova, Anastasia Kovaleva, Karina Kulakova, Olga Tikhonova and Irina Neganova

Abstract

The application of patient-specific human induced pluripotent stem cells (hiPSCs) has a great perspective for the development of personalized medicine. More than 10 hiPSCs clones can be obtained from one patient but not all of them are able to undergo directed differentiation with the same efficiency. Beside, some clones are even refractory to certain directions of differentiation. Therefore, the selection of the "best" or "true" hiPSC clone is very important, but this remains a challenge. Currently, this selection is based mostly on the clone's morphological characteristics. Earlier, using methods of mathematical analysis and deep machine learning, we showed the fundamental possibility for selecting the best clone with about 89% accuracy based on only two to three morphological features. In this chapter, we will expand on how the morphological characteristics of various hiPSCs clones, the so-called "morphological portrait," are reflected by their proteome. By reviewing previously published data and providing the new results, we will highlight which cytoskeletal proteins are responsible for the establishment of the "good" morphological phenotype. Finally, we will suggest further directions in this research area.

Keywords: hiPSCs, hESCs, machine learning, best clone, morphological phenotype, proteome, cytoskeleton

1. Introduction

High-quality clones of human pluripotent cells (hPSCs) are of great importance for research in both basic and translational medicine due to their capacity to differentiate to all cell types of the human body and unlimited self-renewal. Unfortunately, currently available reprogramming methods to generate human induced pluripotent stem cells (hiPSCs) are stochastic, and that causes the presence of a large percentage of partially reprogrammed cells and cells with a low level of pluripotency [1]. The purification of culture is an important requirement to obtain high-quality clones. Usually, this includes either gene expression profiling or evaluation of the

cellular morphology by visual inspection or image analysis. However, both of these approaches have limitations. Namely, gene expression profiling gives a direct readout of stemness and differentiation, but it is destructive to cells. At the same time, visual morphological analysis of cells or their images is a nondestructive method, but it is prone to errors and misinterpretation. This explains the urgent need for the development of the noninvasive evaluation of the pluripotent cell cultures, that is. able to link cell morphology to the level of pluripotency. In the first part of the present chapter, we will discuss the morphological features of hPSCs and methods for their automated evaluation.

Currently, there are publications in the literature on the employment of live-cell imaging analysis along with deep machine learning for the development of automated software for the recognition of the best clones. Certainly, these scientific works, discussed in the first part of the chapter, represent only the first attempts of computer image analysis application for the selection of the best clones or identification of cells that have not undergone complete reprogramming. One question remains to be resolved in this method: to what extend can we make general conclusions based on the data from few studies, even with a large number of samples?

More than 500 distinct human embryonic stem cell lines (hESCs) have been generated to date, but only less than 100 lines are available now for general research as fully characterized lines (NIH stem cell registry, https://grants.nih.gov/stem_cells/registry/current.htm). In addition, there are multiple patient-specific human induced pluripotent stem cells (hiPSCs) lines and the list of these lines continues to grow. By 2020, about 131 studies were classified as clinical trials involving human pluripotent stem cells (hPSCs, comprising both hiPSCs and hESCs) [2]. The analysis published by Deinsberger and colleagues [2] revealed that the number of clinical trials involving hiPSCs was substantially higher than the one involving hESCs (74.8% vs. 25.2%). However, when counting only interventional studies, it appears that the majority (73.3%) was done with the use of hESCs. Application of patient-specific hiPSCs helps to overcome both ethical and immunological issues but hESCs are still widely used in the field of translational and regenerative medicine, disease modeling, and drug screening. Importantly, both hPSC types are very similar in their morphological characteristics not being molecular equivalents [3, 4].

Regardless of the common morphological features of hiPSCs and hESCs, it is well-documented that major line-to-line morphological variability exists even in the same culture conditions and with the use of the same propagation technique [5]. This fact raises the question of whether there are common morphological features that distinguish a "good" hPSCs clone from a "bad" one. Finding the answer to this question is extremely important as the maintenance of hPSCs in culture is not only expensive but it is also very labor intensive. The development of an automated quality control protocol can improve the utility of the high-quality clinical-grade cells.

A noninvasive method of visual inspection of the morphological appearance remains the main criteria used to select the best hPSCs clone. However, until now, it was not clear which parameters of morphology are closely associated with the pluripotent state.

Recently, we analyzed morphological parameters of several hPSCs lines of various passages. We first extracted the parameters from phase-contrast images and constructed classification models of colonies by morphological phenotype [6], and then we used image analysis with convolutional neural networks (CNNs) [7]. Further to this, expression analysis of 11 pluripotency markers genes allowed us to identify phenotype-specific sets of genes that could be used for the selection of the

best clones, meaning the fundamental possibility of constructing a morphological "portrait" of a colony informative for its automatic identification. Additionally, we performed a proteomic analysis of several hPSCs samples from various lines used before for the computational analysis and showed that cells with different phenotypes from various lines cluster at the proteome level in accordance with their morphological phenotype [7].

Multiple studies have provided datasets of comparative proteomes of various hPSC lines. Several studies used proteomic approaches to find proteins that regulate pluripotency [8–10] or to conduct a comparative proteomic analysis of supportive and unsupportive matrix substrates for hESCs maintenance [11]. In addition, several papers described quantitative proteomic analysis of hESCs differentiation [12, 13]. However, only in 2019, the first paper appeared on the analysis and comparison of the proteomic landscapes of 20 hiPSCs lines classified as stable and unstable based on colony morphology. This study has shown that different morphological "portraits" of colonies are associated with different proteomic profiles and different competencies for directed differentiation [14]. Furthermore, it has been shown that a direct relationship exists between pluripotent markers (DNMT3B, DPPA4, SALL4, CD9) and morphological "portraits" of various lineages [6, 14].

In this chapter, we will review the current knowledge about how automated evaluation of the morphological portrait is used to control the hPSC phenotype, and how it is connected to the proteomic analysis. Next, we will present our own proteomic data analysis of hPSCs in respect to their morphological phenotype. We will pay special attention to the cytoskeleton proteins, as some of them turned out to be the top candidates in determining the best cell and colony morphology. The future will tell us if the hiPSC technology will ultimately overcome the current challenges and will finally make its way into routine clinical application with the help of automated recognition of the best clone based on the morphological selection.

2. Morphological features of human pluripotent stem cells and methods for their automated evaluation

Currently, work with hPSCs begins with the assessment of their morphology by an expert to determine if there are signs of spontaneous differentiation or other unwanted changes. Established standard criteria for morphological features of hPSCs during their expansion can be described as: (a) a high nucleus/cytoplasm ratio, (b) prominent nucleoli, (c) formation of compact and round colonies with flat and densely packed cells with scant cytoplasm. Additional important marker is the presence of a clear and smooth colony edge [6, 15, 16]. As hPSC colonies propagate in culture, cells might spontaneously deviate from pluripotency toward a differentiated state. In that case, cell morphology changes dramatically, and it is very noticeable; the cells in the colony start to distribute sparser, the distance between the cells expands and cells significantly increase in size, undergoing a characteristic shape change [6]. In addition, undifferentiated hPSCs have more relaxed chromatin than differentiated; during the differentiation process, nucleoli become unclear and invisible under phase contrast microscopy [16]. Notably, only a very skilled expert can notice this alteration; therefore, evaluation of the cultures by the observation of the colonies morphology by an expert obviously depends on the expert's skills. Undoubtedly, the safe application of hPSCs in the clinic requires the creation of a cell evaluation method, which would be less dependent of the expert's skills.

In recent years, several image analysis approaches have been developed. Machine learning, which involves pattern recognition and computational learning, is one of the most widely used strategies. In addition to pattern recognition, some of machine-learning algorithms classify cells into several quality classes, which are related to non-morphological image features, such as the distribution of luminance intensity. The fully automated system has been reported for morphology-based evaluation of iPSC cultures that consists of time-lapse microscopy and image analysis software [17, 18]. The system acquires low-light phase-contrast images of iPSC growth collected during a period of several days, measures geometrical- and texture-based features of the colonies throughout time, and derives a set of six biologically relevant features to automatically rank the quality of the cell culture. This method has shown that hiPSCs that are classified visually could be adequately distinguished with local binary patterns and an intensity histogram [18]. The classifier presented in that work successfully identifies different cell stages for a wide range of scenes that can include different-sized colonies, varying amounts of dead cells and debris, and differentiated cells within colonies [18].

As mentioned before, in case of cell differentiation, nuclear structures reconfigure dynamically. The method published by Tokunaga and colleagues [19] for discrimination of the bona fide hiPSCs from non-reprogrammed ones, is based only on the fine differences of the nuclear morphology between cells. Namely, this work has demonstrated that specific quantitative parameters contributing to morphological discrepancies reside in the nuclear sub-domains. Analysis of nuclear morphologies revealed dynamic and characteristic signatures, including the linear form of the promyelocytic leukemia (PML)-defined structure in hiPSCs, which was reversed to a regular sphere upon differentiation. Thus, this data confirmed that hiPSCs have a markedly different overall nuclear architecture that may contribute to highly accurate discrimination based on the cell reprogramming status [19].

Similarly, the paper by Kato et al. [20] demonstrated a noninvasive image-based evaluation method for detecting partially differentiated colony morphology in heterogeneous colony populations *via* live image analysis. The authors analyzed eight major parameters comprising 27 sub-parameters selected as essential for further analysis of 303 hiPSC colonies. The data showed a relationship between image features and gene expression by analyzing the expression of hiPSC colonies classified by using spatial frequency. Next, colony morphology classification based on the statistical analysis of the live-cell images with unbiased morphological parameters was compared with classification based on global gene expression profiles of individual colonies. Classification utilizing statistical analysis produced similar results as compared to classification based on gene expression profiles. Thus, authors concluded that quantitative morphological evaluation method facilitates the noninvasive analysis of hiPSC conditions and demonstrates its utility in recognition hPSCs heterogeneity.

The paper by Wakui and colleagues [17] aimed at establishing quality classification of hiPSC images into three classes (poor, moderate, and good) by evaluating the biological features used in the visual inspection. Three features associated with biological structures such as the number of nucleoli, the crack area rate, and the differentiating cellular nuclei area rate were chosen by the expert. Importantly, these features were effective for quality evaluation by the visual inspection. As mentioned before, the number of nucleoli is a feature indicating a non-differentiated state, and cells with many nucleoli are considered to be of good quality. In contrast, the crack area rate and the differentiating cellular nuclei area rate are indicators of deviation from

pluripotency. The method identifies three feature detectors and the cell quality classifier, the inputs of which are the outputs of the detectors. Then, in the image analysis method, the feature detectors and the classifier are applied to each of the regions of interest (150 pixels, 50 μm) of a phase contrast image. For machine learning of the nucleolus detectors, the nucleoli dataset was used as training data. The crack detector and the differentiating cellular nuclei detector were tuned with the masked dataset. The cell quality classifier was developed with the labeled dataset. Nucleoli observed in undifferentiated cells are nearly oval-shaped, 3- to 6-μm in diameter, and appeared black under phase-contrast observation. For confirming the classification capability of these three features, the distributions of the features for each cell quality class of a respective cell line were investigated and the accuracy for cell quality classification that was equivalent to visual inspection with respect to the three hiPSC lines was confirmed [17].

Interestingly, the paper by Nishimura and colleagues remains the only paper, which is based upon the morphology of a cellular organelle; it describes the use of the mitochondria distribution and state for distinguishing reprogrammed mouse PSCs [21]. The authors reported the development of an imaging system, termed phase distribution (PD) imaging system, which visualizes subcellular structures quantitatively in unstained and unlabeled cells. The PD image and its derived PD index reflected the mitochondrial content, enabling quantitative evaluation of the degrees of somatic cell reprogramming and mouse PSCs differentiation [21]. The dynamic changes in mitochondrial biogenesis and antioxidant enzymes are well-documented during the spontaneous differentiation of hESCs, as well as during the reprogramming process [22]. Unlike in PSCs, in the somatic cells mitochondria are numerous and large, reflecting their dependence on oxidative phosphorylation for efficient energy production. It is well-known that the reprogramming of the somatic cells into iPSCs is accompanied by a metabolic shift from oxidative phosphorylation to glycolysis, concomitant with changes in structure and function of mitochondria [23, 24]. Indeed, iPSCs that are reprogrammed to different degrees show an inverse relationship between their pluripotency and mitochondrial activity [25]. Thus, morphological changes of subcellular structures such as mitochondria may serve as an additional useful marker to evaluate the pluripotency of reprogrammed cells.

Our own data on morphological parameters from three lines (hESC line H9, hiPSC line AD3, and hiPSC line HPCASRi002-A) revealed that several morphological criteria can be used to distinguish between "good" and "bad" phenotypes (**Figure 1A**) [6], thus demonstrating that these are strong and reliable criteria for determining the phenotype of hPSCs. We tested seven morphological parameters in total as possible predictors in the neural network-based classification models of the colony phenotype. The models aimed to predict the probability of the colony phenotype (either 'good' or 'bad') and were trained on the morphological parameter values of colonies or cells. A minimal model was selected for each data type that contained a minimal number of predictors and still provided the prediction accuracy close to that in the model with all predictors included. For the colony morphology data, we found a minimal model of four input parameters (Perimeter, Minor Axis, Shape Factor, and AIS) that showed 74% accuracy on average, while only two parameters (Perimeter and Shape Factor) were enough to provide a 68% average accuracy in the minimal classification model for the cellular morphological data (**Figure 1A**).

As an alternative approach for the colony phenotype prediction, we applied convolutional neural networks (CNNs) directly to the phase-contrast images of colonies, omitting the intermediate step of extracting the morphological parameters from the

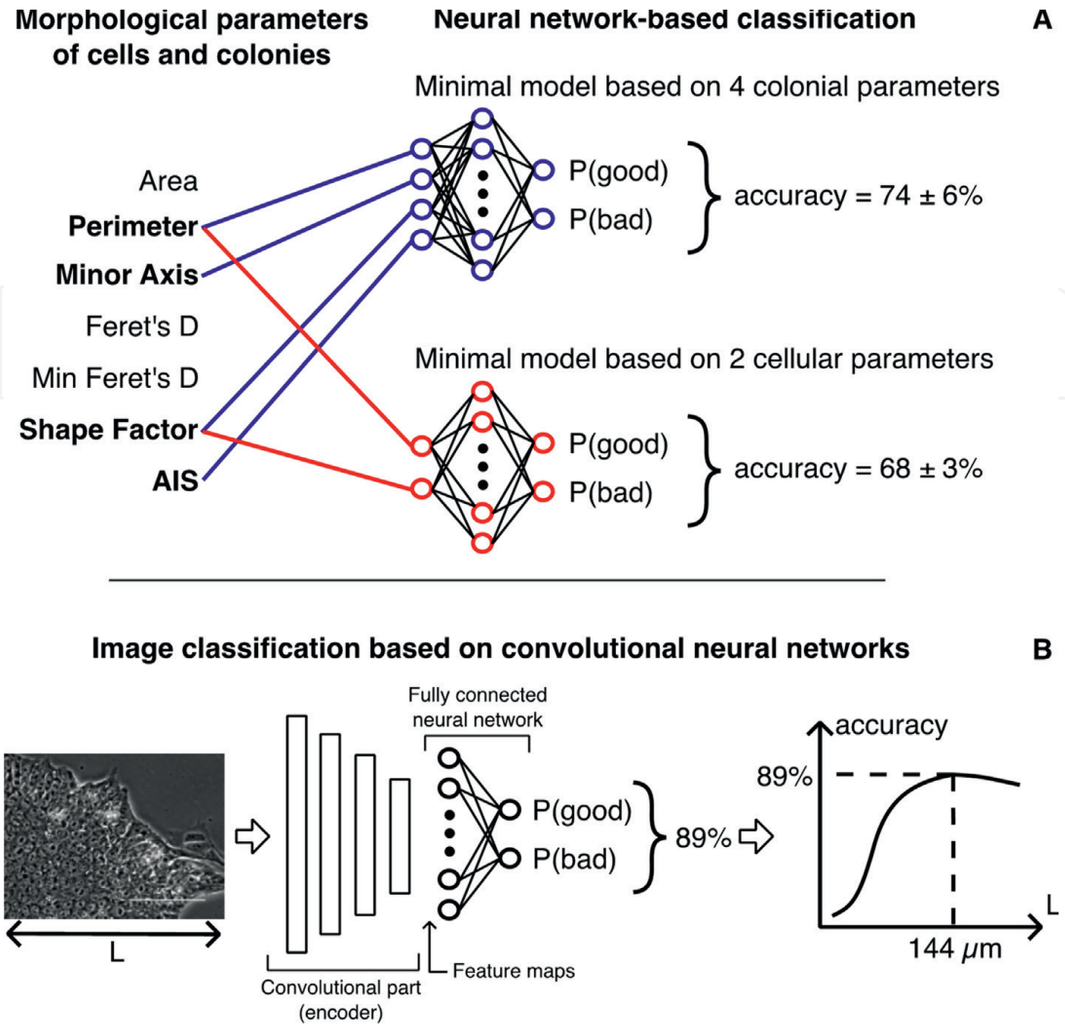

Figure 1.
Two approaches to colony phenotype prediction using automatic classification [6, 7]. A: Minimal classification models based on neural networks use four morphological parameters of colonies or two morphological parameters of cells as predictors. The corresponding parameters are shown in bold and connected with the classifier input with blue or red lines for the colonial or cellular data-based models, respectively. The output of the models is the probabilities that the colony will have a good or bad phenotype (P(good) or P(bad), respectively). B: CNN trained directly on colony images. The convolutional part of the CNN extracts the most representative feature maps from the images. These features are then passed to the input of a fully connected network trained to separate the phenotypes. CNNs trained on datasets with images of various linear size L show the prediction accuracy that has a maximum at the size L ~ 144 μm, which can be interpreted as the most informative spatial scale. AIS, area of intercellular space (for colonies only).

images (**Figure 1B**). CNNs extract the informative features, called "feature maps," from the images and use these features to predict the phenotype at the output. We trained the CNN-based classification model on phase-contrast images of the H9 hESC line colonies and obtained an 89% accuracy in phenotype prediction [7].

The morphological "portrait" of the colony that can be associated with clonality is a complex trait, with no clear spatial scale that could unambiguously separate the natural morphological variability within the colony from signs of clonality loss. Trying to answer the practical question about the most informative spatial scale at which the colony phenotype could be recognized by the automated classifiers most effectively, we trained CNNs on multiple datasets containing images of various linear size. We found an optimal image size of ~144 μm providing the best classification

Morphological data	Phenotypes	Classification data	Performance	Other details	Source
Six morphological parameters extracted from the analysis of hiPSC colonies generated from human fibroblasts using two different reprogramming strategies	12 colony types based on morphological and gene expression patterns, across which non-ESC-like colonies were contrasted against iPSC colonies	-	Criteria to determine iPSC colonies were formulated in terms of specific values (mean ± s.d.) of morphological parameters and demand for cells to be positive for endogenous Sox2 and Cdx2	First evidence that the expression of either Oct3/4 or Nanog is not appropriate for the identification of iPSCs, but rather the expression of endogenous Sox2 and Cdx2 is a reliable marker of iPSCs	Wakao et al. [27]
Six geometric- and texture-based features of iPSC colonies extracted from time-lapse videos of 12 clones from four iPSC lines	Good, fair, and poor	Classification tree based on the probability distribution of six features per phenotype value	Accuracy = 0.80–0.89	The degree of cell compaction and the doubling time were shown to be the features with the highest predictive power	Maddah et al. [18]
Automatically extracted features from phase contrast images for iPSC lines (201B7 and 253G1), newly generated iPSC lines (1H–4H), non-iPSC lines (15B2 and 2B7), and somatic cells (human mammary epithelial cells, HMECs)	Properly and improperly reprogrammed cells (iPSCs vs. non-iPSCs)	Supervised machine learning algorithm wndchrm (weighted neighbor distances using a compound hierarchy of algorithms representing morphology)	Accuracy = 0.87–0.96	Specific morphology quantification suggested that signals contributing to morphological discrepancies reside in nuclear sub-domains	Tokunaga et al. [19]
128-dimensional vector of features extracted by the scaled invariant feature transformation (SIFT) applied to phase contrast images of hiPSCs.	Good, semigood, and bad	*k*-nearest neighbors, multiclass support vector machines, and other classification methods	Accuracy up to 0.62	*k*-nearest neighbors method showed the best performance	Joutsijoki et al. [28]

Morphological data	Phenotypes	Classification data	Performance	Other details	Source
120 colony morphology parameters (reduced to 27 parameters) extracted from live-cell phase contrast images of aberrant 201B7-1A subclone (with unusual undifferentiated ESC- like colony morphology), its healthy parent 201B7 line (with typical ESC-like colony morphology), 253G1 cell line and its subclone 253G1-B1	No phenotyping. Colonies were clustered into five major clusters ('morphological categories') according to their morphological parameters	Unsupervised hierarchical clusterization of colony morphological parameters	-	Cluster analysis results for colony morphologies were reproduced by individual gene expression profiles	Kato et al. [20]
3 colony features (number of nucleoli, the crack area rate, and the differentiating cellular nuclei area rate) extracted from phase contrast images of MRC5, Edom, and 201B7 hiPSC cell lines	Good, moderate, and poor	Supporting vector machine	Accuracy = 0.86	Image analysis framework was created for automated extraction of the three features and classifier application to regions of interest for a phase contrast image	Wakui et al. [17]
Full data comprise low-magnification phase-contrast images and a fluorescence channel for alkaline phosphatase staining of mouse ESC colonies. The only morphological parameter used for phenotype classification is colony circularity.	Pluripotent colonies, mixed colonies, and differentiated cells	Pluri-IQ software, which can automatically quantify the percentage of pluripotent, mixed, or differentiated cells through culture images, with the following cascade modules: segmentation, machine learning (random forest classifier), validation, and automatic scoring	Accuracy > 0.90	Pluri-IQ uses as input large images and has advantages compared with other similar software	Perestrelo et al. [29]

Morphological data	Phenotypes	Classification data	Performance	Other details	Source
Authors derived a phase distribution (PD) index and a 3D extracted PD (ePD) image from an improved retardation modulated differential interference contrast (RM-DIC) imaging systems, applied to mouse iPSCs. Both the PD index and ePD image were designed to reflect subcellular structures, specifically the mitochondrial content.	Qualitative phenotypic states of different degrees of somatic cell reprogramming, ESC differentiation, and pluripotency	-	Various quantitative measures for correspondence between the PD index or ePD image with various phenotypic values	The PD index and ePD image were shown to reflect different degrees of somatic cell reprogramming, ESC differentiation, and pluripotency	Nishimura et al. [21]
Automatic feature maps extraction as a part of classification model training for phase contrast images of diseased iPSCs expressing the Huntington's disease phenotype	Four morphological classes: debris, dense, differentiated, and spread	CNN, with data augmentation by supplementing a minimal biological dataset *via* image generation using generative adversarial networks. Model based on Markov chain stochastic processes is used to account the influence of temporally constrained differentiation on classification model training.	Recall (true positive rate) = 0.88–0.94	Presented work highlights the importance of exploiting temporal relationships between image classes, which is an example of using 'domain knowledge' in combination with deep learning	Witmer and Bhanu [30]
Seven morphological parameters of colonies and cells from phase-contrast images for H9 hESC line, AD3 hiPSC line, and HPCASRi002-A hiPSC line	Good and bad	Neural networks	Accuracy = 0.74 ± 0.06 for colonial parameters and 0.68 ± 0.03 cellular parameters	Four parameters for colony morphology and two parameters for cellular morphology are enough	Krasnova et al. [6]

Morphological data	Phenotypes	Classification data	Performance	Other details	Source
Feature maps extracted from phase-contrast images of iPSC clones derived from peripheral blood mononuclear cells	Undifferentiation, cracked, built-up, differentiation	Feature maps extraction by VQ-VAE encoder (unsupervised learning) + support vector machine for phenotype classification based on extracted feature maps	Accuracy = 0.89	Support vector regression method was used to predict the expression of 218 genes based on feature maps extracted from images, with $0.3 < R^2 < 0.69$	Wakui et al. [31]
Feature maps are automatically extracted from phase contrast images for H9 hESC line during model training	Good and bad	CNN	Accuracy = 0.89	Most informative spatial scale was determined	Mamaeva et al. [7]

Table 1.
Selected results of phenotype classification based on morphological features of hPSCs colonies.

performance (**Figure 1B**). This size is intermediate between typical colony and cell sizes, reflecting the fact that both cellular and colonial information should be taken into account.

We linked these results to the transcription studies by measuring the expression of 10 pluripotency markers (*DNMTB3, SALL4, IGFR1, CD9, DPPA4, OCT4, REX1, NANOG, SOX2,* and *KLF4* genes) in colonies from the hESCs (H9) and hiPSC (AD3 and HPCASRi002-A) lines with different phenotypes [6]. We found that the *SALL4*, *DNMTB3*, *REX1*, *DPPA4*, and *SOX2* genes demonstrated differentiated expression in colonies of different phenotypes, thus confirming that the phenotypes did represent the pluripotency status of the colony.

Finally, the question can be asked, whether the cultivation conditions, namely various culture media and matrixes, affect the morphological parameters important for morphological phenotype recognition. By other worlds, is evaluation method based on the morphologies of various hPSC lines applicable under different culture conditions? The paper by Harkness and colleagues [26] addresses the effect of five different media, namely, mTESR1, Essential E8, StemPro (SP), mouse embryonic fibroblasts conditional media (CM) and StemMacs iPS-Brew XF (SM), on the morphological parameters of the three established hESCs lines (MEL1, WA09, ESI-hES3). These lines were routinely grown on Matrigel (Corning) in mTESR1 media before switching to a different media. As a result, the authors observed distinct and measurable differences in nuclear and cell morphology between different culture conditions. In CM and SP cultures, authors noticed a looser colony structure and a flatter appearance when compared to mTESR1, E8, or SM media. The morphological parameters such as nuclear area, cell area, cell roundness, and cell spread in all three lines demonstrated an overall decrease, while in the least defined media, CM, in all cell lines the cells became larger. Moreover, the nuclear/cytoplasmic ratio varied between the lines, suggesting that media composition can affect the cell's parameters and may cause cytoskeletal remodeling. Furthermore, high content imaging demonstrated that hESCs grown in different media exhibit significantly different cytoskeletal architecture while maintaining their pluripotent status, suggesting that cytoskeleton has become more stable in xeno-free media [26]. Thus, the detailed analysis provided in this research let to conclude that morphological alterations of cell phenotype can be associated with the changes of cell culture conditions. However, it can be assumed that when changing from culture medium to another, cells undergo a period of adaptation and, perhaps, after a certain number of passages, they will restore their previous morphological parameters. However, this needs further verification.

Thereby, to create a reliable system for recognizing the best clones, further studies of different hPSC lines during their cultivation on various matrixes and media are required. The creation of a single database that combines data on morphological parameters from numerous lines will improve the methods of automatic clone's recognition for their reliable application in clinic.

In **Table 1**, we summarized some findings about phenotype classification based on morphological features of hPSCs colonies.

3. Morphological phenotypes of the hPSCs reflected in different proteomic landscapes

Proteomics analysis provides an excellent tool for large-scale quantification and benchmarking of cells and an opportunity to understand deeper the rules that govern hPSCs morphology. Compared to other ~omics, such as transcriptomics and genomics

approaches, proteomics analysis measures the translated proteins. Most of the previous studies have used proteomics approaches to identify proteins important for stem cell pluripotency maintenance and for lineage differentiation [8–10, 13]. Some studies have explored the membrane proteins [32] or the hESCs phosphoproteome [8, 9] in comparison to hiPSCs proteome and phosphoproteome [33]. In addition, molecular differences of the proteome level between hiPSCs of different somatic origin were described [10]. A comparative proteomic analysis has been published comparing supportive and unsupportive extracellular matrix substrates used for hESCs maintenance [11]. All these studies revealed a huge number of proteins known to be important for hPSCs maintenance, namely, cell cycle and DNA damage repair proteins, proteins involved in integrin binding, intracellular vesicle trafficking proteins, RNA binding, adaptor proteins and histones, proteins of exosomes biogenesis and tumorigenesis, zinc finger proteins, mitochondrial proteins, and many others.

The goal of our study was to analyze the hPSCs proteome in accordance with the selected morphological phenotypes [6, 7]. Thus, we compared a proteomic "portrait" of the "true" or the best hPSC colonies versus the "bad" ones.

In the hESCs (H9) samples, we have identified in total 1791 proteins in a good agreement with the Van Hoof and colleagues [9] who have identified 1775 proteins from undifferentiated hES cell line HES-2. Our data demonstrated a clear separation of the samples in accordance with their morphological phenotypes [7], in agreement with the previously published data of Bjørlykke and colleagues [14], thus emphasizing that good and bad morphological populations are molecularly distinct. Comparative proteome analysis of the hESC (H9) colonies with the good morphological portrait compared to colonies with poor morphology and signs of spontaneous differentiation showed that 63 proteins are downregulated and 25 proteins are upregulated (**Figure 2**) [7].

In the context of the identified proteins that determine the morphology of hPSCs, we were especially interested in cytoskeletal proteins since they form the structural network of the cell. In addition, the migration and spread of motile cells, such as hPSCs, over the surface of the substrate accompanied by the reorganization of their actin network. Among 25 upregulated proteins, four belong to cytoskeletal proteins (MYH7, RDX, CNN3, and AIF1L). The other one is the tight junction protein 1, or ZO1 (TJP1), one of the functions of which is to organize the components of tight intermediate junctions and bind them to the cortical actin cytoskeleton. In our analysis, MYH7 appears on the top position among the upregulated proteins (**Figure 2**). The MYH7 gene, known as myosin beta heavy chain (MHC-β), is classified as a type I fiber. Myosins are a large family of proteins that share the common features of ATP hydrolysis, actin binding, and potential for kinetic energy transduction. They composed of a pair of myosin heavy chains (MYH) and two pairs of nonidentical light chains. At least 10 different MYH isoforms have been described in mammalian cells, and the role for the identified in hESCs MYHs, such as MYH16, MYH15, MYH10, MYH9, and MYH7, is aviating to be discovered. This protein is a critical component of the sarcomere's structure and interacts with other key cytoskeletal proteins such as actin, troponin, and myosin-binding protein C (MYBPC3). Its role was shown in directed differentiation of hiPSCs into cardiomyocytes [34] but has not been studied in hPSCs. The dynamics of actin-myosin contraction are directly regulated by the amount of alpha-actinin-3 (ACTN3), which forms cross-links, and the absence of ACTN3 disrupts the symmetry of the actin network in cells. Human ESCs exhibit basal-apical polarity, junctional complexes, integrin-dependent matrix adhesion, and E-cadherin-dependent adhesion, all of which are characteristics of the epiblast epithelium of a mammalian embryo.

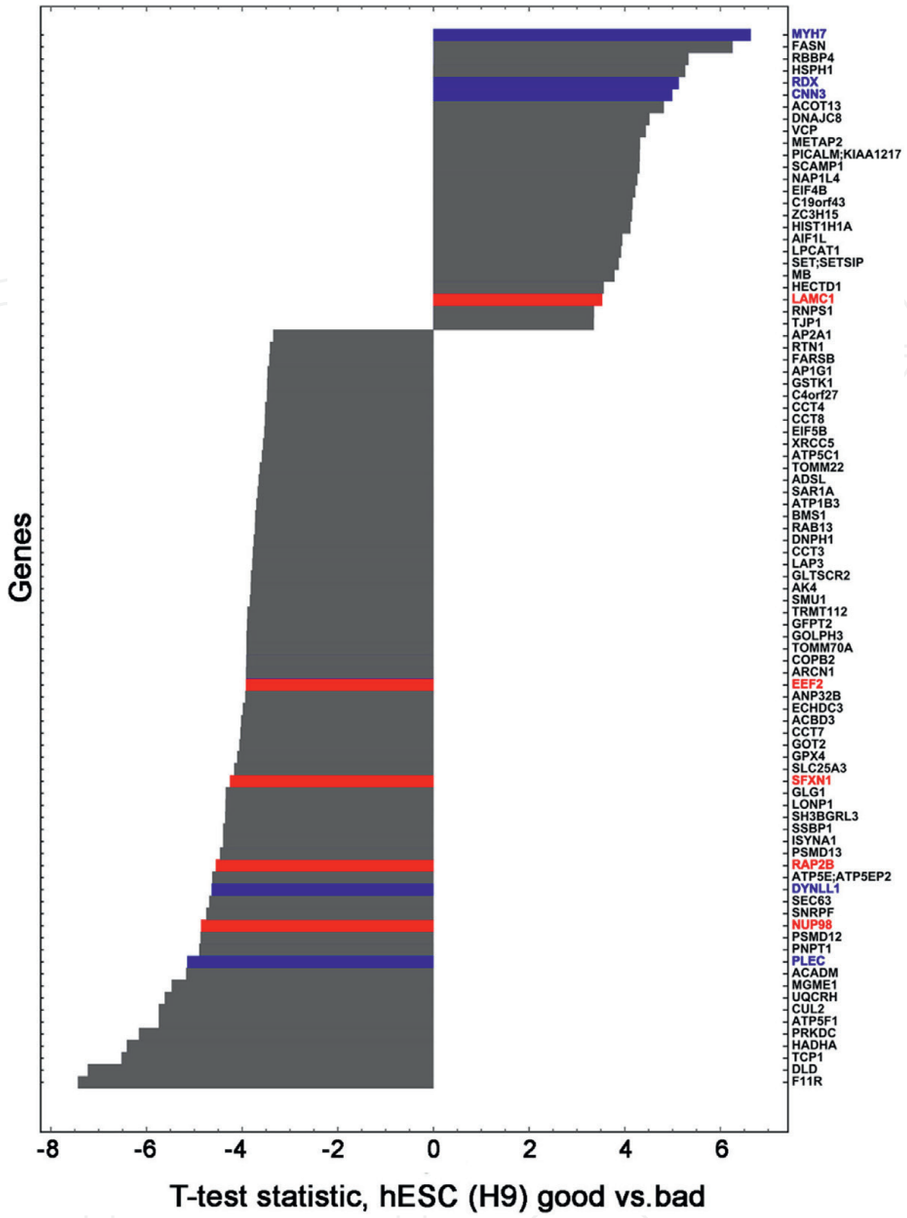

Figure 2.
Z-score-ranked distribution plot for the proteins of the hESC (H9) colonies with the "good" morphological portrait compared to colonies with "bad" morphology. Cytoskeletal proteins are marked blue, and proteins identified via comparison of the "good" morphological hESC H9 samples versus two hiPSC lines with the same characteristics are marked red.

When hPSCs are subject to enzymatic digestion during propagation of the colonies, epithelial structures are destroyed, which leads to programmed cell death; here, actin-myosin contraction is a critical effector of the cell death response to enzymatic dissociation [35]. With this regard, inhibition of the myosin heavy chain ATPase, inhibition of the myosin heavy chain, and inhibition of the myosin light chain (MLC) have been shown to increase the survival and cloning efficiency of individual hPSCs [36]. ROCK inhibition decreases phosphorylation of MLC, suggesting that inhibition of actin-myosin contraction is also the mechanism through which ROCK inhibitors increase cloning efficiency of hESCs [37]. In addition, ROCK1/ROCK2 silencing demonstrated that ROCKs regulate MYH function through MLC phosphorylation in hESCs, which, in turn, leads to membrane blebbing and cell death [36]. Lastly, MYH9 and MYH10

are the most highly expressed MYHs with the conserved sites in hESCs. Treatment of hESCs with MYH9/MYH10 siRNAs demonstrated severe phenotypic changes after 96 hours of transfection but increased cell attachment, survival, and cloning efficiency [36]. On the other side, as mentioned above, MYH7 is regarded as a mesenchymal and specifically myocardial marker gene [38]. Our data is also in a good agreement with the earlier work, which demonstrated that a high level of MYH7 protein was detected in hESCs but not in hiPSCs, while MYH9 was identified in both cell types [12]. Obviously, the role of MYH7 needs to be elucidated further. We were able to detect a very thin network of MYH7 colocalized with F-actin fibers but observed its destruction in hESCs (H9) clones with bad morphology (**Figure 3**); supporting the data from

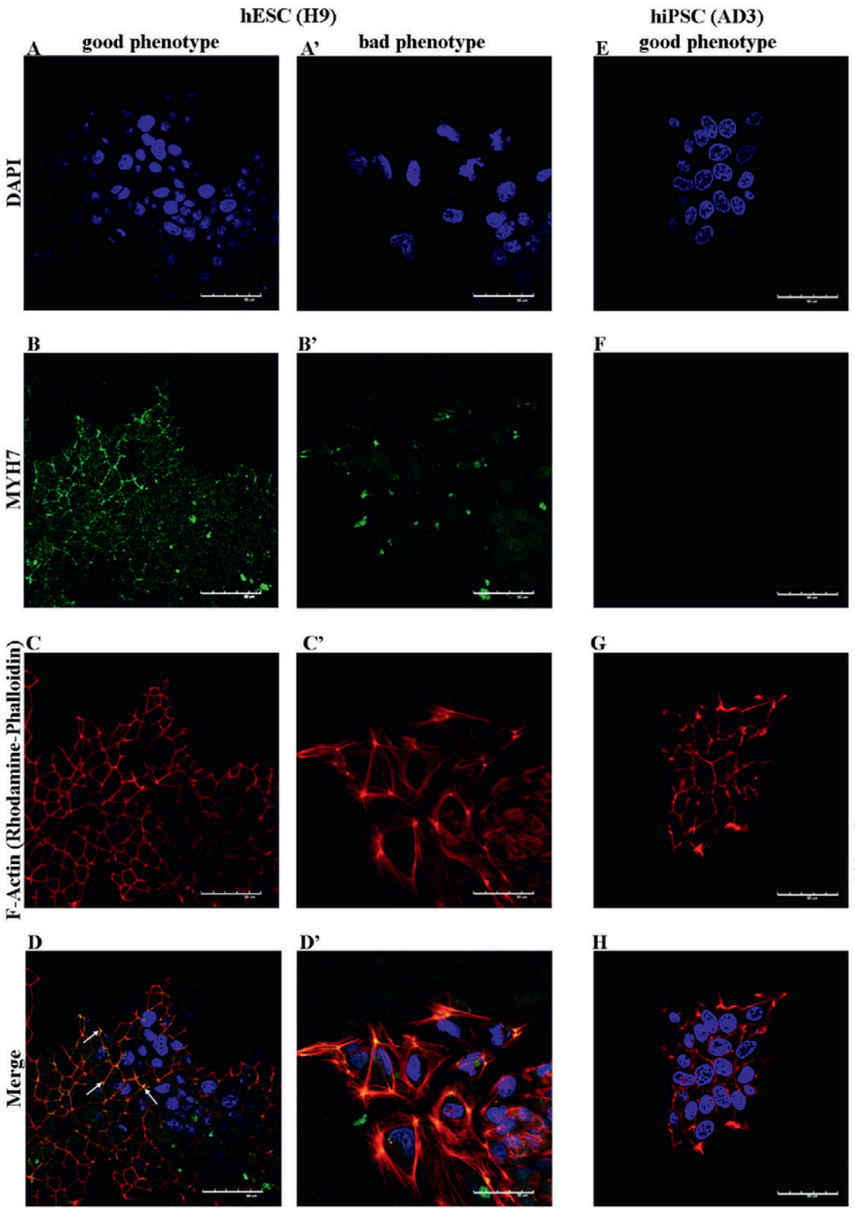

Figure 3.
Colocalization of the MYH7 with F-actin in the organized cytoskeleton network in colonies with "good" phenotype and complete loss of such structural organization and colocalization in "bad" hESCs. Thin white arrows are pointing on the colocalization of F-actin and MYH7. Immunofluorescence was performed using anti-MYH7 antibodies (Santa Cruz Biotechnology, sc-53,089), secondary anti-mouse IgG tagged with Alexa Fluor 488 (Abcam, ab150113), and Rhodamine-Phalloidin (Invitrogen, R415). Scale bar 50 μm.

proteome study, we did not detect MYH7 staining in hiPSCs. The emergence of MYH7 as a top candidate to support the best hESCs morphology might reflect the complexity of the hESCs proteome.

Radixin (RDX) was ranked fourth in our list (**Figure 2**). Radixin is a cytoskeletal protein that may play an important role in binding actin to the plasma membrane. Its exact role for hPSCs has not been explored. However, cellular functions, such as migration and adhesion, require a highly dynamic cytoskeleton behavior. Linker proteins of the ERM family (ezrin/radixin/moesin) can interact with both F-actin and several transmembrane proteins, providing a connection between extracellular cues and the cytoskeleton. The involvement of ERM proteins in a variety of cell functions in the embryonic and early postnatal brain, including axonal outgrowth, morphological rearrangement, cell migration, and signaling, have been described [39]. It is important to note that radixin has been shown to concentrate in the cleavage furrow of dividing cells and may have a role in proliferation [40, 41], the high speed of which is important for pluripotency maintenance [42].

Lastly, throughout the top represented cytoskeletal proteins in hESCs with a good morphological phenotype, we want to discuss CNN3, Calponin 3 (**Figure 2**). Calponin is an actin filament-associated regulatory protein expressed in smooth muscle and multiple types of non-muscle cells. It is capable of binding to actin, calmodulin, and tropomyosin. Three homologous genes, *CNN1*, *CNN2,* and *CNN3*, encoding calponin isoforms 1, 2, and 3, respectively, are present in vertebrates. All three Calponin isoforms are actin-binding proteins with functions in inhibiting actin-activated myosin ATPase and stabilizing the actin cytoskeleton, while each isoform executes different physiological roles based on their cell type-specific expressions. Calponin 1 (*CNN1*) is specifically expressed in smooth muscle cells and plays a role in smooth muscle contractility. Calponin 2 (*CNN2*) is expressed in both smooth muscle and non-muscle cells and regulates multiple actin cytoskeleton-based functions. Calponin 3 (*CNN3*) participates in actin cytoskeleton-based activities in embryonic development and myogenesis. Experiments with cytotrophoblasts from human placenta demonstrated that *CNN3* gene knockdown promoted actin cytoskeletal rearrangement, suggesting *CNN3* to be a negative regulator of trophoblast fusion [43]. With the course of trophoblastic cell differentiation, CNN3 undergoes downregulation. In the trophoblastic cells, membrane flexibility is necessary for membrane fusion [43]. However, whether *CNN3* expression affects the flexibility of the hPSCs plasma membrane is not known, but it may be suggested that regulation of actin cytoskeletal rearrangement by *CNN3* is required for hPSCs. Recently, Calponin 3 was studied in the U2OS osteosarcoma cells, where RNAi knockdown studies revealed that CNN3 is a dynamic component of stress fibers and is required for controlling proper contractility of the stress fiber network [44]. Importantly, the role for CNN3 was also shown for the maintenance of the lens epithelial phenotype where downregulation of CNN3 expression induced changes in cell shape, reorganization of actin cytoskeleton, and formation of focal adhesions resulting in activation of mechanosensitive transcription factor Yap in association with decreased E-cadherin and β-catenin expression [45]. Whether or not the high level of CNN3 in hESCs is associated with the focal adhesion and E-cadherin maintenance remains to be elucidated. Our immunofluorescence study supported obtained proteomic data and revealed a colocalization of the CNN3 with F-actin in the organized cytoskeleton network in colonies with good morphological appearance and complete loss of such structural organization and colocalization in "bad" hPSCs (**Figure 4**).

Among the top downregulated cytoskeletal proteins in hESCs with good morphology appeared DYNLL1 (Dynein light chain 1) and PLEC (PLECTIN) (**Figure 2**).

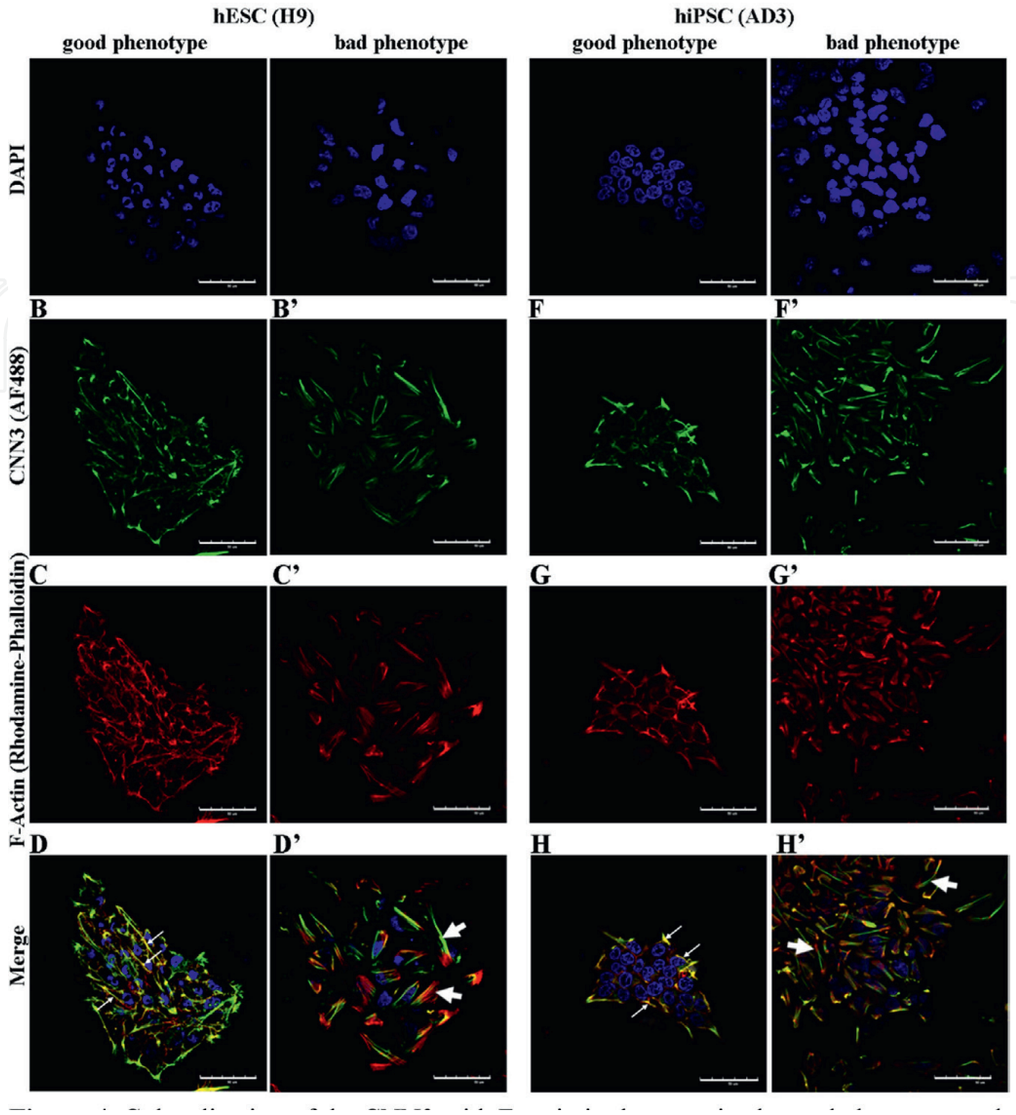

Figure 4.
Colocalization of the CNN3 with F-actin in the organized cytoskeleton network in colonies with "good" phenotype and complete loss of such structural organization and colocalization in "bad" hPSCs. Thin white arrows are pointing on colocalization of F-actin and CNN3; thick white arrows show the absence of colocalization. Immunofluorescence was performed using anti-CNN3 rabbit antibodies (ATLAS, HPA051237), secondary anti-rabbit IgG tagged with Alexa Fluor 488 (Abcam, ab150077), and Rhodamine-Phalloidin (Invitrogen, R415). Scale bar 50 μm.

Cytoplasmic DYNEIN1 acts as an engine for intracellular retrograde mobility of vesicles and organelles along microtubules. Plectin maintains tissue integrity and associate with intermediate filaments (IF). It acts as a cytoskeletal cross-linking agent and signaling scaffold, influencing both the mechanical and dynamic properties of the cytoskeleton. As a member of the cytolinker protein family, plectin has a multidomain structure that is responsible for its ability to bind to many cytoskeletal proteins. It binds not only to all types of IFs, actin filaments, and microtubules but also to transmembrane receptors, nuclear envelope components, and several kinases with known roles in cell migration, proliferation, and energy metabolism. The exact role of plectin in cytoskeletal dynamics is not studied for hPSCs, but in view of its downregulation for a good morphological phenotype, it can be assumed that lower level of protein expression may play a role in the cytoskeletal plasticity of these cells.

In **Figure 5A**, we demonstrate biological processes that we have identified to be related to the upregulated proteins in hESCs H9 line with good morphology. As can be seen, among the most important of them are cellular component biogenesis and assembly, organelle organization, epithelium development, cytoskeleton organization with DNA packaging, and chromatin organization. It is important to highlight that among the most important processes are up-regulation of actin filament-based processes and actin cytoskeleton organization. Among biological processes associated with downregulated proteins (**Figure 5B**), we identified cellular metabolic processes, nitrogen compound metabolic processes, cellular localization, protein transport, DNA metabolic processes, and many others related to control of the cellular metabolism. Also, cellular component analysis (**Figure 5C**) revealed cytoplasm, actin cytoskeleton,

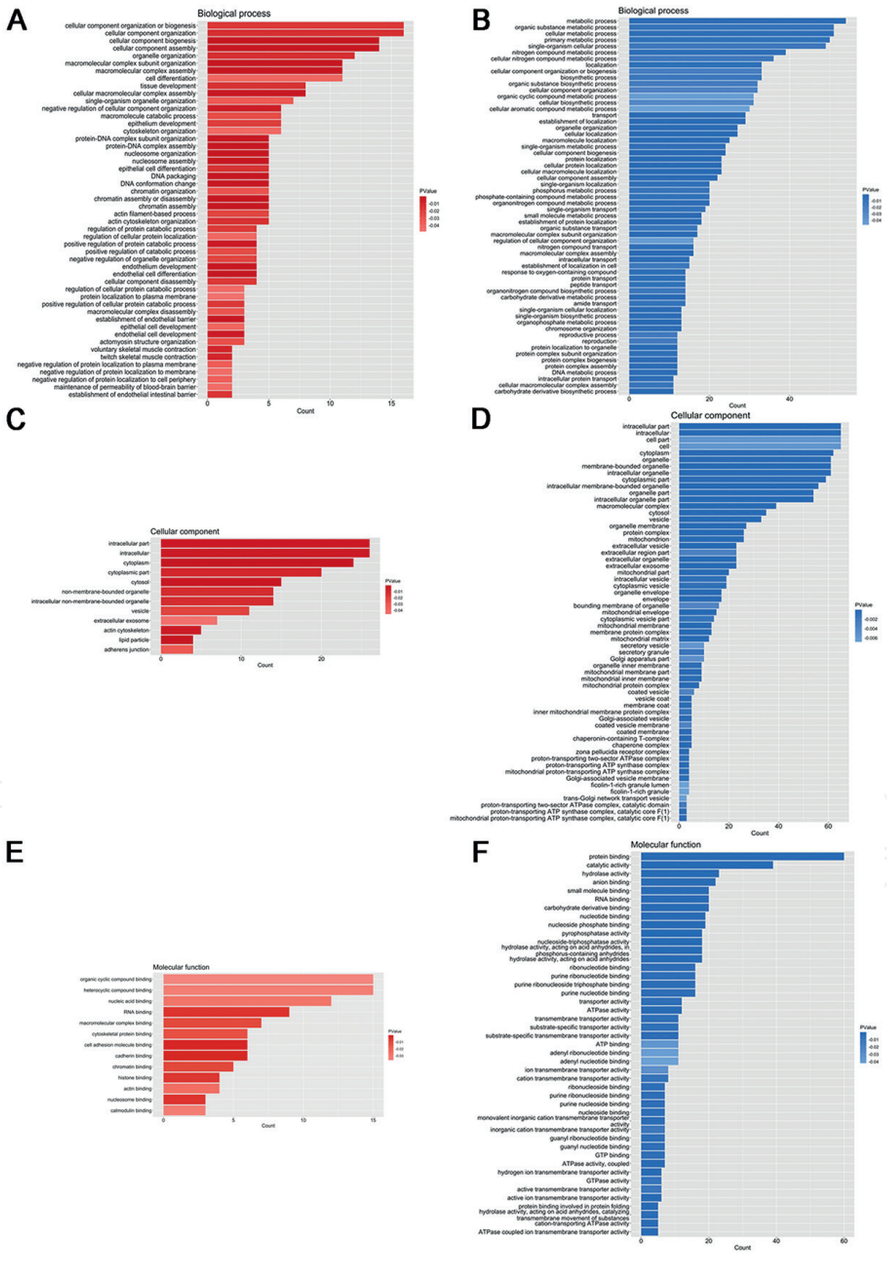

Figure 5.
Biological functions, cellular functions, and molecular processes associated with the up- and downregulated proteins in hESCs H9 cell line. Red bars (A, C, and E) refer to upregulated proteins; blue bars (B, D, and F) refer to downregulated proteins.

and adherence junction among most upregulated processes, while cellular processes related to cell, membrane-bounded organelle, vesicle, mitochondrion, mitochondrial, and organelle envelope appeared to be downregulated (**Figure 5D**). Importantly, molecular functions associated with "good" morphology include cytoskeletal protein binding, cell adhesion molecular binding, cadherin binding, actin binding together with chromatin and histone binding (**Figure 5E**), while among downregulated cellular functions appeared ribonucleotide, purine nucleotide binding, ATP binding, and GTP binding (**Figure 5E**).

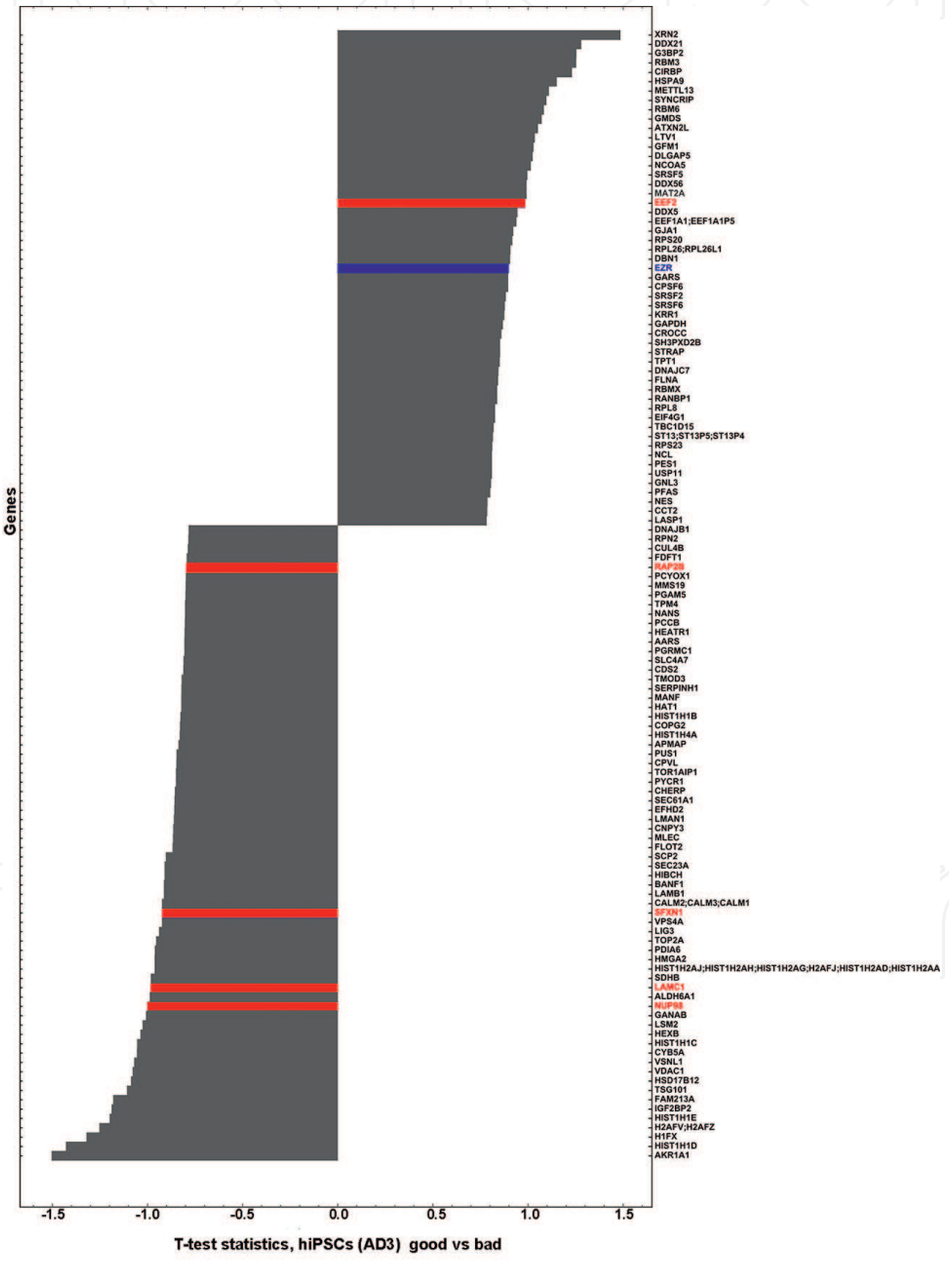

Figure 6.
Z-score-ranked distribution plot for the proteins of the hiPSC (AD3) colonies with the "good" morphological portrait compared to colonies with "bad" morphology. The EZRIN protein is marked blue, and proteins identified via comparison of the "good" morphological hESC H9 samples versus two hiPSC lines with the same characteristics are marked red.

In addition to hESC line H9, we analyzed by the same proteomic approach two hiPSC lines of different origins, namely obtained from the neonatal fibroblasts line AD3 and a patient-specific hiPSCs line HPCASRi002-A (CaSR) [6, 7]. Morphological evaluation of the "good" and "bad" hiPSC clones, as well as comparisons of their proteomic landscapes was performed as for hESC and H9 samples [6, 7]. Interestingly, in the same analysis of proteins associated with cytoskeletal function among experimental groups of the hiPSCs lines, EZR (EZRIN) turned out to be the top-upregulated protein (**Figure 6**). Ezrin, also known as cytovillin or villin-2, is a cytoplasmic peripheral membrane protein and functions as a substrate for tyrosine kinase in microvilli. Its significance for hPSCs morphology has not been studied. Earlier, in support of our data, EZRIN was demonstrated as one of the most prominent cytoskeletal proteins by proteomic profiling of hESCs at the first 48 hours of the early differentiation stage [12], suggesting that it may be expressed differently in clones with "good" and "bad"

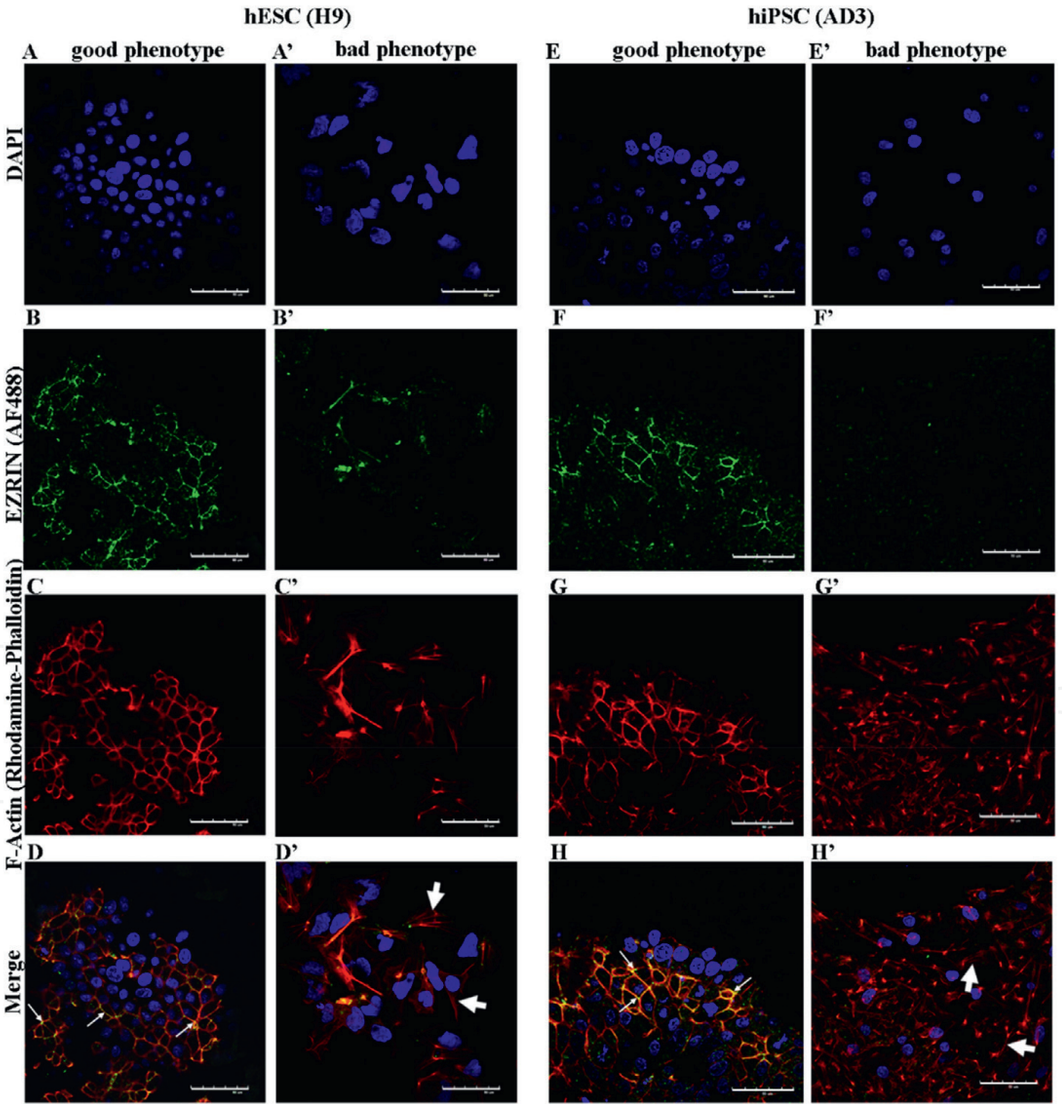

Figure 7.
Association of the EZRIN cytoskeletal network with F-actin in hESC and hiPSC clones with "good" morphological phenotype and complete distraction of this network in "bad" clones. Thin white arrows are pointing on colocalization of F-actin and EZRIN; thick white arrows show the absence of colocalization. Immunofluorescence was performed using anti-EZRIN antibody 3C12 (Invitrogen, 35–7300, secondary), secondary anti-mouse IgG tagged with Alexa Fluor 488 (Abcam, ab150113), and Rhodamine-Phalloidin (Invitrogen, R415). Scale bar 50 μm.

phenotype. In support of our proteome data, by employing specific anti-EZRIN antibody, we were able to detect the association of the EZRIN with F-actin in clones with good morphological phenotype and complete absence of such association and very weak pattern of staining in "bad" clones (**Figure 7**).

Interestingly, a very close in concept earlier work of Bjørlykke and colleagues [14] performed on 20 hiPSC lines different on their morphological appearance did not recognize the same cytoskeleton proteins among abundant upregulated or downregulated proteins. However, KERATIN 19 (KRT19), a member of the keratin family of the intermediate filament proteins responsible for the structural integrity of the epithelial cells was identified among upregulated ones as well as ADD2, a member of the cytoskeleton-associated proteins (ADDUCINS) that promotes the assembly of the spectrin–actin network [14]. Among abundant downregulated proteins PALLADIN (PALLD) and FIBRONECTIN1 (FN1) along with the MRC2, extracellular matrix remodeling protein, appeared as significantly downregulated [14]. Fibronectins bind cell surfaces and various compounds as collagen, fibrin, and actin. These proteins involved in cell adhesion and maintenance of the cell shape. Palladin as a cytoskeleton protein involved in the organization of the actin network, motility, and adhesion. Importantly, both named proteins have a role in cell morphology. Keeping in mind the importance of the colony-defined edge as meaningful morphological characteristic of a good hPSCs phenotype, one can recognize an importance of MRC2 for the establishment of the hPSCs phenotype as MRC2 is a member of the mannose receptor family proteins and plays a role in the extracellular matrix remodeling.

Importantly, it appears that only five proteins (SFXN1, LAMC1, RAP2B, NUP98, and EEF2) were identified *via* comparison of the "good" morphological hESC H9 samples versus two hiPSC lines with the same characteristics (**Figure 8**). Wherein, H9 samples contained 83 unique proteins and hiPSCs–116, but none of the identified proteins is a cytoskeletal protein.

Eukaryotic translation elongation factor 2 (EEF2), the GTP-binding translation elongation factor family member and an essential factor for protein synthesis appeared upregulated in hiPSCs while was downregulated in hESCs samples with good morphology (**Figures 2** and **6**). EEF2 is known as a positive regulator of apoptosis [46]. In a highly proliferative cells, EEF2 maintains genomic integrity by arresting the cell cycle at G2/M phase in response to ionizing radiation to prevent

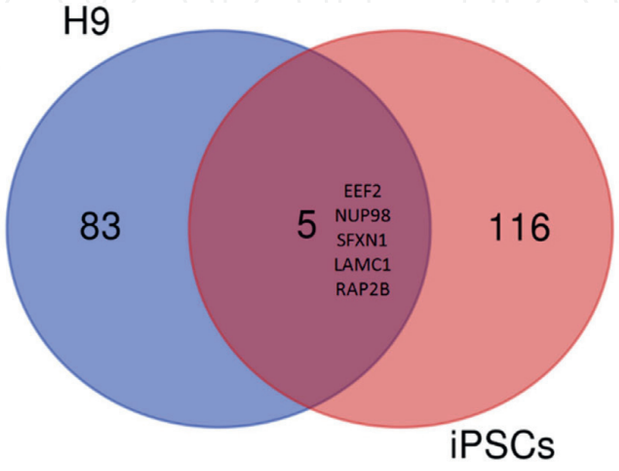

Figure 8.
Comparison of the "good" morphological hESC H9 samples versus two hiPSC lines with the same characteristics revealed five common proteins.

mitotic catastrophe [46]. The rapid proliferation of hPSCs is due to their unique cell cycle regulation. The interplay between cyclins, cyclin-dependent kinase (CDK), and cyclin-dependent kinase inhibitors is important for tight regulation of cell cycle progression in these cells [47]. Moreover, the cell cycle regulation is not only tightly related with pluripotency but also the cell cycle regulators have important functions in DNA damage response (DDR) [42]. Since maintaining genomic stability in hPSCs plays a pivotal role in their self-renewal and stemness, the role of EEF2 should be assessed in the nearest future as in terms of therapeutic application, genomic stability is the key to reducing the risks of cancer development due to abnormal cell replication.

Laminin subunit gamma 1 (LAMC1) belongs to Laminins, a family of extracellular matrix glycoproteins, which are the major non-collagenous constituent of basement membranes. Basement membranes are thin sheets of specialized extracellular matrix (ECM), underlying all epithelia and some other cell types. Laminins are important regulators of cellular functions such as cell adhesion, differentiation, migration, signaling, and metastasis. Human PSCs not only have characteristics typical for epithelial cells [48, 49] but they also rely upon ECM proteins for the support of their niche [50]. Human ESCs produce Laminin α1, α5, β1, and γ1 chains and deposit them as Laminin-511 into hESC-produced ECM. Importantly, Laminin-511 supports hESCs growth in defined medium equally well as Matrigel [50]. Indeed, LAMC1, as well as LAMB1, have been detected in the hESCs by proteomic analysis [11]. However, in our analysis, LAMC1 appears among downregulated proteins in hiPSCs with good phenotype in contrast to hESCs (**Figures 2** and **6**) regardless that all cell lines were grown on the same basement membrane matrix-Matrigel (Corning) with mTESR1 media [6]. As much as hiPSCs are not identical to hESCs, the identified differences may indicate the need for further research in the direction of the hPSCs niche supporting factors for better support of their *in vitro* maintenance.

SFXN1, RAP2B, and NUP98 are expressed in both hESCs and hiPSCs with "good" morphological phenotypes (**Figures 2** and **6**).

Sideroflexin 1(SFXN1) is an integral component of the mitochondrial inner membrane, and it is important for D-serine and L-serine transmembrane transporter activity.

Ras-related protein Rap-2b (RAP2B) is a member of the Ras family of small GTP-binding proteins, and it is involved in innate immune response and ERK signaling, both of which are important players during the reprogramming process. Also, RAP2B may play a role in cytoskeletal rearrangements and may regulate cell spreading through activation of the effector Traf2- and Nck-interacting kinase (TNIK) [51]. Moreover, RAP2B is expressed at high level in various human tumors, where its involvement in cellular spreading and migration was demonstrated more recently [52].

Nuclear pore complex protein Nup98 (NUP98) plays a role in the nuclear pore complex (NPC) assembly and/or maintenance. Nuclear pore complex (NPC) proteins are well-known for their critical roles in regulating nucleocytoplasmic traffic of macromolecules across the nuclear envelope. Several findings suggest that some nucleoporins, including Nup98, have additional functions in developmental gene regulation. Nup98 exhibits transcription-dependent mobility at the NPC but can also bind chromatin away from the nuclear envelope, and it is frequently involved in chromosomal translocations [53]. Importantly, acting as transcription factor, Nup98, could interact directly with histone-modifying enzymes CBP/p300 and histone deacetylases (HDACs), the role of which for hPSCs is well established. However, while the role of Nup98 as a multifunctional protein in macromolecular export has been

studied extensively [53], its precise role in hPSCs has not been elucidated and awaits further discovery.

Thus, we can conclude that despite the significant differences in the protein content for the two studied pools (hESCs vs. hiPSCs), when comparing cells of different types within the same experimental group of "good" morphological phenotype, only five differentially expressed proteins were found out of 1933 reliably identified proteins. This may indicate the similarity of the mechanisms that regulate the "good" morphology of hPSCs.

4. Conclusions and future perspectives

The development of reliable methods for estimating the quality of the hPSCs cultures is an urgent requirement for their reliable use in the clinic. Currently, much attention is paid to the creation of the automatic methods for selecting the best clones based on their images, as noninvasive methods for their evaluation. The first section of our chapter is devoted to these methods with a particular emphasis of our approaches [6, 7] based on the analysis of the morphology of colonies and cells. However, the search for the new approaches to analyze morphological parameters should not stop and the question of the regulation of the cell morphology deserves a separate chapter.

Our proteomic data for the first time demonstrated cytoskeletal proteins as top effectors of the "good" morphological hPSCs phenotype. As discussed above, most of these cytoskeletal proteins have not been studied in detail in hPSCs. The molecular differences on the proteome level between hiPSCs and hESCs lines, as reported in multiple publications, may be related to many factors such as time in culture, methods of cells propagation, general culture conditions, as well as different somatic origin of hiPSCs, the level of pluripotency, and many others [10, 54, 55]. Regardless of the used approaches and cell lines, all proteomics results revealed a large proportion of cytoskeletal proteins, thus highlighting cytoskeletal remodeling as a prominent characteristic for hPSCs phenotype [8–10, 12, 13]. That is not surprising, as the actin cytoskeleton network, consisting of actin filaments and crosslinking and motor proteins, regulates the shape of the most cells.

Understanding the mechanisms responsible for the dynamic changes of the colony morphology from the "good" to the "bad" is an important prerequisite for the safe clinical application of these cells, not only because the differentiation potential of hPSCs is deeply associated with the colony morphology but also because the morphological changes occur quicker and well before significant changes in the pluripotency markers expression profiles can be detected [47, 56]. Human PSC colonies demonstrate fast changes of morphological parameters during the exponential growth, and essential differences in their structure associated with the colony area, mean nuclei area, and mean distance between nearest neighbors were shown to be good indicators to detect possible changes of the pluripotency status [57]. So far, we are only making the first steps toward the complete understanding of this process.

Based on our data, we propose to expand the panel of hPSCs markers used to identify the "best" morphology phenotype to include the cytoskeletal proteins, namely MYH7, RDX, and CNN3 for evaluation of the best hESCs and EZRIN for evaluation of hiPSCs. Obviously, the quest for the reliable markers for the identification of the best morphology has to continue.

Eventually, the development of more complex automated approaches for comparative analysis of cells will provide the best quality control of clones, which will thus ensure their continued safe application in regenerative medicine.

Acknowledgements

This study was funded by the Russian Science Foundation, grant number 21-75-20132 for IN. Proteomic analysis was performed using the equipment of the "Human Proteome" Core Facility Center of the Institute of Biomedical Chemistry (Moscow, Russia). The authors express their deep gratitude to Prof. S.L. Kiselev and A.V. Panova for providing the hiPSC line HPCASRi002-A (CaSR), and to Dr. Ivan Diakonov from Imperial College London for valuable comments and suggestions during the paper preparation. We are also thankful to the staff of the Institute of Cytology core microscopic facility M.L. Vorobev and G.I. Stein.

Conflict of interest

The authors declare no conflict of interest.

Author details

Vitaly Gursky[1†], Olga Krasnova[1†], Julia Sopova[1], Anastasia Kovaleva[1], Karina Kulakova[1], Olga Tikhonova[2] and Irina Neganova[1*]

1 Institute of Cytology, Saint Petersburg, Russia

2 Institute of Biomedical Chemistry, Moscow, Russia

*Address all correspondence to: irina.neganova@incras.ru

† These authors contributed equally.

References

[1] Hanna J et al. Direct cell reprogramming is a stochastic process amenable to acceleration. Nature. 2009;**462**(7273):595-601

[2] Deinsberger J, Reisinger D, Weber B. Global trends in clinical trials involving pluripotent stem cells: A systematic multi-database analysis. NPJ Regenerative Medicine. 2020;**5**:15

[3] Chin MH et al. Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. Cell Stem Cell. 2009;**5**(1):111-123

[4] Tan Y et al. Comparative study using Raman microspectroscopy reveals spectral signatures of human induced pluripotent cells more closely resemble those from human embryonic stem cells than those from differentiated cells. Analyst. 2012;**137**(19):4509-4515

[5] Kilpinen H et al. Common genetic variation drives molecular heterogeneity in human iPSCs. Nature. 2017;**546**(7658):370-375

[6] Krasnova OA, Gursky VV, Chabina AS, Kulakova KA, Alekseenko LL, Panova AV, et al. Prognostic analysis of human pluripotent stem cells based on their morphological portrait and expression of pluripotent markers. International Journal of Molecular Sciences. 2022;**23**:12902

[7] Mamaeva A, Krasnova O, Khvorova I, Kozlov K, Gursky V, Samsonova M, et al. Quality control of human pluripotent stem cell colonies by computational image analysis using convolutional neural networks. International Journal of Molecular Sciences. 2023;**24**:140

[8] Brill LM et al. Phosphoproteomic analysis of human embryonic stem cells. Cell Stem Cell. 2009;**5**(2):204-213

[9] Van Hoof D et al. Proteomics and human embryonic stem cells. Stem Cell Research. 2008;**1**(3):169-182

[10] Pripuzova NS et al. Development of a protein marker panel for characterization of human induced pluripotent stem cells (hiPSCs) using global quantitative proteome analysis. Stem Cell Research. 2015;**14**(3):323-338

[11] Soteriou D et al. Comparative proteomic analysis of supportive and unsupportive extracellular matrix substrates for human embryonic stem cell maintenance. The Journal of Biological Chemistry. 2013;**288**(26):18716-18731

[12] Novak A et al. Proteomics profiling of human embryonic stem cells in the early differentiation stage. Stem Cell Reviews and Reports. 2012;**8**(1):137-149

[13] Jadaliha M et al. Quantitative proteomic analysis of human embryonic stem cell differentiation by 8-plex iTRAQ labelling. PLoS One. 2012;**7**(6):e38532

[14] Bjorlykke Y et al. Reprogrammed cells display distinct proteomic signatures associated with colony morphology variability. Stem Cells International. 2019;**2019**:8036035

[15] Healy L, Ruban L. Atlas of Human Pluripotent Stem Cells in Culture. New York, NY: Springer US: Imprint: Springer; 2015. p. 1 online resource (XV, 206 pages 285 illustrations, 279 illustrations in color)

[16] Yu J et al. Induced pluripotent stem cell lines derived from human somatic cells. Science. 2007;**318**(5858):1917-1920

[17] Wakui T et al. Method for evaluation of human induced pluripotent stem cell quality using image analysis based on the biological morphology of cells. Journal of Medical Imaging (Bellingham). 2017;**4**(4):044003

[18] Maddah M et al. A system for automated, noninvasive, morphology-based evaluation of induced pluripotent stem cell cultures. Journal of Laboratory Automation. 2014;**19**(5):454-460

[19] Tokunaga K et al. Computational image analysis of colony and nuclear morphology to evaluate human induced pluripotent stem cells. Scientific Reports. 2014;**4**:6996

[20] Kato R et al. Parametric analysis of colony morphology of non-labelled live human pluripotent stem cells for cell quality control. Scientific Reports. 2016;**6**:34009

[21] Nishimura K et al. Live-cell imaging of subcellular structures for quantitative evaluation of pluripotent stem cells. Scientific Reports. 2019;**9**(1):1777

[22] Cho YM et al. Dynamic changes in mitochondrial biogenesis and antioxidant enzymes during the spontaneous differentiation of human embryonic stem cells. Biochemical and Biophysical Research Communications. 2006;**348**(4):1472-1478

[23] Folmes CD et al. Somatic oxidative bioenergetics transitions into pluripotency-dependent glycolysis to facilitate nuclear reprogramming. Cell Metabolism. 2011;**14**(2):264-271

[24] Prigione A et al. The senescence-related mitochondrial/oxidative stress pathway is repressed in human induced pluripotent stem cells. Stem Cells. 2010;**28**(4):721-733

[25] Nishimura K et al. A role for KLF4 in promoting the metabolic shift via TCL1 during induced pluripotent stem cell generation. Stem Cell Reports. 2017;**8**(3):787-801

[26] Harkness L et al. Media composition modulates human embryonic stem cell morphology and may influence preferential lineage differentiation potential. PLoS One. 2019;**14**(3):e0213678

[27] Wakao S et al. Morphologic and gene expression criteria for identifying human induced pluripotent stem cells. PLoS One. 2012;**7**(12):e48677

[28] Joutsijoki H et al. Machine learning approach to automated quality identification of human induced pluripotent stem cell colony images. Computational and Mathematical Methods in Medicine. 2016;**2016**:3091039

[29] Perestrelo T et al. Pluri-IQ: Quantification of embryonic stem cell pluripotency through an image-based analysis software. Stem Cell Reports. 2018;**11**(2):607

[30] Witmer A, Bhanu B. Generative adversarial networks for morphological-temporal classification of stem cell images. Sensors (Basel). 2021;**22**(1):206

[31] Wakui T et al. Predicting reprogramming-related gene expression from cell morphology in human induced pluripotent stem cells. Molecular Biology of the Cell. 2023;**34**(5):ar45

[32] Harkness L et al. Identification of a membrane proteomic signature for human embryonic stem cells independent of culture conditions. Stem Cell Research. 2008;**1**(3):219-227

[33] Phanstiel DH et al. Proteomic and phosphoproteomic comparison of human ES and iPS cells. Nature Methods. 2011;**8**(10):821-827

[34] My I, Di Pasquale E. Genetic cardiomyopathies: The lesson learned from hiPSCs. Journal of Clinical Medicine. 2021;**10**(5):1149

[35] Harb N, Archer TK, Sato N. The Rho-Rock-Myosin signaling axis determines cell-cell integrity of self-renewing pluripotent stem cells. PLoS One. 2008;**3**(8):e3001

[36] Chen G et al. Actin-myosin contractility is responsible for the reduced viability of dissociated human embryonic stem cells. Cell Stem Cell. 2010;**7**(2):240-248

[37] Watanabe K et al. A ROCK inhibitor permits survival of dissociated human embryonic stem cells. Nature Biotechnology. 2007;**25**(6):681-686

[38] Li G et al. Transcriptomic profiling maps anatomically patterned subpopulations among single embryonic cardiac cells. Developmental Cell. 2016;**39**(4):491-507

[39] Sato N et al. A gene family consisting of ezrin, radixin and moesin. Its specific localization at actin filament/plasma membrane association sites. Journal of Cell Science. 1992;**103**(Pt 1):131-143

[40] Sato N et al. Radixin, a barbed end-capping actin-modulating protein, is concentrated at the cleavage furrow during cytokinesis. The Journal of Cell Biology. 1991;**113**(2):321-330

[41] Persson A, Lindberg OR, Kuhn HG. Radixin inhibition decreases adult neural progenitor cell migration and proliferation in vitro and in vivo. Frontiers in Cellular Neuroscience. 2013;**7**:161

[42] Neganova I. The role of cell cycle regulation on reprogramming efficiency. In: Birbrair A, editor. Elsevier Series "Advances in Stem Cell Biology". Cambridge, Massachusetts, U.S.A.: Academic Press; 2020. pp. 1-42

[43] Shibukawa Y et al. Calponin 3 regulates actin cytoskeleton rearrangement in trophoblastic cell fusion. Molecular Biology of the Cell. 2010;**21**(22):3973-3984

[44] Ciuba K et al. Calponin-3 is critical for coordinated contractility of actin stress fibers. Scientific Reports. 2018;**8**(1):17670

[45] Maddala R et al. Calponin-3 deficiency augments contractile activity, plasticity, fibrogenic response and Yap/Taz transcriptional activation in lens epithelial cells and explants. Scientific Reports. 2020;**10**(1):1295

[46] Liao Y et al. Paradoxical roles of elongation factor-2 kinase in stem cell survival. The Journal of Biological Chemistry. 2016;**291**(37):19545-19557

[47] Neganova I et al. Expression and functional analysis of G1 to S regulatory components reveals an important role for CDK2 in cell cycle regulation in human embryonic stem cells. Oncogene. 2009;**28**(1):20-30

[48] Ullmann U et al. Epithelial-mesenchymal transition process in human embryonic stem cells cultured in feeder-free conditions. Molecular Human Reproduction. 2007;**13**(1):21-32

[49] Van Hoof D et al. Feeder-free monolayer cultures of human embryonic stem cells express an epithelial plasma membrane protein profile. Stem Cells. 2008;**26**(11):2777-2781

[50] Vuoristo S et al. Laminin isoforms in human embryonic stem cells: Synthesis,

receptor usage and growth support. Journal of Cellular and Molecular Medicine. 2009;**13**(8B):2622-2633

[51] Fu CA et al. TNIK, a novel member of the germinal center kinase family that activates the c-Jun N-terminal kinase pathway and regulates the cytoskeleton. The Journal of Biological Chemistry. 1999;**274**(43):30729-30737

[52] Di J et al. Rap2B promotes cell proliferation, migration and invasion in prostate cancer. Medical Oncology. 2016;**33**(6):58

[53] Franks TM, Hetzer MW. The role of Nup98 in transcription regulation in healthy and diseased cells. Trends in Cell Biology. 2013;**23**(3):112-117

[54] Benevento M, Munoz J. Role of mass spectrometry-based proteomics in the study of cellular reprogramming and induced pluripotent stem cells. Expert Review of Proteomics. 2012;**9**(4):379-399

[55] Kim SY et al. Comparative proteomic analysis of human somatic cells, induced pluripotent stem cells, and embryonic stem cells. Stem Cells and Development. 2012;**21**(8):1272-1286

[56] Neganova I et al. CDK1 plays an important role in the maintenance of pluripotency and genomic stability in human pluripotent stem cells. Cell Death & Disease. 2014;**5**(11):e1508

[57] Orozco-Fuentes S et al. Quantification of the morphological characteristics of hESC colonies. Scientific Reports. 2019;**9**(1):17569