Predicting phenotype using morphological cell responses to nanotopography

Marie F.A. Cutiongco[1], Bjørn S. Jensen[2], Paul M. Reynolds[1], Nikolaj Gadegaard[1,*]

[1]Divison of Biomedical Engineering, School of Engineering, University of Glasgow, Glasgow, United Kingdom

[2]School of Computing Science, University of Glasgow, Glasgow, United Kingdom

* corresponding author

Corresponding author:

Nikolaj Gadegaard, PhD

Division of Biomedical Engineering

Oakfield Avenue, Rankine Building

Glasgow, United Kingdom G12 8LT

Email: Nikolaj.Gadegaard@glasgow.ac.uk

Tel no: 01413305243

1

Abstract

1   Cells respond in complex ways to topographies, making it challenging to identify a

2   direct relationship between surface topography and cell response. A key problem is

3   the lack of informative representations of topographical parameters that translate

4   directly into biological properties. Here, we present a platform to relate the effects of

5   nanotopography on morphology to function. This platform utilizes the 'morphome', a

6   multivariate dataset containing single cell measures of focal adhesions, the

7   cytoskeleton, and chromatin. We demonstrate that nanotopography-induced

8   changes in cell phenotype are uniquely encoded by the morphome. The morphome

9   was used to create a Bayesian linear regression model that robustly predicted

10 changes in bone, cartilage, muscle and fibrous tissue gene expression induced by

11 nanotopography. Furthermore, the morphome effectively predicted nanotopography-

12 induced phenotype within a complex co-culture microenvironment. Thus, the

13 morphome enables the cell function-oriented exploration of new topographies, with

14 potential applications in the development of novel surface-patterned biomaterials for

15 tissue implants.

16

2

Introduction

Biomedical implants continue to be developed to improve patient outcomes. One way to enhance implant efficacy and tissue regeneration is to vary substrate texture with nanotopographies. Topographies at the cell-material interface are widely shown to direct cell behaviour: nanopillars change cell morphology[1]; nanogratings drastically alter lipid metabolism[2], and pluripotent[3-5] and multipotent cell differentiation[6]; nanogradients improve stem cell cardiomyogenic differentiation; and subtle changes to nanopit geometric arrangement switches human mesenchymal stem cells from multipotent to osteogenic fate[7-10]. Morphological responses to nanotopography are manifested through varying focal adhesion size[17-19], orientation[19-21], and composition[22,23] and changes in actin contractility and nuclear deformation[24].

A quantitative relationship exists between a material's physicochemical structure and its biological activity. Rational drug design has long relied on molecule solubility, ionisation and lipophilicity to predict activity[11]. Protein engineering has similarly modelled protein-peptide interactions from protein structure[12]. Cell metabolic activity correlates with synthetic polymer composition, glass transition temperature, and water contact angle[13]. Meanwhile, bacterial attachment can be predicted from descriptors of secondary ionic hydrocarbon chains[14]. In contrast to active biomolecules, the mechanotransductive effects of topography on cell response do not intuitively relate to topography length scale, isotropy, geometry, and polarity. This limits the discovery of functional topography to the screening of libraries for hits using a single, representative cell type[15-18]. Among its limitations (which include cost, inefficiency and the sampling of a small topography space), this screening approach disregards the cell specificity of response to nanotopogrpahy. Thus, it is vitally important to develop a systematic method to capture cell phenotypes induced by topography.

Here, we demonstrate a platform that encompasses cell phenotype and morphological response induced by nanotopography using image-based cellular features. We obtained measurements of focal adhesions, actin cytoskeleton and

1   chromatin, referred to here as the 'morphome', from single cells on

2   nanotopographies. We observed a unique morphological signature based on cell

3   identity and nanotopography in the morphome. We also identified changes in the

4   transcription of myogenic, osteogenic, chondrogenic and fibroblastic genes across all

5   combinations of cells and nanotopographies. We used a Bayesian linear regression

6   model to directly relate the morphological and phenotypic changes induced by

7   nanotopography. Using the morphome as predictors and without prior knowledge

8   about the nanotopography, the model robustly predicted gene expression induced by

9   nanotopography. A new morphome dataset from a co-culture of osteoblastic and

10   fibroblastic cells verified that the model positively predicts bone tissue formation

11   induced by nanotopography, even within a complex microenvironment. The lack of

12   topography-specific parameters in the model and the presence of the

13   mechanosensing machinery in all cell types highlights the broad applicability of this

14   platform to many biomaterial and cell systems, thereby supporting the function-

15   oriented exploration and rational design of new topographies.
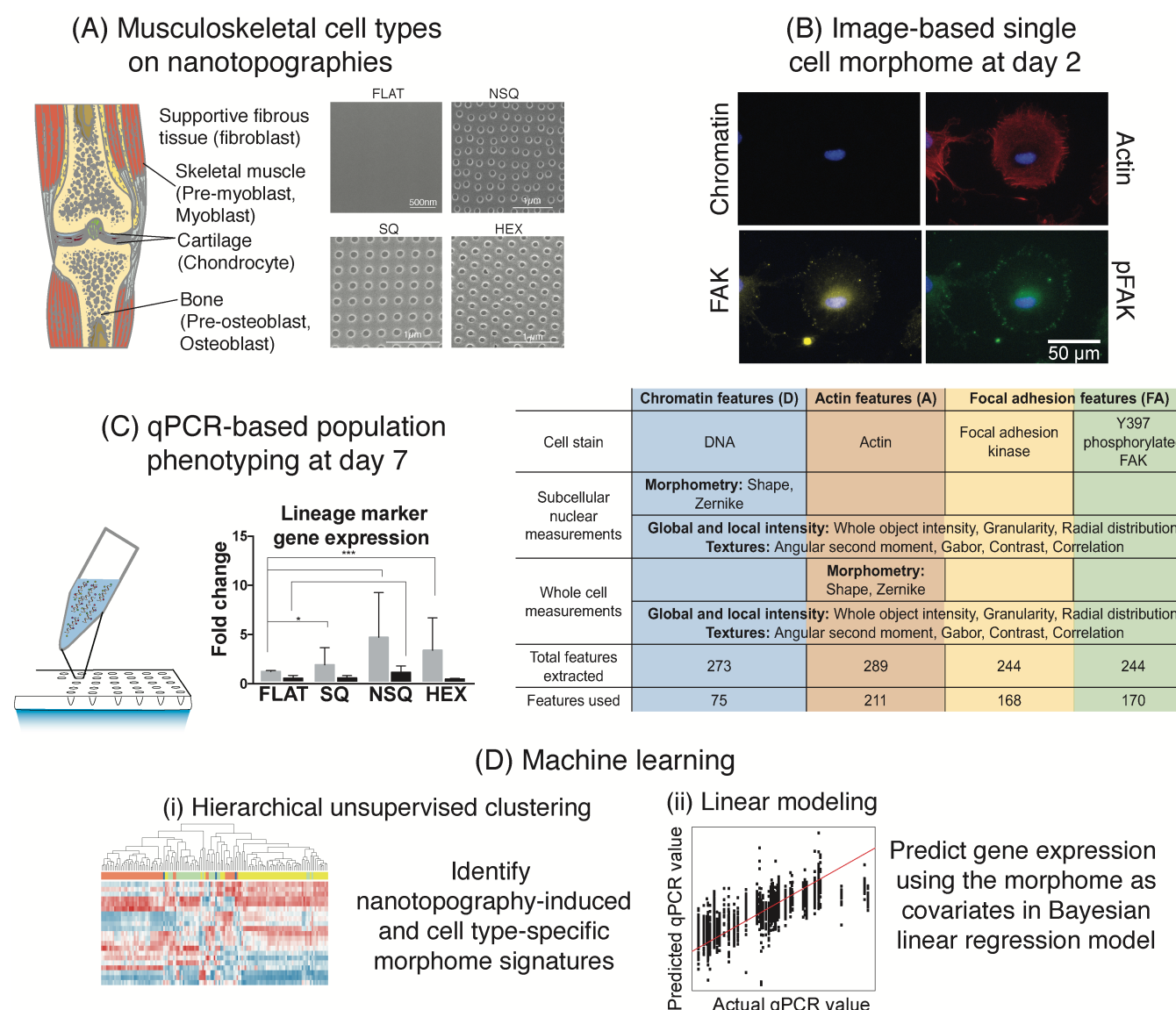
16

17   **Results**

18

19   **Measuring morphological cell responses to topography: the morphome**

20   We used nanopit topographies consisting of 120 nm diameter, 100 nm depth and

21   with a 300 nm centre-to-centre distance in a square array (SQ)[7], hexagonal array

22   (HEX)[25-27], and arranged with centre-to-centre distance offset from 300 nm by 50 nm

23   in both x and y directions (NSQ array)[7,8]. An unpatterned ('FLAT') surface was used

24   as a control (Figure 1A).

25

26   We employed cells of the musculoskeletal system due to the diverse responses of

27   muscle, bone, cartilage and fibrous cell types to nanotopographies[27-29]. Mouse

28   myoblasts, osteoblasts, chondrocytes, fibroblasts, and pre-osteoblast[30] and pre-

29   myoblast[31] progenitors, were grown on nanotopographies. Responses from

30   combinations of each cell type on all nanotopographies were measured, providing 24

31   unique combinations of cell type and nanotopography (Figure S1). Effects of

32   nanotopography on morphological characteristics such as cell and nuclear area, and

1   in pFAK activation levels, actin intensity and focal adhesion area were evident

2   (Figure S2-S4).

3

4   Morphology (shape and geometry of different compartments), texture (spatial

5   patterns of fluorescence), intensity (total fluorescence value), and radial distribution

6   of intensity (measuring radial arrangement of fluorescence), were analysed for

7   chromatin, actin, focal adhesion kinase (FAK) and phosphorylated FAK (pFAK)

8   signals at day 2 of culture (Supplementary data). The morphome consisted of 624

9   single cell measurements ("features"), with 75 chromatin and nuclear features, 211

10   actin and whole-cell features, 168 FAK features, and 170 pFAK features (Figure 1B).

11   Quantitative polymerase chain reaction (qPCR) was then used to assess changes in

12   lineage marker expression induced by nanotopography by day 7. Machine learning

13   was then applied on the morphome: (i) hierarchical clustering was used to uncover

14   unique patterns of morphological features that distinguish cell type-specific

15   responses on nanotopographies, and nanotopography-specific morphological

16   changes induced within the cell; (ii) Bayesian linear regression was used to predict

17   myogenic, osteogenic, chondrogenic and fibrogenic gene expression induced by

18   nanotopography using the morphome as predictors (Figure 1D).

5

(A) Musculoskeletal cell types on nanotopographies

Supportive fibrous tissue (fibroblast)

Skeletal muscle (Pre-myoblast, Myoblast)

Cartilage (Chondrocyte)

Bone (Pre-osteoblast, Osteoblast)

FLAT   NSQ
SQ   HEX

(B) Image-based single cell morphome at day 2

Chromatin   Actin
FAK   pFAK

50 μm

(C) qPCR-based population phenotyping at day 7

Lineage marker gene expression

Fold change

***

*

FLAT SQ NSQ HEX

	Chromatin features (D)	Actin features (A)	Focal adhesion features (FA)	
Cell stain	DNA	Actin	Focal adhesion kinase	Y397 phosphorylate FAK
Subcellular nuclear measurements	**Morphometry:** Shape, Zernike			
	Global and local intensity: Whole object intensity, Granularity, Radial distribution **Textures:** Angular second moment, Gabor, Contrast, Correlation			
Whole cell measurements		**Morphometry:** Shape, Zernike		
	Global and local intensity: Whole object intensity, Granularity, Radial distribution **Textures:** Angular second moment, Gabor, Contrast, Correlation			
Total features extracted	273	289	244	244
Features used	75	211	168	170

(D) Machine learning

(i) Hierarchical unsupervised clustering

Identify nanotopography-induced and cell type-specific morphome signatures

(ii) Linear modeling

Predicted qPCR value

Actual qPCR value

Predict gene expression using the morphome as covariates in Bayesian linear regression model

Figure 1. An overview of the morphome platform for relating cell phenotype and morphological features changed by nanotopography. (A) Osteoblastic, myoblastic, chondroblastic and fibroblastic cells were cultured on nanotopographies to obtain 24 combinations of cell type and topography.(B) At day 2, various measures of chromatin (blue), actin (red), focal adhesion kinase (FAK, yellow) and phosphorylated FAK (pFAK, green) were obtained from images of single cells to create the 'morphome'. (C) At day 7, lineage-specific gene expression induced by nanotopography was measured using population-based qPCR. (D) Machine learning for data-driven exploration and model building using the morphome. (i) Hierarchical clustering uncovered unique patterns that delineate cell type and nanotopography within the morphome. (ii) Bayesian linear regression created a predictive model that related the morphome to gene expression and phenotype.
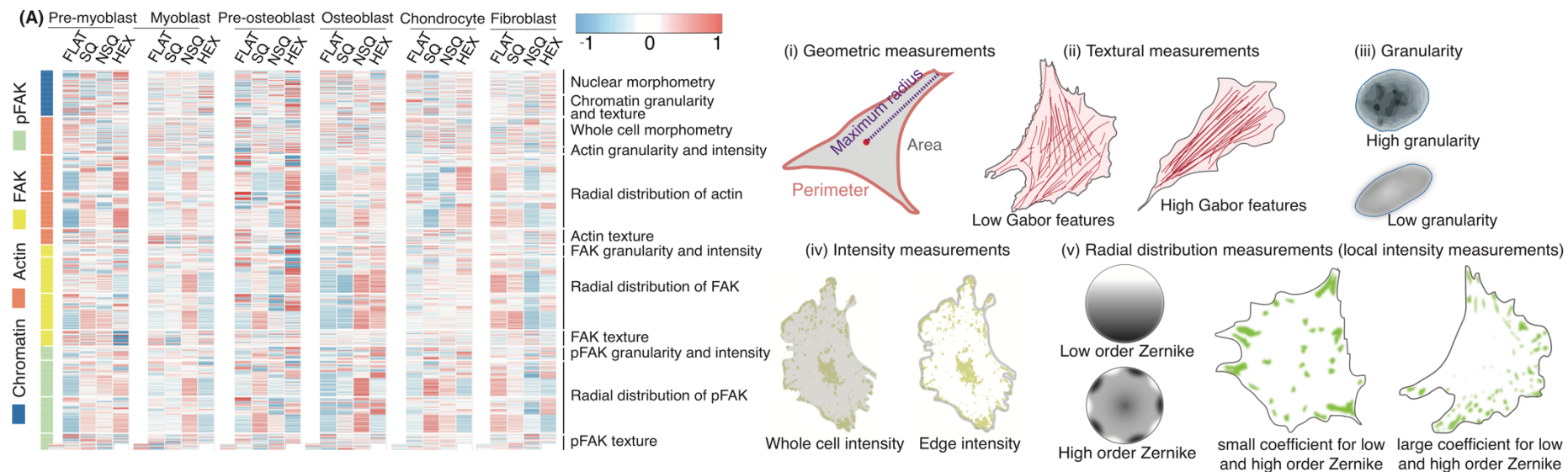
1 **The morphome depicts unique cell signatures induced by nanotopographies**

2 Patterns of nanotopography-induced morphological changes were visible from the

3 morphome (Figure 2A). Immediately apparent were large blocks of actin, FAK and

4 pFAK measurements with similar values within a cell type on a specific

5 nanotopography. These features correspond to increasingly complex measures of

6 texture, granularity and radial intensity distribution for chromatin, actin, FAK and

7 pFAK (Table S1). Frequency of pixel gray levels measures texture and thus of

8 homogeneity, with high texture values indicating coarseness. Granularity measures

9 an object's coarseness, with higher values indicating heterogeneity of pixel

10 intensities and coarser texture. The Zernike coefficient measures the spatial

11 arrangement of intensity as it resembles the increasingly complex Zernike

12 polynomials (Figure 2A). Interestingly, higher Zernike polynomials resemble the

13 punctate shape and spatial distribution of focal adhesions[32].

14

15 **Nanotopography induces cell type-specific gene expression changes**

16 Gene expression was used to determine the effect of nanotopographies on cell

17 phenotype (Figure S5). We discuss here the changes induced by nanotopography

18 on lineage-specific gene expression relevant to the cell type. Pre-myoblasts showed

19 significantly higher expression of the early lineage marker *MYOD1*, and of the late

20 markers *MYOG* and *MYH7* when cultured on SQ surfaces relative to FLAT surfaces

21 (Figure 2B-2D). Both pre-osteoblasts and osteoblasts showed increased expression

22 of early (*RUNX2*, *SP7*) and late (*BGLAP*, *SPP1*) osteogenic markers when cultured

23 on NSQ relative to FLAT (Figure 2E-2H), in line with previous studies[7,8,10,33].

24 Chondrocytes cultured on HEX showed increased expression of *COL2A1* (early

25 marker) and *ACAN* (late marker) compared to those cultured on FLAT (Figure 2I-

26 2K). Meanwhile, fibroblasts showed increased expression of pathogenic fibrosis

27 markers, *TGFB1I1*, *COL3A1* and *ELN*[34,] on all nanotopographies compared with

28 FLAT (Figure 2M-2O). In general, a cell type-specific response to nanotopography

29 was observed. Each lineage was notably enhanced in a specific cell type and

30 nanotopography combination: SQ stimulated the myoblast phenotype, NSQ

31 enhanced the osteoblast phenotype, HEX stimulated the chondrocyte phenotype,

32 while all surfaces except FLAT stimulated the fibrotic phenotype.

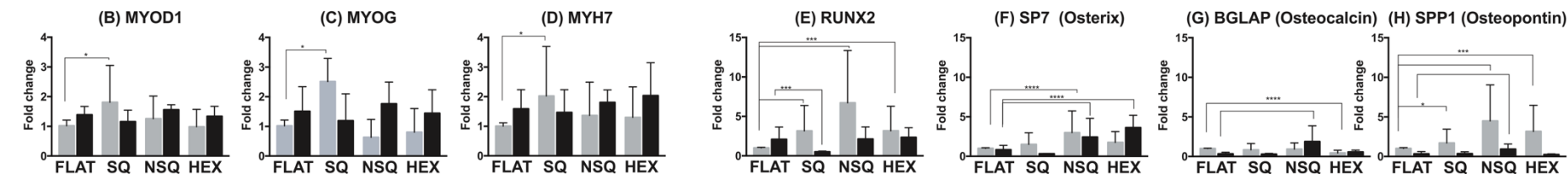**(A)** Pre-myoblast Myoblast Pre-osteoblast Osteoblast Chondrocyte Fibroblast

FLAT SQ NSQ HEX (each panel)

pFAK / FAK / Actin / Chromatin

Nuclear morphometry
Chromatin granularity and texture
Whole cell morphometry
Actin granularity and intensity
Radial distribution of actin
Actin texture
FAK granularity and intensity
Radial distribution of FAK
FAK texture
pFAK granularity and intensity
Radial distribution of pFAK
pFAK texture

(i) Geometric measurements
Maximum radius
Area
Perimeter

(ii) Textural measurements
Low Gabor features
High Gabor features

(iii) Granularity
High granularity
Low granularity

(iv) Intensity measurements
Whole cell intensity
Edge intensity

(v) Radial distribution measurements (local intensity measurements)
Low order Zernike
High order Zernike
small coefficient for low and high order Zernike
large coefficient for low and high order Zernike

Myogenic   Precursor   Differentiated   **Osteogenic**

(B) MYOD1   **(C) MYOG**   **(D) MYH7**   **(E) RUNX2**   **(F) SP7 (Osterix)**   **(G) BGLAP (Osteocalcin)**   **(H) SPP1 (Osteopontin)**

Fold change — FLAT SQ NSQ HEX

Chondrogenic   **Fibrotic**

(I) SOX9   **(J) COL2A1**   **(K) ACAN**   **(L) COL10A1**   **(M) TGFB1I1**   **(N) COL3A1**   **(O) ELN**

Fold change — FLAT SQ NSQ HEX

8
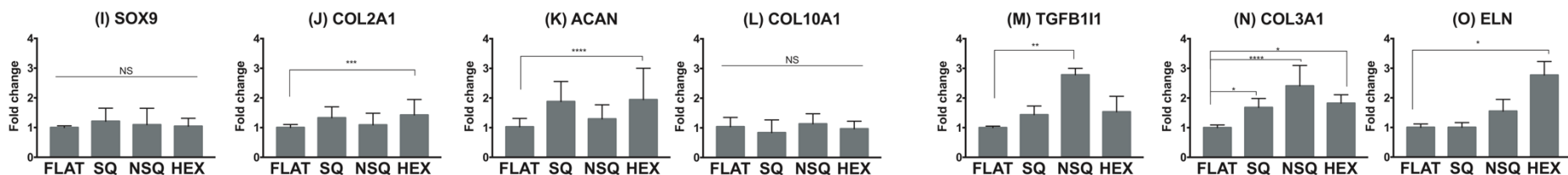
Figure 2. Unique responses of cells to nanotopography at single-cell and population levels. (A) Heat map of the morphome across cell types and nanotopographies. The morphome consisted of 624 features that quantitatively measure the cell and nucleus geometries, as well as chromatin, actin, FAK, pFAK characteristics within single cells. The colour and intensity of each tile represents the average value of the feature for a particular cell type and nanotopography combination. Each feature has an average of 0 and a standard deviation of 1 after normalization across all cell types. Schematic diagrams of representative morphome features are included in (i) to (v). (B-O) QPCR of musculoskeletal genes in response to nanotopographies. Measurement of (B-D) myogenic markers from pre-myoblasts and myoblasts, (E-H) osteogenic markers from pre-osteoblasts and osteoblasts, (I-L) chondrogenic markers from chondrocytes, (M-O) fibrotic markers from fibroblasts on nanotopographies. Gene expression is listed left to right in order of increasing maturity for the given cell lineage. QPCR measurements shown here were normalized to the reference gene and cell type on FLAT. QPCR measurements across all 24 combinations of cell type and topography are shown in Figure S3. All qPCR measurements are given as mean ± standard deviation from 2 independent experiments (n=6). Significance levels obtained from one-way ANOVA with Tukey's post-hoc test, denoted by * (p<0.05), ** (p<0.01), *** (p<0.001), and **** (p<0.0001).

1 **Distinct nanotopographical responses of single cells are reflected in the**

2 **morphome**

3

4 A subset of the morphome, consisting of 185 features, varied significantly across

5 nanotopographies (Figure 3, Supplementary Data). Hierarchical clustering revealed

6 that measurements of actin and pFAK radial distribution were most dissimilar to each

7 other (Figure S6). Notably, the morphome of cells on FLAT already show that unique

8 cell type signatures are manifested in the morphome.

9

10 Hierarchical clustering within each cell type reflected the cell type specificity of

11 nanotopography response. Here, we highlight morphome for cell type and

12 nanotopography combinations that induced highest lineage-specific gene

13 expression. When compared to FLAT, pre-myoblasts on SQ showed high average

14 values of: focal adhesion textures (Figure 3A, cluster 2); pFAK radial distribution

15 (cluster 4); nuclear morphometry; and chromatin textures (cluster 4). The morphome

16 of pre-myoblasts cultured on SQ reflects the need for FAK phosphorylation and for

17 its preferential localization at stress fiber edges, which is necessary for myotube

18 differentiation[36,37]. In contrast, myoblasts on SQ showed a particularly high average

19 value for chromatin granularity and nuclear morphometry (cluster 4), and near-zero

20 values for radial distribution of actin and of focal adhesions (cluster 1). High

21 chromatin granularity observed for both pre-myoblasts and myoblasts on SQ

22 denotes chromatin heterogeneity and condensation and transcriptional activity,

23 which is reportedly higher prior to myotube formation[38,39].

24

25 Pre-osteoblasts on SQ and NSQ had high average values for pFAK radial

26 distribution, intensity, granularity and texture, and high average values for granularity

27 of chromatin and actin (clusters 1-5). However, pre-osteoblasts on SQ had higher

28 order pFAK and FAK radial distribution than on NSQ, which induced the highest

29 expression of osteogenic markers. The morphome of pre-osteoblasts grown on NSQ

30 featured radially variable actin that resemble bone cells, which have high contractility

31 and actin stress fibers[40].

32

1   For osteoblasts, the differences between the SQ and NSQ morphome were more

2   prominent: NSQ induced lower average values of focal adhesion granularity,

3   chromatin texture and nuclear morphometry (cluster 1), and higher average values

4   for focal adhesion radial distribution (clusters 4-7) compared to SQ (Figure 3D). The

5   osteoblast morphome on NSQ indicates that focal adhesions localize at regular

6   intervals along the periphery, which is associated with osteogenesis[41]. Furthermore,

7   changes in nuclear morphometry attributed to spreading after growth on stiff

8   surfaces is also associated with osteogenic differentiation[42].

9

10   Chondrocytes on HEX, which significantly increased chondrogenic marker gene

11   expression relative to FLAT, showed high average values of radial distribution,

12   texture and granularity of actin and FAK, high average nuclear morphometry, and

13   low average values of pFAK and chromatin measurements (Figure 3E, clusters 1-3).

14   These characteristics reflect the morphological changes (including reduced

15   contractility and stress fiber formation, increased cell circularity, and decreased cell

16   spreading[40], low FAK phosphorylation[43] and poor focal adhesion formation[44]) of stem

17   cells undergoing chondrogenesis.

18

19   The morphome of fibroblasts cultured on FLAT had high average values of both actin

20   and focal adhesion measurements (Figure 3F, clusters 3 and 5). The highly uniform

21   radial arrangement of focal adhesions and actin of cells on FLAT indicate reduced

22   polarization and contractile morphology of fibroblasts activated to a fibrotic state[45].

23   Inflammation pathways are reportedly increased in fibroblasts on HEX[46], inducing

24   low adhesion that is reflected in low actin and focal adhesion radial distribution

25   (clusters 3-9). Fibroblasts grown on NSQ and HEX showed low average values of

26   focal adhesion and actin radial distribution but high values when grown on SQ

27   (cluster 2-5).

28

29   Overall, the morphome reflected cell-type specific responses to nanotopography.

30   The morphome also exhibited patterns from different cell types (e.g. pre-osteoblasts

31   vs osteoblasts) despite similarity in origin (Table S2). A multi-class logistic regression

32   classifier was able to accurately distinguish 6 different cell types using the

33   morphome (Figure S7). The radial arrangement of actin and focal adhesions were a

11

1   critical distinction of the musculoskeletal cell types induced by nanotopography,

2   while the arrangement of actin fibers into stress fibers or into cortical, circular

3   bundles provided information on various cell states.

4

5   We also clustered the morphome based on cell type (Figure S8-S9, Supplementary

6   data). Patterns emerge in the morphome in direct response to nanotopography: NSQ

7   induced high average values of pFAK radial distribution, texture and granularity; and

8   HEX induced high average values of actin radial distribution. Correlation between the

9   dendrograms confirm that the morphome clusters of different nanotopographies were

10 dissimilar to each other (Table S3). The morphome enabled discrimination between

11 4 nanotopographies with 68% overall accuracy (87% classification rate for HEX)

12 using multi-class logistic regression (Figure S10).
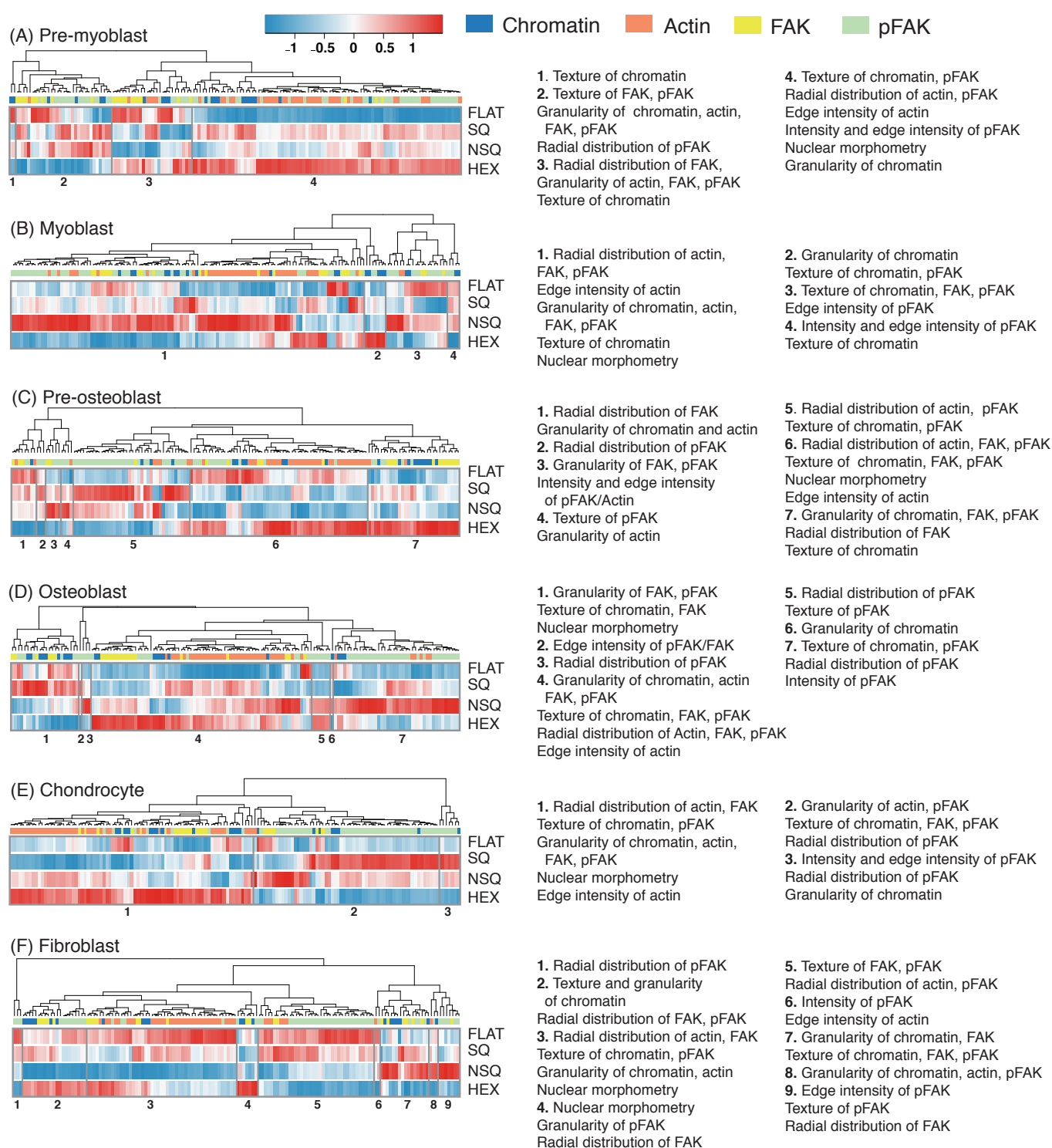
Figure 3. Nanotopography induced changes in single cells are evident in their morphomes. (A-F) Cell type-specific hierarchical clustering of the morphome averaged for each cell type. The total number of cells analyzed from two independent experiments are: (A) n=877 pre-myoblasts; (B) n=931 myoblasts; (C) n=644 pre-osteoblasts; (D) n=728 osteoblasts; (E) n=619 chondrocytes; and (F) n=1140 fibroblasts. The colour and intensity of each tile represents the average value of the feature. Each feature has an average of 0 and a standard deviation of 1 by normalizing within each cell type. The nanotopography specific morphome is made up of 21 chromatin, 42 actin, 23 FAK and 60 pFAK features. Features included were changed significantly across topographies (p < 0.05 using one-way ANOVA).

1   **The morphome predicts nanotopography induced gene expression changes**

2   The Spearman rank correlation revealed that varying degrees of correlation exist

3   between morphome features and gene expression (Figure S11). We hypothesized

4   that the features of the morphomes would sufficiently encompass cell phenotypes

5   induced by nanotopography thus allowing prediction. We utilized Bayesian linear

6   regression to predict gene expression using the morphome features as predictors

7   (for the explicit model definition, see Materials and Methods). A Bayesian linear

8   regression model reflects uncertainty in the estimation of regression weights

9   compared to point value estimates using maximum likelihood regression. Gene

10 expression was modelled independently of each other, thereby creating 14 different

11 equations with variable weighting of the morphome features. Importantly, the model

12 was trained without any prior knowledge of topography type, instead relying on the

13 morphome to encode for this information.

14

15 The morphome clearly captured gene expression changes induced by

16 nanotopography (Figure 4A). The heterogeneity inherent in single cells, usually

17 uncaptured by population measurements of gene expression, are apparent in the

18 variance of the predictions using the morphome and the model. The mean absolute

19 error (MAE) for prediction of all genes was between 10% (for prediction of *MYOD1*,

20 *MYOG* and *MYH7*) and 21% (for prediction of *COL3A1*, Table S4). Features with the

21 highest contribution in predicting the 14 different genes were predominantly

22 measurements of FAK texture and radial distribution, actin texture, and chromatin

23 granularity (Figure S12, Supplementary data). pFAK activation, as indicated by

24 pFAK/FAK intensity ratio, consistently contributed to the prediction of all 14 genes.

25 pFAK was particularly important to the model due to its relevance in contractility

26 induced by nanotopography[19], fibrosis and scar tissue formation[48], *in vitro*

27 osteogenesis[49,50], and chondrogenic maintenance[43]. Through the variability in the

28 model, we obtain realistic expectations of single cell predictions from population level

29 measurements

30

31 The sensitivity and predictive power of the morphome was verified by iteratively

32 training a model without one combination of cell type and topography (Figure 4E).

33 Drastic increases in MAE were observed when lineage-specific genes were

1   predicted from models that excluded the particular cell type lineage being tested,
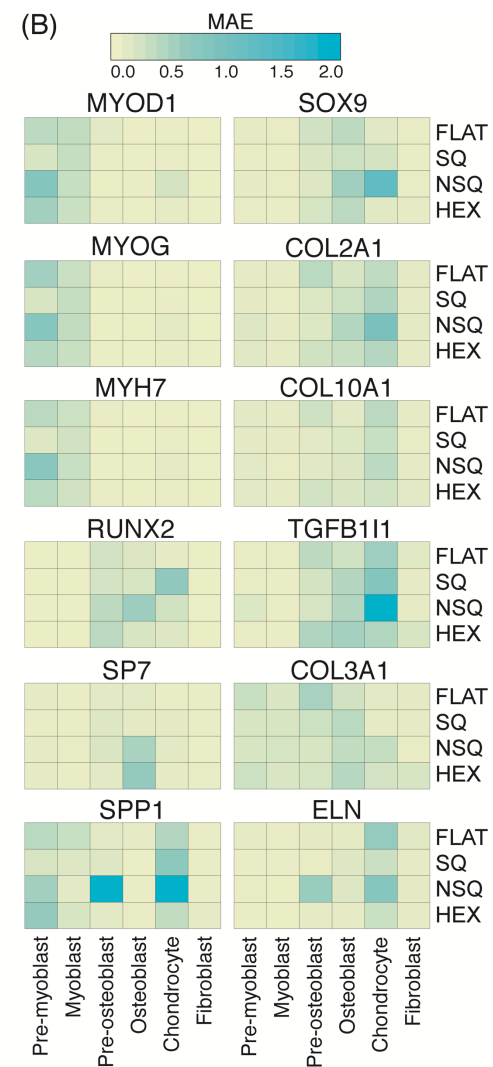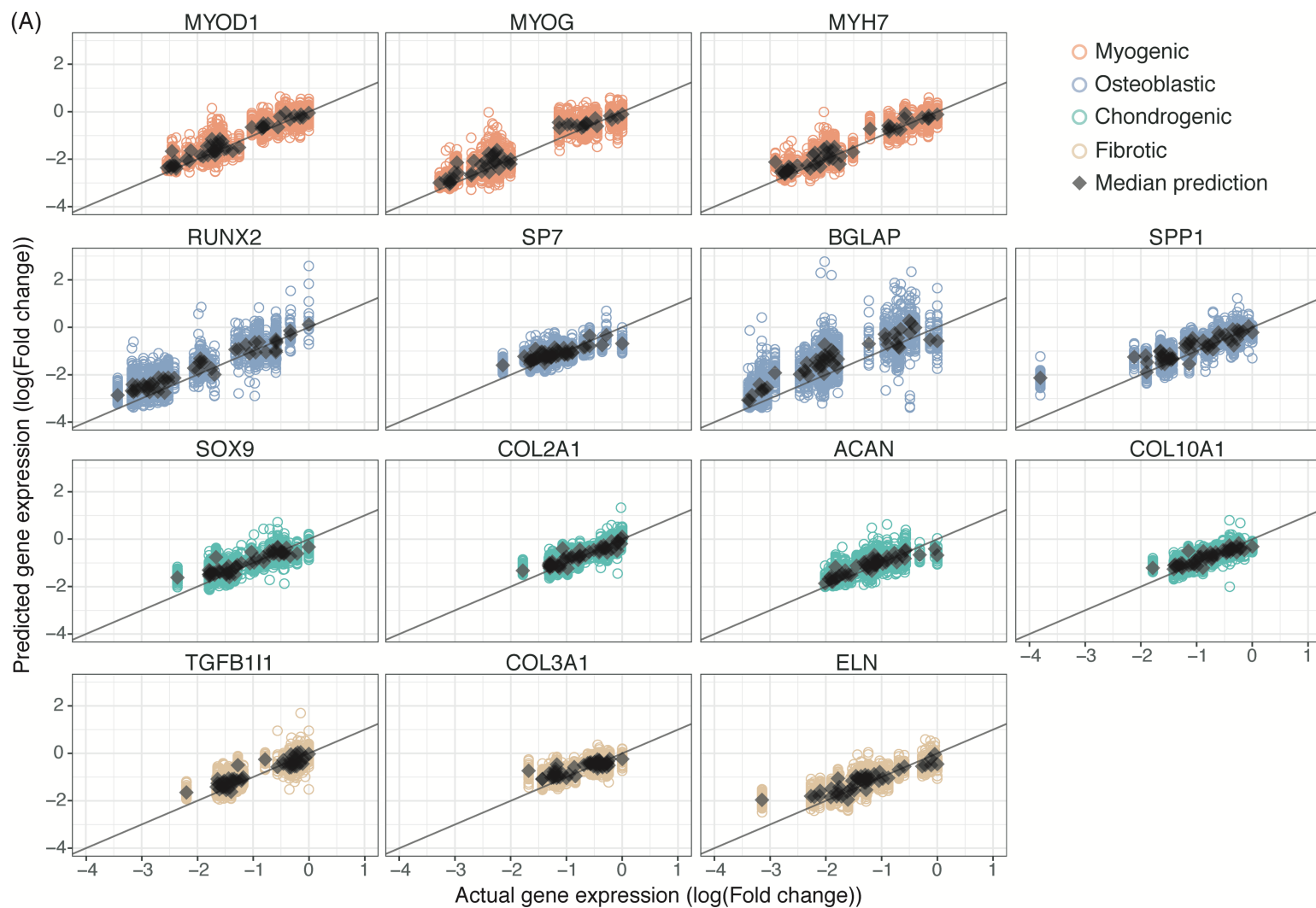
2   regardless of nanotopography.

15

Figure 4. The morphome reliably predicts nanotopography-induced cell phenotype and gene expression from specific cell type and topography combinations. (A) The morphome was used to train a Bayesian linear regression model that predicted myogenic, osteogenic, chondrogenic, and fibrotic gene expression. The expression of each gene (response) was trained against linear combinations of morphome features (predictors) and without any prior knowledge on topography parameters. The linear regression model was trained using 60% of the dataset and tested using 40% of the data. Scatterplots show actual and predicted gene expression values by using the test set as input to the model. Mean absolute error (MAE) was obtained by averaging the differences between actual and predicted qPCR values. (B) Testing the predictive power of the morphome by leave-one-out validation. To test the predictive power and bias of the morphome, the linear regression model was retrained after exclusion of one combination of cell type and topography. The excluded cell type and topography dataset was used for prediction, from which MAE was calculated. The tile position denotes the cell type and nanotopography combination that was excluded in the model and used for testing, while the color of each tile denotes the MAE.

1

1 **The morphome encompasses cell response to topography within a complex**

2 **environment**

3 We demonstrate the application of the linear regression model by predicting the

4 outcome of pre-osteoblasts and fibroblasts co-cultured on nanotopographies. A new

5 morphome were obtained from all cells on the entire nanotopography (Figure S13).

6 This co-culture morphome was then used in the linear regression model to predict

7 gene expression (Figure S14).

8

9 For visualization, the sum of predicted osteogenic (*RUNX2, SP7, BGLAP* and *SPP1*)

10 and fibrotic (*TGFB1I1, COL3* and *ELN*) genes was plotted against the spatial

11 coordinates of the pre-osteoblasts and fibroblasts. Osteogenic gene expression was

12 highest on NSQ, which also induced concentrated areas of highly expressing cells

13 (Figure 5A). These areas might represent hotspots of osteogenic paracrine signaling

14 induced by the NSQ nanotopography[9]. In contrast, osteogenic gene expression was

15 low and homogenous on the FLAT, SQ and HEX topographies.

16

17 Fibrotic gene expression showed more spatial variability across the

18 nanotopographies but was also maximized on the NSQ nanotopography, and largely

19 overlaps with the spatial pattern of osteogenic gene expression (Figure 5B). This is

20 attributable to the synergistic interaction of osteoblasts and fibroblasts on osteogenic

21 differentiation and mineralization[51]. The predicted effect of high osteogenic gene

22 expression induced by NSQ was verified in the increased mineralization and bone

23 tissue formation compared to FLAT at 28 days (Figure S15). Our results

24 demonstrate the capability of the morphome-based platform to capture topography

25 induced responses within complex microenvironments reminiscent of *in vivo* settings.

26

27 The morphome and its application in the linear regression model enables the

28 simultaneous assessment of morphological and functional effects of nanotopography

29 at the single cell level. For instance, nanotopographies show no discernible effects

30 on nuclear perimeter or on the ratio of pFAK/FAK intensity (Figure 5C). Yet there

31 was a clear gradient in predicted osteogenic gene expression: as the nucleus

32 reduced in size and pFAK activation rises, osteogenic gene expression increased.

33 This gradient was less clear with fibrotic gene expression. However, the

1   nanotopographies have discernably different effects on actin and pFAK radial

2   distribution, with cells on SQ showing the lowest values (Figure 5D). These

3   nanotopography-specific cell clusters also revealed low osteogenic gene expression

4   yet high fibrotic gene expression. The morphome clearly enables new insight into

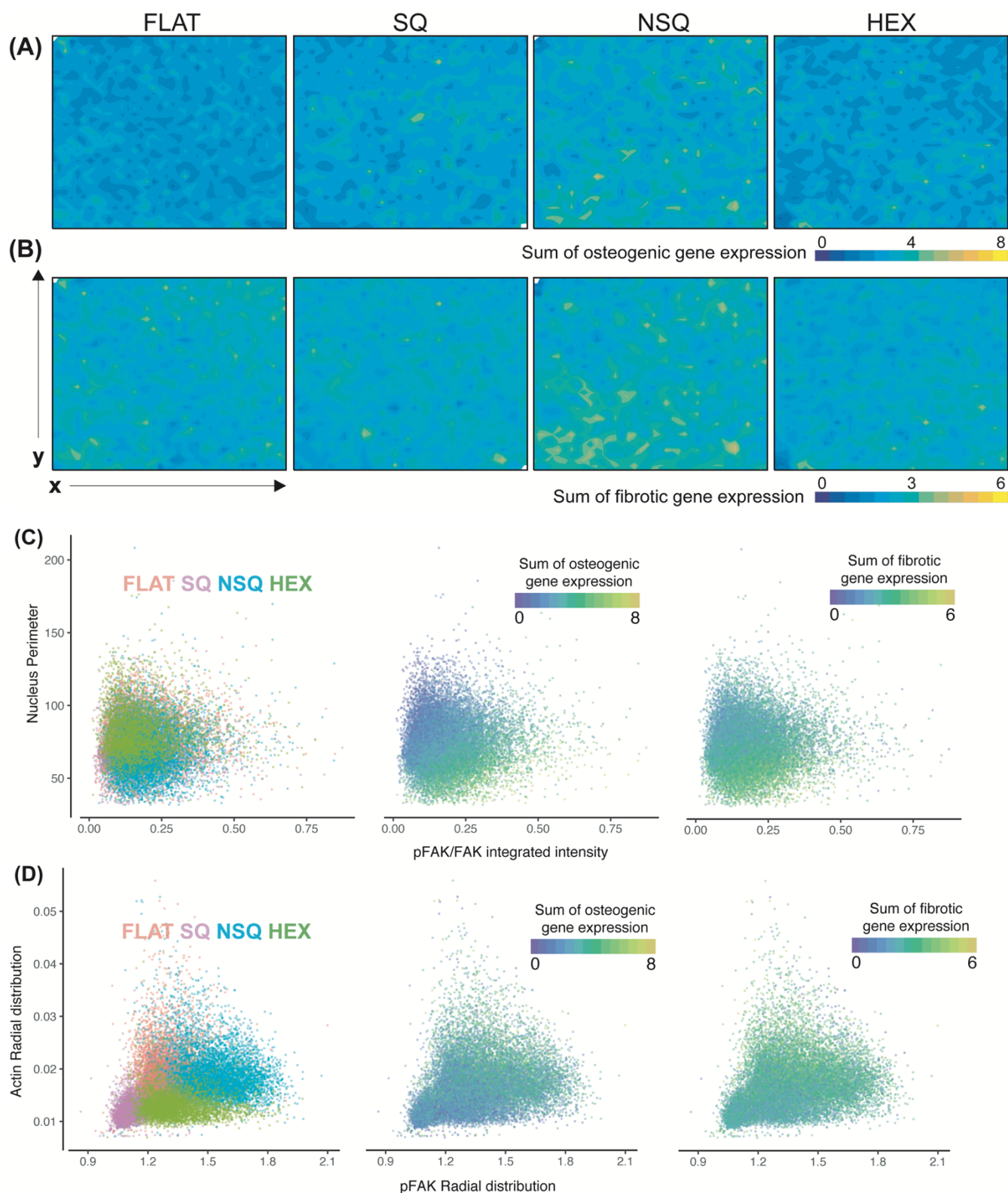5   single cell responses induced by nanotopography.

Figure 5. Morphological and functional response of single cells induced by nanotopography.
(A-B) Predicted functional response of a pre-osteoblast and fibroblast co-culture to nanotopography. Contour plots show the sum of predicted (A) osteogenic (RUNX2, SP7, BGLAP, SPP1) and (B) fibrotic (TGFB1I1, COL3, ELN) gene expression for individual cells on

FLAT, SQ, NSQ and HEX topographies. Pre-osteoblast and fibroblast cells were co-cultured on FLAT, SQ, NSQ and HEX nanotopographies for 2 days, and their morphome obtained from the entire grid. This new morphome was then used as input in the linear regression model (shown in Figure 4) to predict gene expression. A 2.12 mm x 2.12 mm grid was imaged for each nanotopography. X and Y axes of each contour plot shows are spatial coordinates on the nanotopogrpahy substrate. (C-D) Morphological and functional information at the single-cell level is provided by the morphome. Each dot in the plot denotes a single cell. The morphome enables simulataneous analysis of morphological and functional changes induced by nanotopography at the single cell. N ≥ 5000 cells per topography.

1 **Discussion**

2

3 In this study, we present a system that robustly encodes nanotopography

4 parameters into functional cell specific output. The morphome, which is the collective

5 morphological measurements of chromatin, actin and focal adhesions within single

6 cells, were found to manifest nanotopography-induced changes in cell morphology

7 and cell function. The information encoded in the morphome underpinned the

8 performance of a Bayesian linear regression model for predicting gene expression.

9

10 Focal adhesion complexes are the primary mechanosensing machinery of the cell

11 that respond to the biophysical microenvironment. Focal adhesion assembly or

12 confinement induced by nanotopography allows spontaneous actin cytoskeleton

13 assembly, nucleus deformation, and cell fate determination[24,52,53]. Cells also change

14 and control their microenvironment, thereby dictating focal adhesion and cellular

15 characteristics to varying degrees[54,55]. Focal adhesion growth is particularly

16 significant in sensing nanoscale changes in ligand arrangement, as this supports

17 force redistribution throughout the cell [56]. This reciprocity between the

18 nanotopographical microenvironment, focal adhesions, cytoskeleton and chromatin,

19 and cell function underline the platform presented in this study.

20

21 Our results demonstrate that quantitative measures of high-level cell response to

22 nanotopography emerge from the morphome. The morphome-based platform

23 reported here offers two unique advantages. First, the design parameters of the

24 nanotopographies were explicitly excluded in predicting cell function. To use

25 structural information from topography (e.g. depth, diameter, pitch, geometry,

1   isotropy, roughness) considerably limits the dataset to only a handful of descriptors.

2   As opposed to polymer-based biomaterials that contain easily interpretable

3   physicochemical properties derived from chemical structure, the effects of

4   topography parameters on hydrophobicity, serum adsorption and cell attachment are

5   less well understood and quantified. Instead, the information contained in the

6   morphome encompass both topography and cell function. Furthermore, by creating a

7   general model that is independent of topographical parameters, this approach can

8   be easily applied to predict cell functions induced by new nanotopographies.

9   Second, in this approach, gene expression-based measures of cell phenotype

10   determined the specialized function of cells. In contrast to our system, many

11   quantitative structure and function relationship studies of polymeric biomaterials

12   have focused only on unspecialized cellular behaviours, such as adhesion and

13   metabolic activity[13,14]. Since gene expression is highly scalable, the linear regression

14   model can be adapted to predict other cell behaviours or phenotypes. This also

15   ensures that the morphome supports a function-focused exploration of the

16   topography space and the rational design of topographies, as opposed to the

17   currently used trial-and-error screening approach.

18

19   Our data reveals that the morphome can also manifest information on the stage

20   within the cellular lineage commitment timeline. The morphome easily captured the

21   higher levels of mature gene expression in cells farther along lineage commitment

22   compared to precursor cells. Additionally, our co-culture experiments robustly

23   predicted the osteogenic properties induced by NSQ. The elevated level of predicted

24   osteogenic gene expression was supported by high mineralization on NSQ after 28

25   days of culture. Our results suggest that the morphome can manifest cellular

26   changes induced by nanotopography, as well as changes driven by chemical or

27   paracrine cues. This property of the morphome can be exploited to predict cell

28   behavior in settings that recapitulate in vivo-like environments with numerous

29   microenvironmental signals.

30

31   The co-culture experiment also demonstrated that the morphome dataset

32   encompasses, at high resolution, structural, functional and spatial information.

33   Structural information is crucial for determining how cells look, a highly valuable

22

1   feature that can also be used to study drug- and disease-induced cell perturbation.

2   The functional information – represented by changes in the expression of 14 different

3   genes – can be observed at single-cell level, reflecting spatial resolution that is

4   invaluable in high-throughput formats. When developed as an automated platform

5   that executes topography production, cell seeding and staining, image acquisition,

6   morphome extraction and gene expression prediction, this approach could increase

7   screening efficiency and make it more cell function-oriented.

8

9   Clearly, morphome capture is crucial to the ability of the linear regression model to

10   predict nanotopography induced cell function. While population-level measures of

11   gene expression strongly indicate cell function, they introduce a measure of

12   uncertainty and biological variability into the linear regression model. Thus, a one-to-

13   one relationship between the morphome and cell function is essential to develop.

14   Non-destructive microscopic and molecular tools[57] that combine spatial and

15   structural information from the morphome with single-cell functional assays are vitally

16   important for establishing quantitative topography structure- and cell-function

17   relationships using the morphome. However, the use of routine methods, such as

18   high-content imaging and qPCR, permits any lab to measure the morphome and to

19   model it against the cell function in question.

20

21   By generating a multivariate morphome dataset and combining it with machine

22   learning, we have created a powerful platform for relating topography structure to cell

23   function. The predictive power of the Bayesian linear regression model we have

24   developed easily lends to sequential experimental design by exploiting the

25   uncertainty and variability within the model[1,2]. This platform can ultimately be used to

26   guide the testing and exploration of new topographies.

27

28   **Materials and methods**

29

30   *Polycarbonate surfaces with nanotopography*

31   Surfaces patterned with 120 nm diameter and 100 nm depth nanopits were

32   fabricated on polycarbonate using injection molding[60]. The following

33   nanotopographies were used: surfaces without nanopits (FLAT); nanopits in a

1   square array with 300 nm center-to-center spacing (SQ); nanopits in a square array

2   with approximately 300 nm center-to-center spacing distorted by 50 nm in both x and

3   y directions (NSQ); nanopits in a hexagonal array with 300 nm center-to-center

4   spacing (HEX). Samples were cleaned in 70% ethanol and dried before treating with

5   O_2 plasma at 120 W for 1.5 mins. Samples were sterilized using UV light in biological

6   safety cabinet for at least 20 mins before cell seeding.

7

8   *Cell culture*

9   Mouse fibroblast cell line NIH3T3 (ATCC) was cultured in reduced sodium

10   bicarbonate content (1.5 g/L) Dulbecco's modified eagle's medium with (DMEM)

11   supplemented with L-glutamate (2mM), 10% bovine calf serum, and 1% penicillin-

12   streptomycin. Mouse C2C12 myoblasts (ATCC) were cultured in DMEM with 20%

13   FBS and 1% penicillin-streptomycin, and differentiated into skeletal muscle using

14   DMEM supplemented with 2% horse serum and 1% penicillin streptomycin. Mouse

15   primary chondrocytes were cultured in minimum essential medium alpha (MEMα)

16   with nucleosides, ascorbic acid, glutamate, sodium pyruvate supplemented with 10%

17   FBS and 1% penicillin-streptomycin. Mouse MC3T3 cells (ATCC) were cultured in

18   MEMα with nucleosides and L-glutamine without ascorbic acid and supplemented

19   with 10% FBS and 1% penicillin-streptomycin. To differentiate MC3T3 into

20   osteoblasts, MC3T3 media was supplemented with 10 nM dexamethasone, 50 µg/ml

21   ascorbic acid and 10 mM beta-glycerophosphate[61]. Lineage committed progenitor

22   cells, referred to here as pre-osteoblasts and pre-myoblasts, were also included in

23   the study to mimic the osteogenic and myogenic regeneration in the adult tissue[30,31].

24

25   *Cell seeding*

26   Cells were harvested from flasks using trypsin in versene buffer and spun down at

27   400 x g for 5 minutes. NIH3T3 and MC3T3 cells were resuspended in complete

28   media and seeded at 4000 cells/cm^2. Chondrocytes and C2C12 were seeded at

29   2500 cells/cm^2. Cells were seeded at different densities to ensure single cells at

30   approximately 30% confluency on each surface after 2 days culture. To ensure

31   homogeneity of seeding, cells were seeded using a cell seeder that controlled fluid

32   flow. For co-culture studies, MC3T3 and NIH3T3 cells were simultaneously seeded

1 at 2000 cells/cm^2 per cell type in MC3T3 growth media. All cells were grown for 2

2 days on nanotopographies before fixation for immunofluorescence staining.

3

4 *Quantitative polymerase chain reaction (qPCR)*

5 After 6 days, total RNA was obtained from lysed cells according to manufacturer's

6 instructions (Promega ReliaPrep Cell Miniprep kit). Gene expression was measured

7 directly from 5 ng RNA using a one-step RT PCR kit with SYBR dye (PrimerDesign).

8 qPCR was run on the BioRad CFX96 platform. Relative gene expression was

9 normalized to the 18S ribosomal RNA reference gene. A list of the forward and

10 reverse primers used to study different mouse genes is given in Table S5. Bar charts

11 for qPCR data were obtained using GraphPad Prism (v7.0). One-way ANOVA with

12 Tukey's post hoc test for multiple comparisons was performed to determine the

13 effect of nanotopography on gene expression compared with FLAT. Statistical

14 significance was considered at $p < 0.05$.

15

16 *Immunofluorescence staining*

17 After 2 days, cells on surfaces were fixed with 4% paraformaldehyde solution in

18 phosphate buffered saline (PBS) at 4°C for 15 minutes. Fixed cells were then

19 permeabilized and blocked with 10% goat serum and 2% bovine serum albumin in

20 PBS for 1 hour at room temperature. Surfaces were stained with the following

21 primary antibodies overnight at 4°C: pFAK Y397 (Abcam 39967, 1:400) and FAK

22 (ThermoScientific 396500, 1:400). Afterwards, Alexa Fluor-conjugated secondary

23 antibodies (ThermoScientific, 1:500) against the host species of the primary antibody

24 were used. Alexa Fluor 549 conjugated phalloidin (ThermoScientific, 1:200) were

25 used to structuralize the actin cytoskeleton. Cells were also stained with DAPI

26 (ThermoScientific) to structuralize nuclei. All surfaces were mounted on 0.17 μm

27 thick Glass coverslips with ProLong mounting medium (ThermoScientific) and dried

28 overnight at 4°C before imaging.

29

30 *Image acquisition and feature extraction*

31 For single-cell population studies, monochrome images of each fluorophore were

32 obtained at 40X magnification (numeric aperture 1.3) using the EVOS FL1 System

33 (ThermoScientific). For co-culture studies, the entire field per nanotopography was

1   imaged and stitched through an automated microscope (EVOS FL2 Auto) with a 40X

2   magnification. All images from one cell type were obtained using the same camera

3   and light settings. Afterwards, images were analysed using CellProfiler (v2.4.0, The

4   Broad Institute). Image processing, including illumination correction and channel

5   alignment, was performed across each independent experiment and each cell type[62].

6   Nucleus and cell body were segmented from each cell in each image. Shape or

7   morphometric measurements, total and local intensities and textures of chromatin,

8   actin, pFAK and FAK were obtained using built-in object measurement modules in

9   CellProfiler[63].

10

11 *Machine learning and multivariate analysis*

12 The morphome initially consisted of a total of 1050 measurements obtained from

13 single cells. Morphome measurements from single-cell populations were combined

14 and were scaled by subtracting the mean and normalizing by the standard deviation

15 of the entire data set. Morphome data from co-culture studies were similarly scaled

16 and normalized using the mean and standard deviation from the initial dataset

17 consisting of homogeneous cell populations. Prior to machine learning, features with

18 zero variance within each batch (e.g. Zernike Phase measurements) were removed

19 from the data set. A Pearson correlation method at significance level 90% was used

20 to remove features with correlation higher than 0.9 without significantly reducing total

21 data variance (Figure S16) using the *KMDA* package[64]. After preprocessing, 624

22 morphome features were used further in the study. To determine the features that

23 were significantly varied either across nanotopography or cell type, a one-way

24 ANOVA was performed (KNIME v3.3.0). The intersection of nanotopography-specific

25 and cell type-specific features were obtained. Hierarchical clustering was performed

26 using Spearman correlation as a distance metric and an average linkage method for

27 cluster linkage using *gplots* package[65]. Cluster membership and stability was

28 obtained from silhouette analysis using the *cluster* package[66]. Dendrogram

29 correlation was performed using *corrplot* package[67].

30

31 *Bayesian linear log-Normal regression*

32 Only morphome features with an absolute Spearman correlation coefficient ≥ 0.7

33 against all examined expression markers were used in the linear regression model

26

1   (Supplementary data). The linear regression model used 243 morphome features,

2   containing 22 nuclear morphometry and chromatin, 71 actin, 75 FAK and 75 pFAK

3   measurements, as predictors of the model. Each replicate of the qPCR data was

4   propagated across each replicate of the single-cell morphome data. For each gene

5   analysed, data were rescaled from 0 to 1 by normalizing to the maximum gene

6   expression.

7

8   Linear regression was performed as a simple approximation of the relationship

9   between the morphome and myogenic, osteogenic, chondrogenic and fibrotic gene

10   expression. Established Bayesian inference methods were used to determine the

11   probabilities of observing gene expression with a given morphome set. We consider

12   a linear model where expression of one gene (response y) was predicted through a

13   linear combination of the morphome features (predictors X) transformed by the

14   inverse identity link function. We assume that y follows a log Normal distribution

15   parametrized by the mean μ and variance σ_i^2:

16

17
$$y_i \sim Lognormal(\mu_i, \sigma_i^2)$$

18

19   And that μ is a linear function of X parametrized by β:

20
$$\mu_i = \alpha + \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n$$

21

22   All model parameters β were assumed *a priori* to come from a normal distribution,

23   parametrized by mean and standard deviation:

24

25
$$\beta \sim Normal(0, 2)$$

26

27   Each gene was trained independently resulting in 14 different linear regression

28   equations. A 60%-40% training and test split for Bayesian linear regression was

29   performed randomly and with stratification using the *caret* package for R[68]. The

30   Bayesian linear model was created using the *brms* package for R[69], which utilizes

31   the Hamiltonian Markov Chain Monte Carlo sampler for estimation of the posterior

32   distribution of β. Bayesian linear modelling was carried out using with 1000 warm-up

1   iterations and 1000 sampling iterations within each chain for 3 independent chains.

2   All models were confirmed to converge to the equilibrium distribution by confirming

3   potential scale reduction statistic split-$\hat{R} \geq 1$, effective sample size was smaller than

4   total sample size and low autocorrelation. Predicted gene expression was performed

5   by using the test set or the morphome obtained from the co-culture study as input to

6   the linear model. Predicted values were averaged across 50 draws from the

7   posterior distribution.

8

9   To determine the predictive power of the morphome, a specific combination of cell

10   type, topography and replicate were iteratively omitted and the remaining dataset

11   was used to refit new models. Thus, 576 additional models were created to test 24

12   different cell type combinations across 12 genes and 2 replicates. The predictive

13   quality of the models was assessed by predicting the expression of all 14 genes from

14   the omitted cell type, topography and replicate dataset. We report the mean absolute

15   error (MAE) of qPCR prediction for each cell type and topography combination

16   averaged across 2 replicates. MAE was calculated as the average across all

17   absolute differences between predicted and actual gene expression.

18

19   *Statistics, visualization and software*

20   Statistical analysis and machine learning were performed using statistical software R

21   (v3.4.3) and its graphical interface RStudio (v1.0)[70]. Scatterplots, boxplots and

22   histograms were generated using *ggplot2* in R[71]. Pearson correlation, Spearman

23   rank correlation and Bayesian linear regression were performed in R (v3.4).

24   Interpolation of x and y coordinates and sum of predicted osteogenic or fibrotic gene

25   expression was performed using bivariate interpolation of a regularly gridded dataset

26   using *akima* package[72]. Afterwards, the contour plots were created using the

27   *fitted.contour* function in the base package of R, with the nuclear centroid position

28   used as spatial coordinates of the cell. Barcharts and one-way ANOVA analysis of

29   qPCR values were obtained using GraphPad Prism (v7.0a).

30

31   **Acknowledgements**

8

9   **Author contributions**

10 MFAC, PMR and NG designed the biological experiments. MFAC and BSJ designed

11 the machine learning analysis. MFAC carried out imaging, image analysis, qPCR,

12 machine learning. PMR fabricated and characterized the nanotopographical

13 surfaces. MFAC, BSJ, PMR and NG wrote the manuscript. All authors have read and

14 approved the manuscript before submission.

15

16 **Disclosure of competing financial interests**

17 The authors have no competing financial interests to disclose.

18

19 **References**

20

21 1.   Reynolds, P. M., Pedersen, R. H., Riehle, M. O. & Gadegaard, N. A Dual
22   Gradient Assay for the Parametric Analysis of Cell-Surface Interactions.
23   *Small* **8,** 2541–2547 (2012).
24 2.   Di Rienzo, C. *et al.* Unveiling LOX-1 receptor interplay with
25   nanotopography: mechanotransduction and atherosclerosis onset. *Sci.*
26   *Rep.* **3,** 1–8 (2013).
27 3.   Chan, L. Y., Birch, W. R., Yim, E. K. F. & Choo, A. B. H. Temporal
28   application of topography to increase the rate of neural differentiation
29   from human pluripotent stem cells. *Biomaterials* **34,** 382–392 (2013).
30 4.   Tan, K. K. B. *et al.* Enhanced differentiation of neural progenitor cells into
31   neurons of the mesencephalic dopaminergic subtype on topographical
32   patterns. *Biomaterials* **43,** 32–43 (2015).
33 5.   Rasmussen, C. H. *et al.* Enhanced Differentiation of Human Embryonic
34   Stem Cells Toward Definitive Endoderm on Ultrahigh Aspect Ratio
35   Nanopillars. *Adv. Funct. Mater.* **26,** 815–823
36 6.   Yim, E. K. F., Pang, S. W. & Leong, K. W. Synthetic nanostructures
37   inducing differentiation of human mesenchymal stem cells into neuronal
38   lineage. *Exp. Cell Res.* **313,** 1820–1829 (2007).

7. McMurray, R. J. *et al.* Nanoscale surfaces for the long-term maintenance of mesenchymal stem cell phenotype and multipotency. *Nature Materials* **10,** 637–644 (2011).

8. Dalby, M. J. *et al.* The control of human mesenchymal cell differentiation using nanoscale symmetry and disorder. *Nature Materials* **6,** 997–1003 (2007).

9. Tsimbouri, P. M. *et al.* Using Nanotopography and Metabolomics to Identify Biochemical Effectors of Multipotency. *ACS Nano* **6,** 10239–10249 (2012).

10. Yang, J. *et al.* Nanotopographical Induction of Osteogenesis through Adhesion, Bone Morphogenic Protein Cosignaling, and Regulation of MicroRNAs. *ACS Nano* **8,** 9941–9953 (2014).

11. Luan, F. & (null), M. C. in *Recent Trends on QSAR in the pharmaeutical perceptions* (ed. Khan, M.) 194–241 (books.google.com, 2012). doi:10.2174/978160805379711201010194

12. Jenson, J. M. *et al.* Peptide design by optimization on a data-parameterized protein interaction landscape. *Proc. Natl. Acad. Sci. U.S.A.* **592,** 201812939 (2018).

13. Kholodovych, V. *et al.* Accurate predictions of cellular response using QSPR: a feasibility test of rational design of polymeric biomaterials. *Polymer* **45,** 7367–7379 (2004).

14. Hook, A. L. *et al.* Combinatorial discovery of polymers resistant to bacterial attachment. *Nat Biotechnol* **30,** 868–875 (2012).

15. Moe, A. A. K. *et al.* Microarray with Micro- and Nano-topographies Enables Identification of the Optimal Topography for Directing the Differentiation of Primary Murine Neural Progenitor Cells. *Small* **8,** 3050–3061 (2012).

16. Kukumberg, M., Yao, J. Y., Neo, D. J. H. & Yim, E. K. F. Microlens topography combined with vascular endothelial growth factor induces endothelial differentiation of human mesenchymal stem cells into vasculogenic progenitors. *Biomaterials* **131,** 68–85 (2017).

17. Unadkat, H. V. *et al.* An algorithm-based topographical biomaterials library to instruct cell fate. *Proc. Natl. Acad. Sci. U.S.A.* **108,** 16565–16570 (2011).

18. Ankam, S. *et al.* Substrate topography and size determine the fate of human embryonic stem cells to neuronal or glial lineage. *Acta Biomaterialia* **9,** 4535–4545 (2013).

19. Teo, B. K. K. *et al.* Nanotopography Modulates Mechanotransduction of Stem Cells and Induces Differentiation through Focal Adhesion Kinase. *ACS Nano* **7,** 4785–4798 (2013).

20. Zhou, Q. *et al.* Directional nanotopographic gradients: a high-throughput screening platform for cell contact guidance. *Sci. Rep.* **5,** 16240 (2015).

21. Balaban, N. Q. *et al.* Force and focal adhesion assembly: a close relationship studied using elastic micropatterned substrates. *Nature Cell Biology* **3,** 466–472 (2001).

22. Kulangara, K., Yang, Y., Yang, J. & Leong, K. W. Nanotopography as modulator of human mesenchymal stem cell function. *Biomaterials* **33,** 4998–5003 (2012).

23. Cavalcanti-Adam, E. A. *et al.* Cell Spreading and Focal Adhesion Dynamics Are Regulated by Spacing of Integrin Ligands. *Biophysical Journal* **92,** 2964–2974 (2007).

24. Natale, C. F., Ventre, M. & Netti, P. A. Tuning the material-cytoskeleton crosstalk via nanoconfinement of focal adhesions. *Biomaterials* **35,** 2743–2751 (2014).

25. Dalby, M. J. *et al.* Nanomechanotransduction and interphase nuclear organization influence on genomic control. *J. Cell. Biochem.* **102,** 1234–1244 (2007).

26. Dalby, M. J., Gadegaard, N. & Wilkinson, C. D. W. The response of fibroblasts to hexagonal nanotopography fabricated by electron beam lithography. *J Biomed Mater Res A* **84,** 973–979 (2008).

27. Dalby, M. J. *et al.* Group analysis of regulation of fibroblast genome on low-adhesion nanostructures. *Biomaterials* **28,** 1761–1769 (2007).

28. Dalby, M. J., Riehle, M. O., Yarwood, S. J., Wilkinson, C. D. W. & Curtis, A. S. G. Nucleus alignment and cell signaling in fibroblasts: response to a micro-grooved topography. *Exp. Cell Res.* **284,** 272–280 (2003).

29. Allan, C. *et al.* Osteoblast response to disordered nanotopography. *J Tissue Eng* **9,** 204173141878409–7 (2018).

30. Quarles, L. D., Yohay, D. A., Lever, L. W., Caton, R. & Wenstrup, R. J. Distinct proliferative and differentiated stages of murine MC3T3-E1 cells in culture: an in vitro model of osteoblast development. *Journal of Bone and Mineral Research* **7,** 683–692 (1992).

31. Wagers, A. J. & Conboy, I. M. Cellular and molecular signatures of muscle regeneration: current concepts and controversies in adult myogenesis. *Cell* **122,** 659–667 (2005).

32. Kumar, R. M. & Sreekumar, K. A survey on image feature descriptors. *Int J Comput Sci Inf Technol* (2014).

33. Tsimbouri, P. *et al.* Nanotopographical effects on mesenchymal stem cell morphology and phenotype. *J. Cell. Biochem.* **115,** 380–390 (2014).

34. Van De Water, L., Varney, S. & Tomasek, J. J. Mechanoregulation of the Myofibroblast in Wound Contraction, Scarring, and Fibrosis: Opportunities for New Therapeutic Intervention. *Advances in Wound Care* **2,** 122–141 (2013).

35. Bollong, M. J. *et al.* Small molecule-mediated inhibition of myofibroblast transdifferentiation for the treatment of fibrosis. *Proceedings of the National Academy of Sciences* **114,** 4679–4684 (2017).

36. Quach, N. L., Biressi, S., Reichardt, L. F., Keller, C. & Rando, T. A. Focal Adhesion Kinase Signaling Regulates the Expression of Caveolin 3 and β1 Integrin, Genes Essential for Normal Myoblast Fusion. *Molecular Biology of the Cell* **20,** 3422–3435 (2009).

37. Clemente, C. F. M. Z., Corat, M. A. F., Saad, S. T. O. & Franchini, K. G. Differentiation of C 2C 12myoblasts is critically regulated by FAK signaling. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* **289,** R862–R870 (2005).

38. Bachmann, P. & Hinrichsen, K. Semiautomated evaluation of morphometric data of myoblasts and myotubes and the calculation of the relative DNA content with the television texture analysis system (TAS). *Histochemistry* **69,** 233–242 (1980).

39. Lassalle, B., Gautron, J., Martelly, I. & Le Moigne, A. Image analysis of rat satellite cell proliferation in vitro. *Cytotechnology* **2,** 213–224 (1989).

40. Treiser, M. D. *et al.* Cytoskeleton-based forecasting of stem cell lineage fates. *Proc. Natl. Acad. Sci. U.S.A.* **107,** 610–615 (2010).

41. Kilian, K. A., Bugarija, B., Lahn, B. T. & Mrksich, M. Geometric cues for directing the differentiation of mesenchymal stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **107,** 4872–4877 (2010).

42. Swift, J. *et al.* Nuclear lamin-A scales with tissue stiffness and enhances matrix-directed differentiation. *Science* **341,** 1240104 (2013).

43. Shin, H. *et al.* Focal Adhesion Assembly Induces Phenotypic Changes and Dedifferentiation in Chondrocytes. *J. Cell. Physiol.* **231,** 1822–1831 (2016).

44. Connelly, J. T., García, A. J. & Levenston, M. E. Interactions between integrin ligand density and cytoskeletal integrity regulate BMSC chondrogenesis. *J. Cell. Physiol.* **217,** 145–154 (2008).

45. Vasilopoulos, Y., Gkretsi, V., Armaka, M., Aidinis, V. & Kollias, G. Actin cytoskeleton dynamics linked to synovial fibroblast activation as a novel pathogenic principle in TNF-driven arthritis. *Annals of the Rheumatic Diseases* **66 Suppl 3,** iii23–8 (2007).

46. Dalby, M. J. Cellular response to low adhesion nanotopographies. *Int J Nanomedicine* **2,** 373–381 (2007).

47. Newman, P. *et al.* Relationship between nanotopographical alignment and stem cell fate with live imaging and shape analysis. *Sci. Rep.* **6,** 37909 (2016).

48. Wong, V. W. *et al.* Focal adhesion kinase links mechanical force to skin fibrosis via inflammatory signaling. *Nat Med* **18,** 148–152 (2011).

49. Castillo, A. B. *et al.* Focal Adhesion Kinase Plays a Role in Osteoblast Mechanotransduction In Vitro but Does Not Affect Load-Induced Bone Formation In Vivo. *PLoS ONE* **7,** e43291–11 (2012).

50. Salasznyk, R. M., Klees, R. F., Williams, W. A., Boskey, A. & Plopper, G. E. Focal adhesion kinase signaling pathways regulate the osteogenic differentiation of human mesenchymal stem cells. *Exp. Cell Res.* **313,** 22–37 (2007).

51. Wang, I.-N. E. *et al.* Role of osteoblast-fibroblast interactions in the formation of the ligament-to-bone interface. *J. Orthop. Res.* **25,** 1609–1620 (2007).

52. Gautrot, J. E. *et al.* The Nanoscale Geometrical Maturation of Focal Adhesions Controls Stem Cell Differentiation and Mechanotransduction. *Nano Lett.* **14,** 3945–3952 (2014).

53. Changede, R., Cai, H., Wind, S. & Sheetz, M. P. Ligand Geometry Controls Adhesion formation via Integrin Clustering. *bioRxiv* (2018). doi:10.1101/435826

54. Li, B., Moshfegh, C., Lin, Z., Albuschies, J. & Vogel, V. Mesenchymal stem cells exploit extracellular matrix as mechanotransducer. *Sci. Rep.* **3,** 2425 (2013).

55. Trappmann, B. *et al.* Extracellular-matrix tethering regulates stem-cell fate. *Nature Materials* **11,** 642–649 (2012).

56. Oria, R. *et al.* Force loading explains spatial sensing of ligands by cells. *Nature* 1–20 (2017). doi:10.1038/nature24662

57. Guillaume-Gentil, O. *et al.* Tunable Single-Cell Extraction for Molecular Analyses. *Cell* **166,** 506–516 (2016).

58. Lindley, D. V. On a Measure of the Information Provided by an Experiment. *The Annals of Mathematical Statistics* **27,** 986–1005 (1956).

59. Chaloner, K. & Verdinelli, I. Bayesian Experimental Design: A Review. *Statistical Science* **10,** 273–304 (1995).

60. Stormonth-Darling, J. M., Saeed, A., Reynolds, P. M. & Gadegaard, N. Injection Molding Micro- and Nanostructures in Thermoplastic Elastomers. *Macromol. Mater. Eng.* **301,** 964–971 (2016).

61. Yan, X.-Z. *et al.* Effects of continuous passaging on mineralization of MC3T3-E1 cells with improved osteogenic culture protocol. *Tissue Engineering Part C: Methods* **20,** 198–204 (2014).

62. Caicedo, J. C. *et al.* Data-analysis strategies for image-based cell profiling. *Nat Meth* **14,** 849–863 (2017).

63. Kamentsky, L. *et al.* Improved structure, function and compatibility for CellProfiler: modular high-throughput image analysis software. *Bioinformatics* **27,** 1179–1180 (2011).

64. Zhan, X., Patterson, A. D. & Ghosh, D. Kernel approaches for differential expression analysis of mass spectrometry-based metabolomics data. *BMC Bioinformatics* **16,** 77 (2015).

65. Warnes, G. R. *et al.* gplots: Various R programming tools for plotting data. *R package version* **2,** 1 (2009).

66. Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M. & Hornik, K. cluster: Cluster Analysis Basics and Extensions. (2017).

67. Wei, T. *et al.* **corrplot: Visualization of a Correlation Matrix**. (2017).

68. Kuhn, M. Building Predictive Models in R Using the caret Package. *Journal of Statistical Software* **28,** 1–26 (2008).

69. Bürkner, P.-C. brms: An RPackage for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software* **80,** 1–28 (2017).

70. Team, R. R: A language and environment for statistical computing. (2017).

71. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag New York, 2009).

72. Akima, H. & Gebhardt, A. akima: Interpolation of Irregularly and Regularly Spaced Data.