Stem Cell Reports

Report

ISSCR

OPEN ACCESS

Structural Phenotyping of Stem Cell-Derived Cardiomyocytes

Francesco Silvio Pasqualini,[1] Sean Paul Sheehy,[1] Ashutosh Agarwal,[1] Yvonne Aratyn-Schaus,[1] and Kevin Kit Parker[1,*]

[1]Disease Biophysics Group, Wyss Institute for Biologically Inspired Engineering, School of Engineering and Applied Sciences, Harvard Stem Cell Institute, Harvard University, Cambridge, MA 02138, USA
*Correspondence: kkparker@seas.harvard.edu
http://dx.doi.org/10.1016/j.stemcr.2015.01.020

SUMMARY

Structural phenotyping based on classical image feature detection has been adopted to elucidate the molecular mechanisms behind genetically or pharmacologically induced changes in cell morphology. Here, we developed a set of 11 metrics to capture the increasing sarcomere organization that occurs intracellularly during striated muscle cell development. To test our metrics, we analyzed the localization of the contractile protein α-actinin in a variety of primary and stem-cell derived cardiomyocytes. Further, we combined these metrics with data mining algorithms to unbiasedly score the phenotypic maturity of human-induced pluripotent stem cell-derived cardiomyocytes.

INTRODUCTION

Several efforts have been reported in the emerging field of structural phenotyping for the integration of image acquisition, processing, and analysis to assess the response of cells and tissues to various challenges (Eliceiri et al., 2012). All of these methodologies are predicated on the assumption that cell shape is an important indicator of the cell pathophysiological state and rely on (1) image-processing algorithms for the extraction of morphological features and (2) machine-learning strategies to mine the cell morphology data (Crane et al., 2012; Jones et al., 2009; Treiser et al., 2010).

Many intracellular structures, such as the contractile cytoskeleton in striated muscles, are also predictors of cell function (Feinberg et al., 2007). Additionally, while a cell specifies along the myocyte lineage (Mummery et al., 2012; Qian and Srivastava, 2013), it also progresses through myofibrillogenesis as force-generating units, known as sarcomeres, self-assemble along the actin cytoskeleton (Grosberg et al., 2011; Parker et al., 2008). Moreover, the contractile proteins of mature myocytes are continuously turned over and their spatial organization remodeled to adapt to pathophysiological stimuli (McCain et al., 2013; Sun et al., 2012; Wang et al., 2014). Therefore, while the presence of contractile proteins is necessary for myocytes function (Cahan et al. 2014; Mummery et al., 2012), it is not sufficient to define the developmental stage (Sheehy et al., 2014), the health status (Wang et al., 2014), or the functional capabilities of these cells (Feinberg et al., 2012).

Here, we designed a set of 11 metrics (Figure S1; Table 1) that intrinsically score myocyte structural phenotypes by the increasing degree of organization and alignment that sarcomeres acquire during myofibrillogenesis. We utilized these metrics to score the phenotypic maturity of primary and stem cell-derived cardiomyocytes based on the degree of sarcomeric structural organization observed in α-actinin images.

RESULTS

Quantitative Analysis of the Contractile Cytoskeleton in Striated Muscle Cells

Sarcomeres are ∼2 μm long ultrastructures delimited by Z-disks that are rich in the contractile protein α-actinin (red in Figure 1A). The localization of this protein can be taken to indicate the maturity of cardiac myocytes (Grosberg et al., 2011): in differentiating cells, α-actinin is diffuse in the cytoplasm (Figure 1Aii); in immature myocytes, it appears as a fibrous structure or as aperiodically spaced puncta known as Z-bodies (Figure 1Aiii); and in mature myocytes, α-actinin localizes into the regular lattice of Z-disks (Figure 1Aiv). Therefore, we focused our analysis (Figure 1B) on identifying how regularly spaced and well-aligned sarcomeric α-actinin-positive structures were in the images. We first associated each sarcomeric α-actinin-positive pixel with the orientation (color coded in Figure 1Ci) of its local neighborhood and then fitted a bimodal distribution to the resulting orientation histogram (red and black curves in Figure 1Cii). This enabled the extraction of several metrics: the global orientational order parameter (Grosberg et al., 2011) (OOP) a value that ranges from 0 to 1 as contractile elements become more aligned and separate OOPs for the two fitted modal distributions, representing Z-disks and Z-bodies, as well as their relative presence. Further, we radially integrated the image power spectrum (Figure 1Ciii), yielding a 1D

Table 1. List of Metrics of Sarcomere Organization Developed, Integrated, or Updated for This Study

Metric Number	Metric Name	Metric Description
1	sarcomere length (SL)	The average distance between Z-disks in the entire field of view.
2	total energy	The total amount of spatially varying immunosignal.
3	sarcomeric energy	The amount of immunosignal with a spatial periodicity given by SL.
4	sarcomeric packing density (SPD)	The fraction of immunosignal that localized in a regular lattice at a distance SL.
5	orientational order parameter (OOP)	The degree of alignment of all foreground elements in the field of view.
6	sarcomeric OOP (OOP1)	The degree of alignment of foreground elements that are oriented orthogonally to the actin bundles (Z-disks).
7	nonsarcomeric OOP (OOP2)	The degree of alignment of foreground elements that are oriented parallel to the actin bundles (Z-bodies).
8	Z-disks relative presence (γ)	The fraction of foreground elements that are recognized as Z-bodies.
9	weighted OOP	Calculated by multiplying the sarcomeric OOP (OOP1) and the weight γ. In this sense, it represents both the abundance and relative alignment of the Z-disks in the imag.
10	coverage quality control	Calculated as the percentage of the image pixels that have an intensity value higher than user-specified threshold. It estimates the α-actinin coverage.
11	coherency quality control	Calculated as the percentage of α-actinin-positive pixels that have a value of the coherency higher than a user-selected threshold. It is useful for artifact removal and for image quality control.

See Figure S1 for a schematic representation of the role of each parameter.

representation (Figure 1Civ) that highlights the relative importance of each spatial frequency. In particular, the peaks (red curve) represent α-actinin-positive elements arranged at a distance on the order of the sarcomere length (SL) and therefore become more prominent as sarcomerogenesis progresses. Through nonlinear fitting, we separated this component from the aperiodic contribution of Z-bodies and other image artifacts (black curve). Relevant metrics in this case included the area (shaded in red) under the periodic component, the total area under the data curve, and their ratio, a quantity we termed sarcomere packing density (SPD). Taken together, this set of metrics has a direct biophysical interpretation: substantially more mature myofibrillar architectures exhibit a regular lattice of well-oriented Z-disks, resulting in elevated values of SPD and OOP. Additionally, this analysis is robust to common imaging artifacts such as out-of-focus blurriness, salt-and-pepper noise, or poor contrast (Figure S2).

Quantitative Analysis of the Contractile Cytoskeleton in Murine Primary and Stem Cell-Derived Single Cardiomyocytes

To test our analysis tool, we asked whether we could quantify the ability of human and murine-induced pluripotent stem cell-derived cardiomyocytes (hiCMs and miCMs, respectively) to replicate the contractile cytoskeletal architecture observed in murine primary cardiomyocytes (mpCMs) in vitro (Parker et al., 2008). Qualitatively, we observed that mpCMs (Figure 1Di) and miCMs (Figure 1Dii) 3 days after seeding on square fibronectin islands showed mature myofibrillar architecture, characterized by uniformly distributed sarcomeric α-actinin-rich striations. Conversely, hiCMs (Figure 1Diii) exhibited sparse Z-disks solely in the perinuclear region and arranged in ring-like myofibrils (red arrow in Figure 1Diii). In addition, close to the hiCM membrane, the actin and sarcomeric α-actinin signals were diffuse (yellow arrows) and resembled the cortical architecture observed in immature and/or migratory cells (Parker et al., 2008; Sheehy et al., 2012). Quantitatively, while local regions of aligned Z-disks could be detected (color-coded insets below the panels in Figure 1D), the Fourier analysis clearly demonstrated a globally reduced periodicity in the sarcomere distribution of hiCMs than observed in miCMs and mpCMs (insets on the right of the panels in Figure 1D). Consistently, the SPDs measured in mpCMs and miCMs were two times higher than in hiCMs (Figure S3). Notably, all myocytes considered in this study were positive for sarcomeric α-actinin, suggesting that they would have been clustered in the same group by traditional assays detecting the presence of this protein or its transcript (Cahan et al. 2014; Mummery et al., 2012).
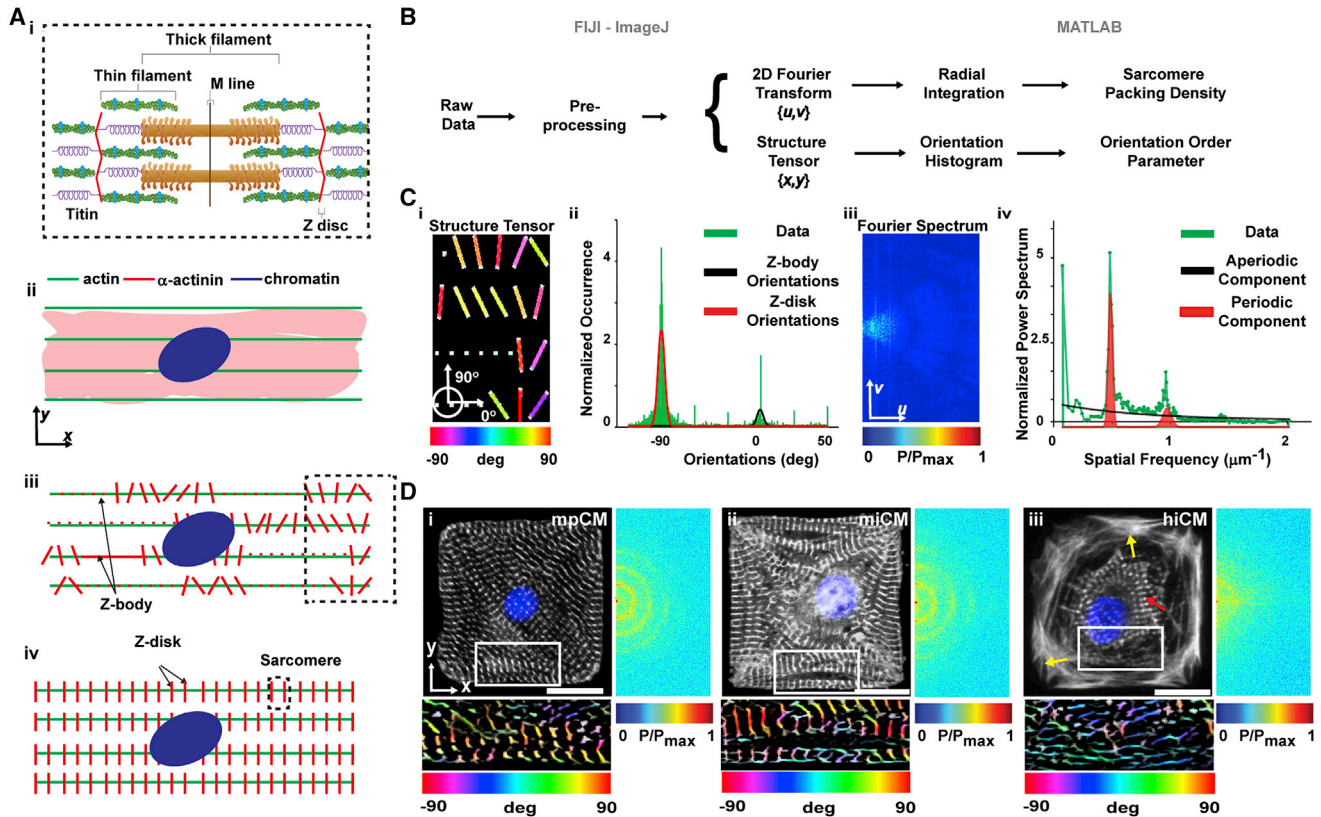
Figure 1. Metrics of Contractile Architecture to Characterize the Progression of Myofibrillogenesis

(A) Schematic representation of a sarcomere (i) and of the distribution of α-actinin (red) during myofibrillogenesis: in the cytoplasm (ii), along the actin (green) filament in the form of Z-bodies (iii), and incorporated into the Z-disks (iv).

(B) Algorithmic detection of the orientation and periodic registration of α-actinin-positive structures using the image spatial (co-ordinates x,y) and Fourier (coordinates u,v) domains.

(C) Color-coded orientations (i, from the inset of synthetic image Figure 1Bii) displayed into a histogram (ii) can be fitted to identify orientations belonging to Z-disks (red) and Z-bodies (black). In parallel, the 2D Fourier power spectrum (iii) was integrated into a 1D curve (iv) and fitted to identify the contribution of periodically spaced Z-disks (red) and aperiodic Z-bodies (black).

(D) α-actinin immunostains (white) of mononucleated (DAPI, blue) murine primary (mpCM, i) and murine (miCM, ii) or human (hiCM, iii) induced pluripotent stem cell-derived cardiomyocytes. The color-coded representation of the α-actinin orientation in the inset is reported below the image. The positive semiplane for the Fourier transform is reported on the right of each image.

Scale bar represent 20 μm. See also Figures S2 and S3.

Quantitative Analysis of the Maturity of the Contractile Cytoskeleton in Human Stem Cell-Derived Cardiomyocytes

We and others have previously shown that extending time in culture (Lundy et al., 2013; McCain et al., 2014a) could be beneficial for obtaining hiCMs with a more mature phenotype than we observed here. However, evaluating the quality of mass-produced stem cell-derived myocytes requires an extensive structure-function characterization and a direct comparison against myocytes exhibiting a postnatal phenotype (Sheehy et al., 2014). We reasoned that since the process of myofibrillogenesis is highly conserved across species (Sissman, 1970) we could design a prescreening tool that estimates the effectiveness of maturation strategies by integrating our metrics of myofi-brillar architecture with machine-learning algorithms for structural phenotyping. While available platforms require a user-selected training set, a set of images that an "expert" assigns to all phenotypic classes (Eliceiri et al., 2012), we took advantage of our ability to recapitulate myofibrillo-genesis in vitro using primary cardiomyocytes from neonate rats (rpCMs) (Agarwal et al., 2013; Parker et al., 2008) to create such a training set. First, we cultured rpCMs and hiCMs as engineered tissues that mimic the native architecture of the myocardium (Sheehy et al., 2014) (Figures 2A and S4). Second, we assigned the sets of features (Figure S1; Table 1) extracted from images collected at 6 (Figure 2Ai), 24 (Figure 2Aii), and 48 hr
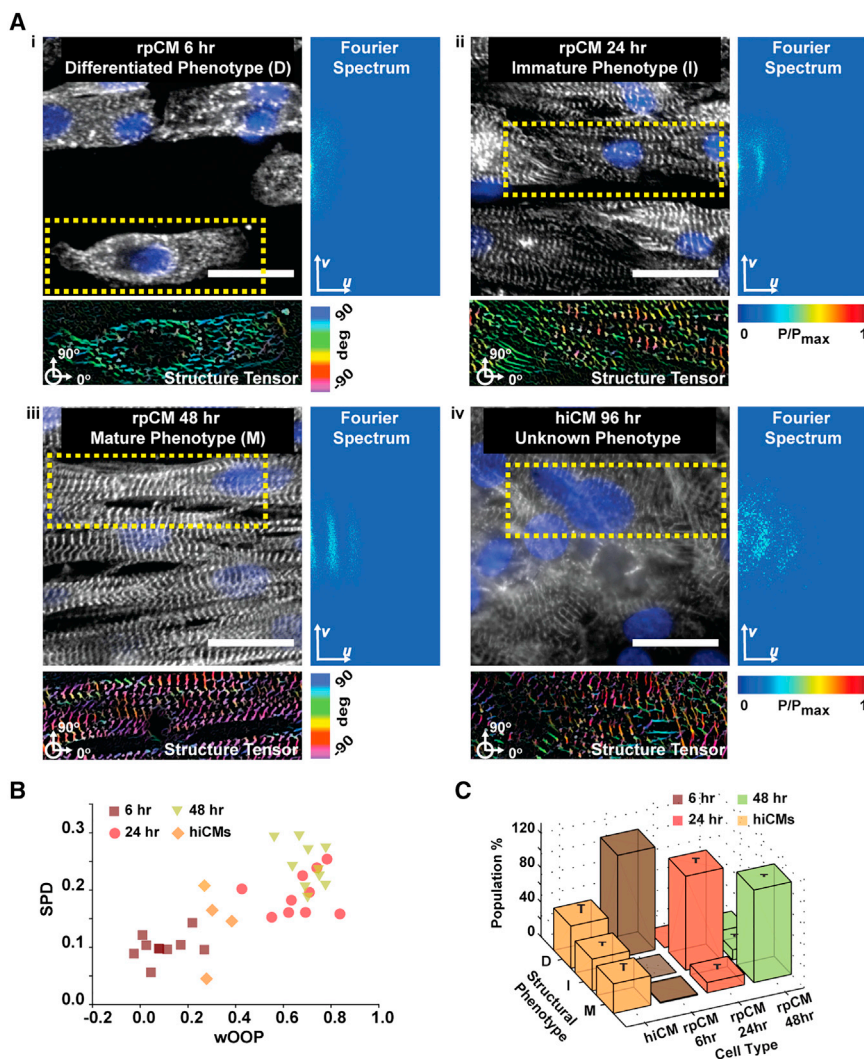
Figure 2. Structural Phenotyping of Stem Cell-Derived Cardiomyocytes

(A) α-actinin (white) and chromatin (blue) images of rpCMs at 6 (i), 24 (ii), and 48 hr (iii) as well as hiCMs at 72 hr (iv) after seeding with color-coded orientations and Fourier representations. Scale bar represents 25 μm.

(B) Scatter plot showing how our metrics of myofibrillar architecture quantitatively captured the progression of myofibrillogenesis in rpCM tissues from differentiated (6 hr, brown squares) to immature (24 hr, red circles) and finally mature (48 hr, green triangles) myocytes. In contrast, the hiCM tissues (orange diamonds) exhibited a relatively immature myofibrillar organization.

(C) A dataset comprising ~120 sarcomeric α-actinin images per conditions (insets in Ai–Aiv) was acquired, and the features extracted from this dataset were utilized to train several classifiers to distinguish the classes of differentiated (D), immature (I), and mature (M) myocytes. The classifiers assigned only ~29% of the 118 hiCM images to the class of myocytes with a mature structural phenotype, with the rest showing differentiated or immature contractile architectures.

Results are shown as mean ± SEM. See also Figure S4.

(Figure 2Aiii) after seeding to the classes of differentiated, immature, and mature myocytes, respectively. We collected digital images from more than 100 cells (insets in Figure 2A and Figure S4A–B) in each condition. Third, we utilized the automatically annotated dataset to train a simple naive Bayes classifier as well as two more advanced supervised learning strategies based on neural networks and tree bagging (Figure S4C). We selected three machine-learning strategies that operate under various assumptions (Table 2) to demonstrate that our classification was not biased (Eliceiri et al., 2012) by the choice of one specific algorithm. Finally, we asked the three classifiers to confirm whether or not hiCMs (Figure 2Aiv) possess a mature structural architecture.

At the tissue level, as rpCMs in culture underwent myofibrillogenesis, we observed that weighted OOP (wOOP) and SPD increased as expected (Figure 2B). In comparison, hiCMs scored values consistent with their immature

myofibrillar organization. Further, all classifiers trained on the rpCMs dataset failed to recognize a mature myofibrillar architecture in the majority of hiCMs images (Figure S4D). Specifically, ~70% of the 118 hiCM images were not classified as mature by the naive Bayes classifier, ~71% by the neural network classifier, and ~77% by the tree bagging classifier. For example, the neural network classified ~40% of hiCMs as differentiated myocytes and ~31% as immature myocytes (Figure 2C). Interestingly, 29% of the cells embedded in anisotropic hiCM tissues did display mature myofibrillar architectures, suggesting a pool of hiCM with enhanced myogenic potential may exist (Hartjes et al., 2014). Thus, we (1) provided quantitative metrics for the organization of the contractile cytoskeleton of primary and stem cell-derived cardiomyocytes and (2) utilized this information to unbiasedly and robustly quantify their maturation.

Table 2. Machine Learning Algorithms Adopted for the Analysis of the Myofibrillogenesis Dataset

Classifier Number	Classifier Name	Pros	Cons
1	naive Bayes	One of the simplest classifiers, based on intuitive probability models, and it is computationally very treatable.	It simplistically assumes that all the features are statistically independent, which may not be true for sarcomeres whose structure, during myofibrillogenesis, becomes more periodic and well-aligned.
2	neural network (NN)	A popular machine learning algorithm for structural phenotyping. Extensive literature shows how a NN classifier can always be constructed, providing that one has a good enough dataset.	The data model is not intuitive. The neural network optimization is not trivial and requires considerations for the dataset size as well as the stochastic initialization.
3	tree bagging	A popular machine-learning algorithm for structural phenotyping. The data model is more intuitive than NN.	Optimization of the tree bagging algorithm is not trivial and requires careful consideration of the sample and tree sizes.

See Figure S4.

DISCUSSION

Quantitative methods to characterize the contractile cytoskeleton of striated muscle cells have been previously proposed. For example, the analysis of the orientation of intracellular elements has been conducted adopting mean orientations (Rao et al., 2013). Unfortunately, the specific circular statistics tests (Berens, 2009) required to compare and contrast these mean orientations are only rarely adopted. By using the OOP values in the range of 0–1, we were able to employ classical statistical tools, such as ANOVA (Figure S3). Additionally, approaches based on nonlinear fitting of multiple Von Mises distributions have been proposed (Rezakhaniha et al., 2012); in the context of fluorescence bioimaging though, these methods may suffer from overfitting issues, given the large numbers of parameters that are needed to identify multiple distributions. Here we restricted ourselves to only two distributions, centered on the two biophysically relevant principal directions of Z-disks and Z-bodies that are orthogonal to one another. We were thus able to limit the number of fitting parameters, largely reducing the risk of overfitting. Fourier analysis has also been previously considered in the estimation of sarcomere length (Lundy et al., 2013; Wei et al., 2010). Our automatic approach offers significant advantages in that the contractile cytoskeleton within the entire cell is considered, reducing the user bias (Eliceiri et al., 2012) introduced by manual selections in the spatial (Lundy et al., 2013) or Fourier (Wei et al., 2010) domains. Moreover, our algorithm not only yields an estimate of the sarcomere length across the entire spatial extension of the cell/tissue, but also reveals the relative presence of well-formed sarcomeres irrespectively from the direction of their alignment. When the myofi-brils are highly aligned (Figure 2Aiii), the Fourier spectrum exhibits peaks along one principal direction; conversely, when sarcomeres are well organized along many directions (Figure 1Di), the Fourier spectrum shows a circular pattern. By integrating across all directions in the Fourier domain, we are able to quantify the sarcomeres translational periodicity across all directions in the spatial domain.

Finally, the method proposed here for the calculation of the SPD significantly improves our previous efforts (McCain et al., 2014b; Wang et al., 2014). By normalizing the energy of the periodic component to the total energy of the sarcomeric α-actinin image, we estimate a signal-to-noise ratio that is bound in the unit interval, a desirable property for many machine-learning algorithms (Shamir et al., 2010).

In conclusion, we have developed 11 metrics to characterize the structural phenotype of primary and stem cell-derived cardiomyocytes in a way that is biophysically related to their functional proficiency (Feinberg et al., 2012). Moreover, by engineering myocyte shape and tissue architecture, we were able to generate a myofibrillogenesis dataset that allows structural phenotyping of stem cell-derived cardiomyocytes in an unbiased fashion and that is largely robust to the choice of a specific machine-learning strategy. Finally, while assessing the quality of human pluripotent stem cell-derived myocytes remains critical, to date, their maturation has been suboptimally estimated (Mummery et al., 2012; Sheehy et al., 2014), as healthy human myocytes are not readily available. Since myofibrillogenesis is an extremely well-conserved physiological process (Sissman, 1970), our method allows for a quantitative characterization of myocytes maturation that naturally overcomes this limitation.

EXPERIMENTAL PROCEDURES

Substrate Engineering

Photolithographic masks bearing desired features were drawn in AUTOCAD (Autodesk). UV light was shone through the mask into a silicon wafer (Wafer World), previously spin coated with SU-8 3005 photoresist (MicroChem). The wafer was then developed in propylene glycol methyl ether acetate (Sigma) and utilized to cast polydimethylsiloxane (PDMS, Sylgard 184; Dow Corning; 10:1 ratio) stamps. Glass coverslips were also coated with PDMS and treated for 8 min in the UV ozone cleaner before coming into contact with PDMS stamps inked with fibronectin (50 μg/ml; BD Biosciences).

Alternatively, coverslips were coated with polyacrylamide hydrogels (streptavidin-acrylamide/bis ratio 7.5/0.3%). To transfer fibronectin islands, the hydrogel was dried (37°C, 10 min), brought into contact with stamps inked with biotinylated fibronectin (Pierce), and finally sterilized with UV exposure (15 min).

Primary Harvest

Ventricular myocytes were isolated from day 2 neonate Balb/c mice and Sprague Dawley rats according to procedures approved by the Harvard University institutional animal care and use committees. Isolated ventricles were incubated in cold (4°C) 0.1% (w/v) trypsin (USB) solution for approximately 12 hr. Ventricular tissue was further exposed to serial treatments (2 min each) of 0.1% (w/v) warm (37°C) collagenase type II (Worthington Biochemical) solution. Isolated rat and mouse cardiomyocytes were seeded onto engineered substrates at a density of 10,000 and 20,000 cells/cm^2, respectively. Culture medium consisted of Medium 199 (Invitrogen) supplemented with 10% (v/v) heat-inactivated fetal bovine serum (FBS), 10 mM HEPES, 20 mM glucose, 2 mM L-glutamine, 1.5 μl vitamin B12, and 50 U/ml penicillin for the first 48 hr. The FBS concentration was then reduced to 2%.

Stem Cell Culture

hiCM and miCMs were kindly provided by Cellular Dynamics and Axiogenesis (CorAt-iPS), respectively. Cells were cultured as per manufacturers' recommendations: hiCMs were seeded in the presence of vendor-provided plating medium; miCMs were positively selected after plating onto 10 mg/ml fibronectin coated flasks. After 72 hr, both cell types were dissociated with 0.05% trypsin-EDTA solution (Invitrogen, 25200-072) and seeded onto the engineered substrates at a density of 10,000 cells/cm^2.

Immunocytochemistry and Imaging

At room temperature (RT), cells were treated with 4% paraformaldehyde and 0.05% Triton X-100 in PBS (v/v) for 10 min and incubated with anti-sarcomeric α-actinin (A7811; Sigma) and anti-fibronectin antibodies (F3648; Sigma) for 1 hr (1:200 dilution). Samples were further treated with DAPI (Invitrogen), Alexa Fluor 633-conjugated phalloidin (A22284, Invitrogen), and Alexa Fluor 488-conjugated goat anti-mouse IgG and Alexa Fluor 546-conjugated goat anti-rabbit IgG secondary antibodies (Invitrogen) for 2 hr at RT. Samples were imaged with a Zeiss LSM confocal microscope (Carl Zeiss Microscopy) equipped with the EC Plan-

Neofluar 40×/1.30 oil DIC M27 objective; 1024 × 1024 pixels per image were acquired for a final pixel size of 160 nm.

Image Processing and Future Extraction

Preprocessing was performed with ImageJ/FIJI (Schindelin et al., 2012). The tubeness (Sato et al., 1998) and OrientationJ (Rezakhaniha et al., 2012) plugins were used to highlight the filamentous structure and calculate the orientations of foreground pixels, respectively. MATLAB (Mathworks) was adopted for feature extraction.

Orientational Order Parameter

The structure tensor method generated orientations $\{\vartheta_1, \vartheta_2, \ldots, \vartheta_N\}$ whose frequency of occurrence we plotted in a histogram, such as in Figures 1Ci, 1Cii, and S3. The OOP was calculated using the mean resultant vector from circular statistics (Berens, 2009)

$$OOP = \frac{1}{N}\left|\sum_{j=1}^{N} e^{i\vartheta_j}\right|, \qquad \text{(Equation 1)}$$

where $i = \sqrt{-1}$ is the complex unit, e is Euler's number (~ 2.71), and ϑ_j is the jth orientation in $\{\vartheta_1, \vartheta_2, \ldots, \vartheta_N\}$. The sum of unit vectors in Equation 1 is bound by 0 (for a set of randomly oriented vectors) and 1 (for a set of perfectly aligned vectors). Further, we fitted the orientation histogram with the following linear mix of Von Mises Distributions

$$f(\theta; \mu_1, \delta_1, \mu_2, \delta_2, \gamma) = \gamma \frac{\exp[\delta_1 \cos(\vartheta - \mu_1)]}{2\pi I_0(\delta_1)} + (1 - \gamma) \frac{\exp[\delta_2 \cos(\vartheta - \mu_2)]}{2\pi I_0(\delta_2)}, \qquad \text{(Equation 2)}$$

where $\mu_{1,2}$ and $\delta_{1,2}$ represent the localization and spread parameters for the Z-disks and Z-bodies peak, respectively, γ indicates the fraction of orientations allocated into the Z-disk peak, and I_0 is the modified Bessel function of order 0. We then sampled 1,000 orientations from the two fitted Von Mises probability density distributions and calculated OOP_1 and OOP_2 for the Z-disk and Z-body peaks, respectively. In addition, we introduced a weighted version of the OOP: $wOOP = \gamma * OOP_1$ that quantifies both the presence (γ) and the alignment (OOP_1) of the Z-disks in the image.

Sarcomeric Packing Density

To calculate the degree of spatial organization of sarcomeres, we first considered the Fourier power spectrum $P(u,v)$ of the preprocessed sarcomeric α-actinin image $I(x,y)$

$$\begin{cases} P(u,v) = |F(u,v)|^2 \\ F(u,v) = \iint_{\mathbb{R}^2} I(x,y)\exp[i\,2\pi(xu + yv)]dxdy \end{cases}. \qquad \text{(Equation 3)}$$

Equation 3 uses the Fast Fourier Transform algorithm to establish a correspondence between the spatial domain of the image $I(x,y)$ and the Fourier domain where the power spectrum $P(u,v)$ is defined. The signal energy (E_{TOT}) was sampled and integrated along 1,024 directions to obtain a 1D representation ($\Gamma(\omega)$) that exhibits periodic peaks (subscript p), in correspondence of spatial frequencies that are integer multiple of ω_0, modulated by an aperiodic noise term (subscript ap).

$$E_{TOT} = \iint_{\mathbb{R}^2} P(u,v) du dv = \int_0^\infty \left[\omega \int_{-90}^{90} P(\omega, \vartheta) d\vartheta \right] d\omega = \int_0^\infty [\Gamma(\omega)] d\omega.$$

(Equation 4)

Equation 4 states that the total energy in the image can be expressed integrating the 2D power spectrum $P(u,v)$ or, alternatively, through the integration of its 1D representation $\Gamma(\omega)$. To approximate this 1D function, we chose the relationships in Equation 5, where $\widehat{\Gamma}_{ap}$ is a decaying exponential and $\widehat{\Gamma}_p$ is the sum of three Gaussian peaks.

$$\begin{cases} \widehat{\Gamma}(\omega; \xi) = \widehat{\Gamma}_p(\omega; \xi_p) + \widehat{\Gamma}_{ap}(\omega; \xi_{ap}) \\ \widehat{\Gamma}_{ap}(\omega; \xi_{ap}) = a \exp(-\omega/b); \ \xi_{ap} = \{a, b\} \\ \widehat{\Gamma}_p(\omega; \xi_p) = \sum_{k=1}^3 a_k \exp\left[-(\omega - k\,\omega_0)^2 / b_k \right]; \ \xi_p = \{a_k, b_k, \omega_0\}_{k=1,2,3} \end{cases}.$$

(Equation 5)

We fitted the function to the data using Matlab "lsqnonlin" function that uses the Trust Region Reflective algorithm (Coleman and Li, 1996). With this approach, we were able to calculate (see Equation 6) the sarcomere length the area under the periodic component E_p (or sarcomeric energy), as well as its ratio with the total area (total energy), a quantity we named sarcomeric packing density (SPD):

$$\begin{cases} E_p = \int_D \widehat{\Gamma}_p(\omega; \xi_p) d\omega \\ SPD = E_p / E_{TOT} \\ SL = \omega_0^{-1} \end{cases}.$$

(Equation 6)

Machine Learning

The naive Bayes, neural network, and tree bagging classifiers were implemented using Matlab built-in functions (see also Supplemental Information). Ten random iterations were seeded to ensure that the results were stochastically robust. For the naive Bayes and tree bagging classifiers, training was performed with a 10-fold cross-validation test. For the neural network, we used 70%, 15%, and 15% of the rpCMs dataset for the training, validation, and testing phases, respectively.

Downloadable Content

The myofibrillogenesis dataset and the ImageJ code utilized in this paper can be downloaded from the Disease Biophysics Group website: http://diseasebiophysics.seas.harvard.edu/.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and four figures and can be found with this article online at http://dx.doi.org/10.1016/j.stemcr.2015.01.020.

REFERENCES

Agarwal, A., Farouz, Y., Nesmith, A.P., Deravi, L.F., McCain, M.L., and Parker, K.K. (2013). Micropatterning Alginate Substrates for In Vitro Cardiovascular Muscle on a Chip. Adv. Funct. Mater. 30, 3738–3746.

Berens, P. (2009). CircStat: A MATLAB Toolbox for Circular Statistics. J. Stat. Softw. 31, 1–21.

Cahan, P., Li, H., Morris, S.A., Lummertz da Rocha, E., Daley, G.Q., and Collins, J.J. (2014). CellNet: network biology applied to stem cell engineering. Cell 158, 903–915.

Coleman, T.F., and Li, Y. (1996). An Interior, Trust Region Approach for Nonlinear Minimization Subject to Bounds. SIAM J. Optim. 6, 418–445.

Crane, M.M., Stirman, J.N., Ou, C.-Y., Kurshan, P.T., Rehg, J.M., Shen, K., and Lu, H. (2012). Autonomous screening of C. elegans identifies genes implicated in synaptogenesis. Nat. Methods 9, 977–980.

Eliceiri, K.W., Berthold, M.R., Goldberg, I.G., Ibáñez, L., Manjunath, B.S., Martone, M.E., Murphy, R.F., Peng, H., Plant, A.L., Roysam, B., et al. (2012). Biological imaging software tools. Nat. Methods 9, 697–710.

Feinberg, A.W., Feigel, A., Shevkoplyas, S.S., Sheehy, S., Whitesides, G.M., and Parker, K.K. (2007). Muscular thin films for building actuators and powering devices. Science 317, 1366–1370.

Feinberg, A.W., Alford, P.W., Jin, H., Ripplinger, C.M., Werdich, A.A., Sheehy, S.P., Grosberg, A., and Parker, K.K. (2012). Controlling the contractile strength of engineered cardiac muscle by hierarchal tissue architecture. Biomaterials 33, 5732–5741.

Grosberg, A., Kuo, P.-L., Guo, C.-L., Geisse, N.A., Bray, M.-A., Adams, W.J., Sheehy, S.P., and Parker, K.K. (2011). Self-organization of muscle cell structure and function. PLoS Comput. Biol. 7, e1001088.

Hartjes, K.A., Li, X., Martinez-Fernandez, A., Roemmich, A.J., Larsen, B.T., Terzic, A., and Nelson, T.J. (2014). Selection via pluripotency-related transcriptional screen minimizes the influence of somatic origin on iPSC differentiation propensity. Stem Cells 32, 2350–2359.

Jones, T.R., Carpenter, A.E., Lamprecht, M.R., Moffat, J., Silver, S.J., Grenier, J.K., Castoreno, A.B., Eggert, U.S., Root, D.E., Golland, P., and Sabatini, D.M. (2009). Scoring diverse cellular morphologies in image-based screens with iterative feedback and machine learning. Proc. Natl. Acad. Sci. USA 106, 1826–1831.

Lundy, S.D., Zhu, W.Z., Regnier, M., and Laflamme, M.A. (2013). Structural and functional maturation of cardiomyocytes derived

from human pluripotent stem cells. Stem Cells Dev. *22*, 1991–2002.

McCain, M.L., Sheehy, S.P., Grosberg, A., Goss, J.A., and Parker, K.K. (2013). Recapitulating maladaptive, multiscale remodeling of failing myocardium on a chip. Proc. Natl. Acad. Sci. USA *110*, 9770–9775.

McCain, M.L., Agarwal, A., Nesmith, H.W., Nesmith, A.P., and Parker, K.K. (2014a). Micromolded gelatin hydrogels for extended culture of engineered cardiac tissues. Biomaterials *35*, 5462–5471.

McCain, M.L., Yuan, H., Pasqualini, F.S., Campbell, P.H., and Parker, K.K. (2014b). Matrix elasticity regulates the optimal cardiac myocyte shape for contractility. Am. J. Physiol. Heart Circ. Physiol. *306*, H1525–H1539.

Mummery, C.L., Zhang, J., Ng, E.S., Elliott, D.A., Elefanty, A.G., and Kamp, T.J. (2012). Differentiation of human embryonic stem cells and induced pluripotent stem cells to cardiomyocytes: a methods overview. Circ. Res. *111*, 344–358.

Parker, K.K., Tan, J., Chen, C.S., and Tung, L. (2008). Myofibrillar architecture in engineered cardiac myocytes. Circ. Res. *103*, 340–342.

Qian, L., and Srivastava, D. (2013). Direct cardiac reprogramming: from developmental biology to cardiac regeneration. Circ. Res. *113*, 915–921.

Rao, C., Prodromakis, T., Kolker, L., Chaudhry, U.A.R., Trantidou, T., Sridhar, A., Weekes, C., Camelliti, P., Harding, S.E., Darzi, A., et al. (2013). The effect of microgrooved culture substrates on calcium cycling of cardiac myocytes derived from human induced pluripotent stem cells. Biomaterials *34*, 2399–2411.

Rezakhaniha, R., Agianniotis, A., Schrauwen, J.T.C., Griffa, A., Sage, D., Bouten, C.V.C., van de Vosse, F.N., Unser, M., and Stergiopulos, N. (2012). Experimental investigation of collagen waviness and orientation in the arterial adventitia using confocal laser scanning microscopy. Biomech. Model. Mechanobiol. *11*, 461–473.

Sato, Y., Nakajima, S., Shiraga, N., Atsumi, H., Yoshida, S., Koller, T., Gerig, G., and Kikinis, R. (1998). Three-dimensional multi-scale line filter for segmentation and visualization of curvilinear structures in medical images. Med. Image Anal. *2*, 143–168.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.

Shamir, L., Delaney, J.D., Orlov, N., Eckley, D.M., and Goldberg, I.G. (2010). Pattern recognition software and techniques for biological image analysis. PLoS Comput. Biol. *6*, e1000974.

Sheehy, S.P., Grosberg, A., and Parker, K.K. (2012). The contribution of cellular mechanotransduction to cardiomyocyte form and function. Biomech. Model. Mechanobiol. *11*, 1227–1239.

Sheehy, S.P., Pasqualini, F., Grosberg, A., Park, S.J., Aratyn-Schaus, Y., and Parker, K.K. (2014). Quality metrics for stem cell-derived cardiac myocytes. Stem Cell Reports *2*, 282–294.

Sissman, N.J. (1970). Developmental landmarks in cardiac morphogenesis: comparative chronology. Am. J. Cardiol. *25*, 141–148.

Sun, N., Yazawa, M., Liu, J., Han, L., Sanchez-Freire, V., Abilez, O.J., Navarrete, E.G., Hu, S., Wang, L., Lee, A., et al. (2012). Patient-specific induced pluripotent stem cells as a model for familial dilated cardiomyopathy. Sci Transl Med *4*, 130ra147.

Treiser, M.D., Yang, E.H., Gordonov, S., Cohen, D.M., Androulakis, I.P., Kohn, J., Chen, C.S., and Moghe, P.V. (2010). Cytoskeleton-based forecasting of stem cell lineage fates. Proc. Natl. Acad. Sci. USA *107*, 610–615.

Wang, G., McCain, M.L., Yang, L., He, A., Pasqualini, F.S., Agarwal, A., Yuan, H., Jiang, D., Zhang, D., Zangi, L., et al. (2014). Modeling the mitochondrial cardiomyopathy of Barth syndrome with induced pluripotent stem cell and heart-on-chip technologies. Nat. Med. *20*, 616–623.

Wei, S., Guo, A., Chen, B., Kutschke, W., Xie, Y.-P., Zimmerman, K., Weiss, R.M., Anderson, M.E., Cheng, H., and Song, L.-S. (2010). T-tubule remodeling during transition from hypertrophy to heart failure. Circ. Res. *107*, 520–531.