bioRxiv preprint doi: https://doi.org/10.1101/867507; this version posted December 6, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

A FIJI Macro for quantifying pattern in extracellular matrix

Esther Wershof[1,2], David J Barry[3], Robert P Jenkins[2], Antonio Rullan[2,4], Anna Wilkins[2], Ioannis Roxanis[5], Kurt I Anderson[3], Danielle Park[6]*, Paul A Bates[1]*, Erik Sahai[2]*

[1] Biomolecular Modelling Laboratory, The Francis Crick Institute, London, United Kingdom
[2] Tumour Cell Biology Laboratory, The Francis Crick Institute, London, United Kingdom
[3] Advanced Light Microscopy Facility, The Francis Crick Institute, London, United Kingdom
[4] The Institute of Cancer Research, United Kingdom
[5] Breast cancer research, The Institute of Cancer Research, United Kingdom
[6] Developmental Signalling Laboratory, The Francis Crick Institute, London, United Kingdom.

* authors for correspondence
danielle.park@crick.ac.uk
paul.bates@crick.ac.uk
erik.sahai@crick.ac.uk

bioRxiv preprint doi: https://doi.org/10.1101/867507; this version posted December 6, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Diverse extracellular matrix patterns are observed in both normal and pathological tissue; however only piecemeal tools currently exist for quantitative analysis of matrix pattern features. Thus, there is a need for an automated pipeline for quantifying these matrix patterns. To this end we have developed an ImageJ plugin called TWOMBLI, which stands for The Workflow Of Matrix BioLogy Informatics. The aim of TWOMBLI is to quantify matrix patterns, including branch points, endpoints, curvature, gaps, and fractal dimension. It is designed to be easy-to-use, quick and versatile. TWOMBLI can be downloaded from https://github.com/wershofe/TWOMBLI together with detailed documentation. Here we present an overview of the pipeline and demonstrate its utility for quantifying extracellular matrix patterns in both experimental and clinical samples.

Introduction

The extracellular matrix (ECM) provides support and structure to multicellular organisms and also guides the migration of cells [1,2], including leukocytes engaged in immune surveillance [3]. The ECM is re-built and remodelled in response to tissue damage, with changes in ECM composition and structure occurring during the aging process and in pathologies, such as cancer. The architecture of the ECM has lately been the subject of renewed focus [4–6]. Numerous methods exist for revealing the organisation of the ECM, including both antibody-based and chemical-based histochemistry protocols, such as picrosirius red staining[7]. However, formal quantification of ECM patterns, both for use in experimental studies and in pathological decision-making is lacking. The generation of metrics that describe ECM pattern could lead to insights into a wide range of fields, ranging from experimentalists interested in cell migration and remodelling of matrices, to clinicians researching conditions such as cancer and fibrosis.

Despite the potential utility of quantitative metrics that describe the diverse extracellular matrix patterns observed in both normal and pathological tissue, current tools typically only report one feature of the ECM pattern. For example, pipelines already exist for quantifying alignment of ECM [8] (CTfire) and for the evaluation of gaps in the ECM[9]; however, they are encoded in different MatLab scripts. OrientationJ [10] is an automated ImageJ plugin that is able to create vector fields and perform directional analysis on fibres, but this does not lend itself to quantifying overall matrix patterns. Hence, current tools for quantifying ECM pattern are fragmented and require heavy manual intervention [11,12]. Furthermore, methods for quantifying other spatial properties of the ECM are lacking. For example, the curvature of fibres or the fractal dimension of the matrix structure could be important characteristics.

As outlined above, there is a need for an end-to-end pipeline for quantifying ECM patterns, which is automated and easy to use on versatile data sets. To this end, we have created the ImageJ macro plugin TWOMBLI. TWOMBLI stands for The Workflow Of Matrix BioLogy Informatics (the name TWOMBLI is also a nod to the American artist Cy Twombly whose works are full of diverse marks and patterning [13]). The aim of TWOMBLI is to quantify matrix patterns

bioRxiv preprint doi: https://doi.org/10.1101/867507; this version posted December 6, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

in an ECM image by deriving a range of metrics, which can then be analysed together with other metrics, such as transcriptomic and proteomic data, and correlated with clinical data.

Results

Our goal was to generate a comprehensive tool for analysis of ECM patterns, whereby a user could enter versatile matrix images (e.g. fibronectin staining, second harmonic collagen imaging, or picrosirius red staining) into the pipeline and derive a meaningful quantification of the matrix patterns as output (Figure 1). Image J was employed as the supporting platform for the plugin, since it enabled us to build on many existing tools and downloads already developed for use with Image J. With a history dating back to 1997, Image J (also known as Fiji when downloaded as a package containing plugins and documentation) is widely used by many scientists and clinicians[14]. We first selected metrics that covered a range of spatial features of matrix patterns (described in the following section), before assembling an integrated FIJI plug-in for generating the metrics. Finally, we tested our pipeline on both experimentally-generated images of ECM and human pathological samples.

Metrics for matrix quantification

Matrix patterns were quantified by deriving a number of metrics that can be split between those describing individual fibres and those describing general ECM pattern: number of fibre end points, number of fibre branchpoints, total length of fibres, fibre curvature, proportion of high-density matrix, ECM fractal dimension, hyphal growth unit (a measure of the number of end points per unit length) and lacunarity (a measure of how the ECM fills the space) [15]. These metrics together could characterise different matrix properties. A schematic diagram of some key metrics is shown in Figure 2 (left-hand panels). Human breast cancer samples stained with picrosirius red to highlight collagen fibres are shown alongside the diagrams to provide examples of tissue with low and high levels of endpoints, high density matrix, curvature, fractal dimension, and branchpoints (Figure 2, right-hand panels).

Metrics relating to individual fibres

The number of endpoints is an intuitive count of the number of ends of the mask fibres in the image. The number of branchpoints is the number of intersections of mask fibres in the image. Total length is the sum of the length of all mask fibres in the image. This quantity can be useful in normalising the number of branchpoints and endpoints. Curvature was measured as the mean change in angle moving incrementally along individual mask fibres by user-specified windows. A high curvature score is indicative of tight curvature, i.e. that which can be described with a small radius.

Metrics of global pattern

High-density matrix (HDM) is a measure of the proportion of pixels in an image above a user-specified threshold intensity, indicating highly collagenous regions. Fractal dimension is an indicator of the self-similarity and complexity of the ECM and is bound between in the range [1,2] for a single 2D image slice. Specifically, the metric used is the box-counting dimension [16]. A grid with squares of side length ϵ is overlaid over the image. The number of squares $N(\epsilon)$ which are occupied by the non-background part of the image N is recorded. As ϵ gets smaller, $N(\epsilon)$ increases. Fractal dimension is then computed as the limit of the following equation:

bioRxiv preprint doi: https://doi.org/10.1101/867507; this version posted December 6, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

$$frac = \lim_{\epsilon \to 0} \frac{logN(\epsilon)}{\log(1/\epsilon)}.$$

In addition, an optional step in the plugin (step 7) allows the user to perform gap analysis on matrix patterns using the Max Inscribed Circle function available from the BIOP plugin in FIJI. The algorithm analyses spaces between objects (in this case, the fibres in the masks derived by TWOMBLI) by fitting circles of decreasing radius to fill in the gaps [17]. Information on the size and location of these circles can provide additional insight into the structure of the matrix patterns.

Development of the TWOMBLI pipeline

TWOMBLI was developed so that a user can input tissue samples stained for ECM components and is then guided through image pre-processing before finally being given a comprehensive output of ECM metrics in csv files with accompanying processed images. Furthermore, an additional script is included in the TWOMBLI repository for optional cropping down of images to regions of interest. TWOMBLI exists as a user-friendly ImageJ plugin that can be downloaded with detailed documentation (see Documentation) from https://github.com/wershofe/TWOMBLI. A tutorial video is also available from https://bmm.crick.ac.uk/TWOMBLI/. TWOMBLI relies heavily on existing ImageJ plugins Ridge Detection [18] and AnaMorf [15] , providing an end-to-end pipeline for users. Figure 3A shows an example of user input together with the different outputs from the pipeline. Of note, the original image is converted into a skeletal mask that captures the matrix filaments. This is used as the basis of all metrics except the high density matrix metric, which is determined via a simple thresholding step. The principal stages of TWOMBLI are: prechecking, pre-processing and processing. The prechecks (steps 0-3) are carried out manually by the user, following prompts, to check the eligibility of potential input images. An image needs to be in focus, have high enough resolution and not contain too many artefacts or regions that are not pertinent for the analysis. A resolution of 1 micron per pixel or higher is recommended. In pre-processing (steps 4-9), the user is guided through selecting a subset of test images and then choosing appropriate parameters for thresholding these test images. These parameters are saved so that in future runs, the user can skip directly to the processing stage. In processing (steps 10-14), all of the images are analysed in batch using the parameters generated in the pre-processing stage. An input image of 1MB in size takes approximately 30 seconds to process on a computer with 2.2 GHz Intel Core i7 and 16GB of memory. For tips on handling larger files and other troubleshooting see Documentation.

Having established the TWOMBLI pipeline, we sought to test it. Figure 3B demonstrates that the pipeline can be used to analyse images of ECM acquired via a wide variety of imaging techniques and of varying file types: including fibronectin staining of *in vitro* generated cell-derived matrices, second harmonic imaging of mouse tissue, and matrix patterns generated by *in silico* simulation of fibroblast behaviour.

Output metrics are able to distinguish between different matrix patterns

In Figure 4A we present specific examples of output metrics of matrix patterns using TWOMBLI. Importantly, the TWOMBLI pipeline is able to quantify differences in endpoints,

bioRxiv preprint doi: https://doi.org/10.1101/867507; this version posted December 6, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

high density matrix, curvature, fractal dimension, and branch points in the representative examples of ECM shown in Figure 2 (highlighted with blue and beige shading in Figure 4). For metrics relating to individual fibres, endpoints and branch points, normalisation to the total fibre length is recommended. The normalised values are provided in brackets in Figure 4. Quantification of matrix from the breast cancer biopsies presented in Figure 2 shows how some metrics such as high-density matrix (HDM) are intuitively different, whilst other metrics such as curvature and branchpoints are more nuanced. This suggests that TWOMBLI could be particularly helpful in identifying patterning properties that are difficult to ascertain by eye (Figure 4A). Next, we wanted to test if TWOMBLI could extract additional information from images that had previously been reported to exhibit differing degrees of anisotropy. To this end, three images of anisotropic matrix and three images of isotropic matrix produced by FDM [4] were used as input and corresponding masks generated (Figure 4B). Encouragingly, the anisotropic matrix patterns were quantified as having lower normalised endpoints, curvature, fractal dimension and normalised branchpoints. Taken together, these analyses demonstrate that TWOMBLI is able to distinguish between clearly different matrix patterns. Importantly, TWOMBI is also able to determine less obvious quantitative difference in matrix patterns and provide additional classes of information not captured using existing tools.

bioRxiv preprint doi: https://doi.org/10.1101/867507; this version posted December 6, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Discussion

Much work has focused on identifying patterns of cells in tissues [19–21]. However, ECM organisation, which plays a crucial part in tissue architecture, has been largely neglected. We have developed TWOMBLI to quantify matrix patterns, enabling further understanding of the relevance of ECM organisation in a wide range of contexts, from tissue damage [22–24] to ageing [25], development [26] and fibrotic disease. Analysis of the metrics generated in multidimensional space will enable a numerical description of the types of ECM observed in both normal physiology and pathology. Furthermore, the inter-relationship between the metrics may shed light on the mechanisms of pattern formation and the limits of classes of pattern. Based on the limited analysis presented here, the inter-relationship of branch point, curvature, and fractal dimension warrants further analysis.

It is our sincere hope that this tool will be of use to researchers working with ECM. In particular, the ability to derive quantitative metrics will enable researchers to objectively compare data sets with associated benefits in the robustness of observations. and to correlate matrix patterns with other classes of data, including transcriptomic, proteomic, and clinical datasets. Linking to gene expression and proteomic data should provide new insights into the underlying mechanisms that generate matrix pattern, while correlation with clinical data should inform us about the consequences of variation in ECM pattern.

bioRxiv preprint doi: https://doi.org/10.1101/867507; this version posted December 6, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Methods

Computational analysis

All computational methods are described in detail in the TWOMBLI documentation which can be found at https://github.com/wershofe/TWOMBLI.

Picrosirius red staining

Samples were stained using ABCAM ab150681 Picrosirius Red Kit. Briefly, stains were deparaffinised and hydrated, we applied Picro-Sirius Red Solution for 60 minutes, rinsed twice in acetic acid, then alcohol dehydration. Slides were then scanned at 10x using Zeiss Axio Scan.Z1.

FDM assay

The FDM assay was performed as described in Park et al (Nat materials paper). Briefly, glass-bottom dishes (MatTek, P35-1.5-14-C) were pre-prepared with 0.2% gelatin solution (1h, 37 °C), then 1% glutaraldehyde for 30 min at room temperature (RT). After PBS buffer solution wash, the plate was incubated with 1 M ethanolamine for 30 min (RT). After two washes with PBS we seeded 7 × 104 cells in media with 100 µg ml−1 ascorbic acid ((+)-sodium L-ascorbate, Sigma, A4034). Cells were kept for 6 days and media changed every 2 days. We used extraction buffer and washed several times with PBS before immunofluorescence for ECM using anti-fibronectin-FITC (1:50 dilution; Abcam, ab72686). Samples were imaged using a Zeiss LSM 780 microscope.

Collagen Imaging

Second harmonic generation imaging for collagen was performed using a Zeiss LSM 780 microscope with Mai Tai multi-photon laser

Supplementary Information

Supplementary Text 1 = TWOMBLI Documentation

Supplementary Video = TWOMBLI Tutorial

bioRxiv preprint doi: https://doi.org/10.1101/867507; this version posted December 6, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Bibliography

1. Frantz, C., Stewart, K. M. & Weaver, V. M. The extracellular matrix at a glance. *J. Cell Sci.* (2010). doi:10.1242/jcs.023820
2. van Helvert, S., Storm, C. & Friedl, P. Mechanoreciprocity in cell migration. *Nat. Cell Biol.* **20**, 8–20 (2018).
3. Hallmann, R. *et al.* The regulation of immune cell trafficking by the extracellular matrix. *Current Opinion in Cell Biology* (2015). doi:10.1016/j.ceb.2015.06.006
4. Park, D. *et al.* Extracellular matrix anisotropy is determined by TFAP2C-dependent regulation of cell collisions. *Nat. Mater.* (2019). doi:10.1038/s41563-019-0504-3
5. Wershof, E. *et al.* Matrix feedback enables diverse higher-order patterning of the extracellular matrix. *PLoS Comput. Biol.* (2019). doi:10.1371/journal.pcbi.1007251
6. Yuan, Y. Spatial heterogeneity in the tumor microenvironment. *Cold Spring Harb. Perspect. Med.* **6**, (2016).
7. Constantine, V. S. & Mowry, R. W. Selective Staining of Human Dermal Collagen: II. The Use of Picrosirius Red F3BA with Polarization Microscopy*. *J. Invest. Dermatol.* **50**, 419–423 (1968).
8. Bredfeldt, J. S. *et al.* Computational segmentation of collagen fibers from second-harmonic generation images of breast cancer. *J. Biomed. Opt.* (2014). doi:10.1117/1.jbo.19.1.016007
9. Tozluoğlu, M. *et al.* Matrix geometry determines optimal cancer cell migration strategy and modulates response to interventions. *Nat. Cell Biol.* **15**, 751–762 (2013).
10. Sage, D. OrientationJ. *Experimental investigation of collagen waviness and orientation in the arterial adventitia using confocal laser scanning microscopy, Biomechanics and modeling in mechanobiology* (2012).
11. Manning, C. S. *et al.* Intravital imaging reveals conversion between distinct tumor vascular morphologies and localized vascular response to Sunitinib. *IntraVital* (2013). doi:10.4161/intv.24790
12. Rezakhaniha, R. *et al.* Experimental investigation of collagen waviness and orientation in the arterial adventitia using confocal laser scanning microscopy. *Biomech. Model. Mechanobiol.* **11**, 461–473 (2012).
13. Rondeau, J. Cy twombly. *Art Institute of Chicago Museum Studies* (2008). doi:10.2307/20205578
14. Schindelin, J., Rueden, C. T., Hiner, M. C. & Eliceiri, K. W. The ImageJ ecosystem: An open platform for biomedical image analysis. *Mol. Reprod. Dev.* **82**, 518–529 (2015).
15. Barry, D. J., Williams, G. A. & Chan, C. Automated analysis of filamentous microbial morphology with AnaMorf. *Biotechnol. Prog.* **31**, 849–852 (2015).
16. Cross, S. S. Fractals in pathology. *J. Pathol.* **182**, 1–8 (1997).
17. Acton, S. E. *et al.* Dendritic cells control fibroblastic reticular network tension and lymph node expansion. *Nature* (2014). doi:10.1038/nature13814
18. Steger, G. An unbiased detector of curvilinear structures. *IEEE Trans. Pattern Anal. Mach. Intell.* (1998). doi:10.1109/34.659930
19. Heindl, A., Nawaz, S. & Yuan, Y. Mapping spatial heterogeneity in the tumor microenvironment: a new era for digital pathology. *Lab. Investig.* **95**, 377–384 (2015).
20. Nawaz, S., Heindl, A., Koelble, K. & Yuan, Y. Beyond immune density: Critical role of spatial heterogeneity in estrogen receptor-negative breast cancer. *Mod. Pathol.*

bioRxiv preprint doi: https://doi.org/10.1101/867507; this version posted December 6, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(2015). doi:10.1038/modpathol.2015.37

21. Roxanis, I., Colling, R., Kartsonaki, C., Green, A. R. & Rakha, E. A. The significance of tumour microarchitectural features in breast cancer prognosis : a digital image analysis. 1–11 (2018). doi:10.1186/s13058-018-0934-x
22. Dallon, J. C., Sherratt, J. A. & Maini, P. K. Mathematical modelling of extracellular matrix dynamics using discrete cells: Fiber orientation and tissue regeneration. *J. Theor. Biol.* (1999). doi:10.1006/jtbi.1999.0971
23. McDougall, S., Dallon, J., Sherratt, J. A. & Maini, P. K. Fibroblast migration and collagen deposition during dermal wound healing: mathematical modelling and clinical implications. *Philos. Trans. A. Math. Phys. Eng. Sci.* **364**, 1385–405 (2006).
24. Hirsch, T. *et al.* Regeneration of the entire human epidermis using transgenic stem cells. *Nature* (2017). doi:10.1038/nature24487
25. Kaur, A. *et al.* Remodeling of the collagen matrix in aging skin promotes melanoma metastasis and affects immune cell motility. *Cancer Discov.* CD-18-0193 (2018). doi:10.1158/2159-8290.CD-18-0193
26. Driskell, R. R. *et al.* Distinct fibroblast lineages determine dermal architecture in skin development and repair. *Nature* **504**, 277–281 (2013).

bioRxiv preprint doi: https://doi.org/10.1101/867507; this version posted December 6, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

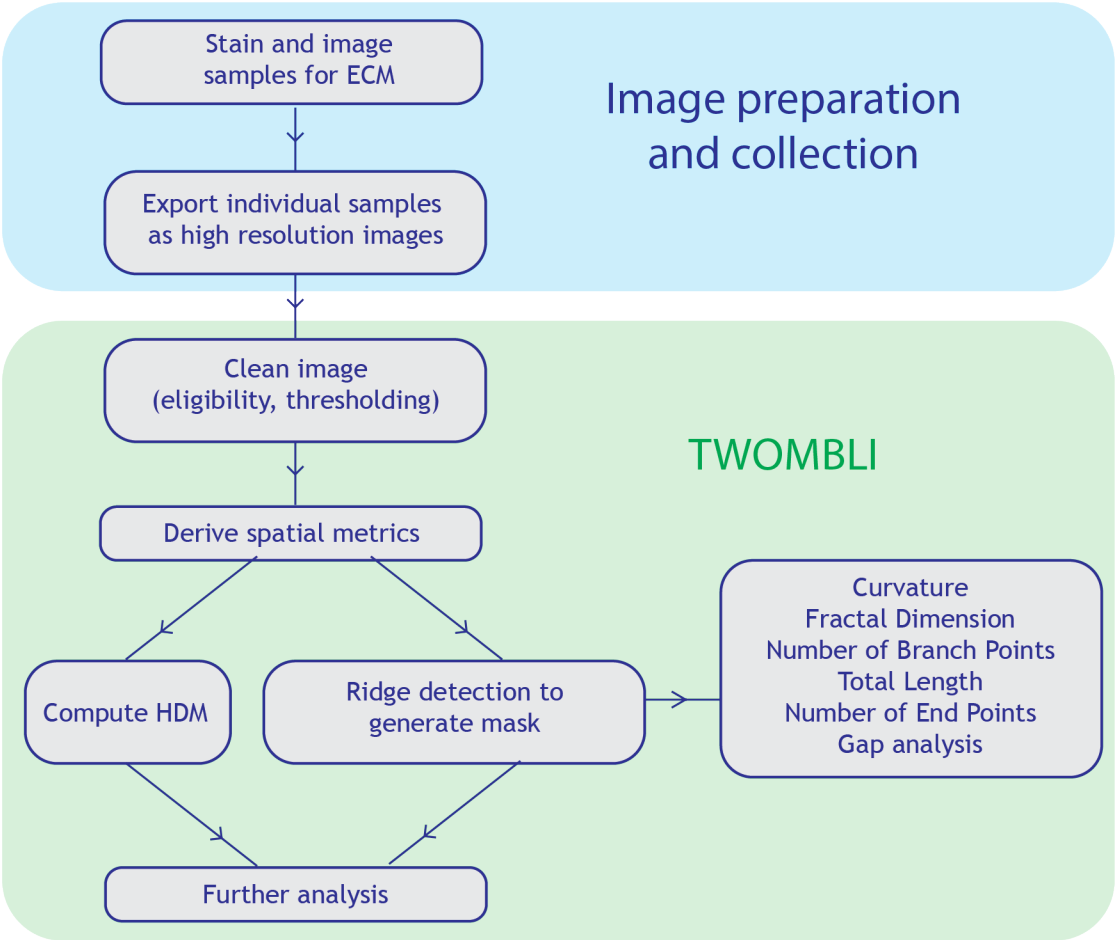


Figure 1: Workflow diagram of quantification of matrix patterns. End-to-end pipeline from obtaining the samples through matrix quantification to survival analysis based on this matrix metrology. A list of metrics is given in the right-most box.

bioRxiv preprint doi: https://doi.org/10.1101/867507; this version posted December 6, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

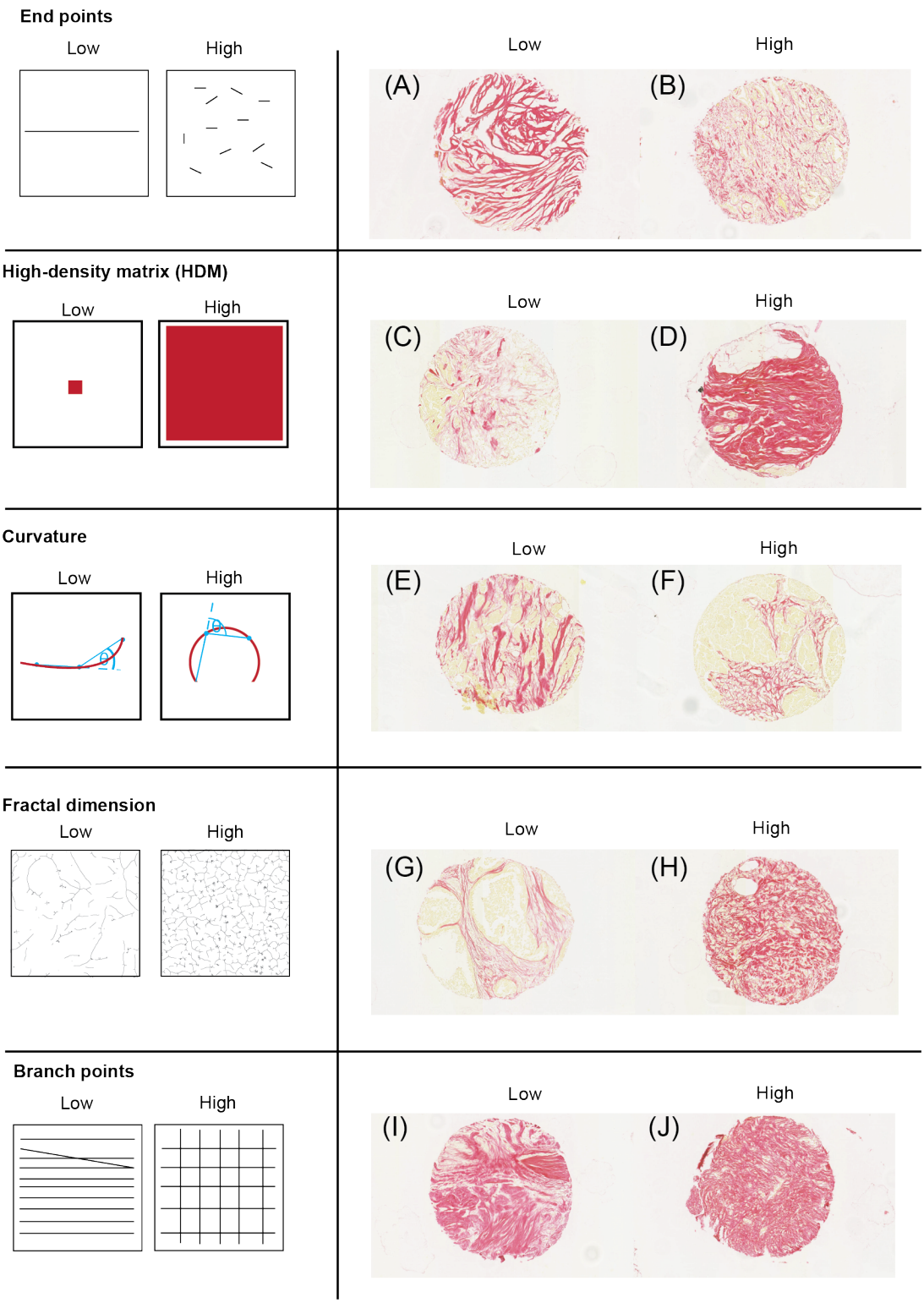
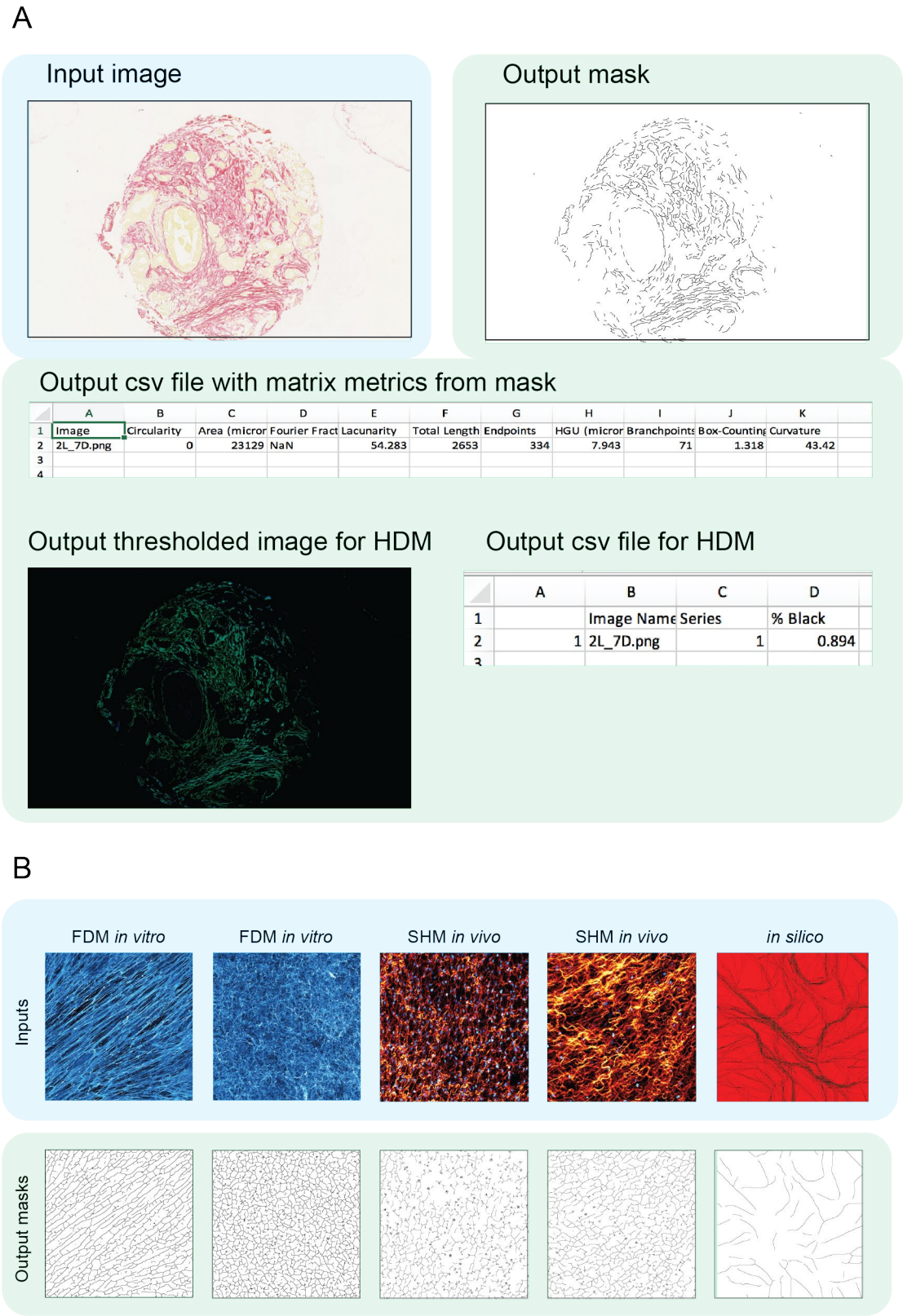


Figure 2: Schematics and example tissue biopsies of ECM metrics.

bioRxiv preprint doi: https://doi.org/10.1101/867507; this version posted December 6, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



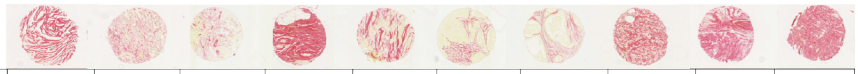
	A	B	C	D	E	F	G	H	I	J	K
1	Image	Circularity	Area (micror	Fourier Fract	Lacunarity	Total Length	Endpoints	HGU (micror	Branchpoints	Box-Counting	Curvature
2	2L_7D.png	0	23129	NaN	54.283	2653	334	7.943	71	1.318	43.42
3											
4											

	A	B	C	D
1		Image Name	Series	% Black
2	1	2L_7D.png	1	0.894
3				

Figure 3: Input and outputs of TWOMBLI pipeline. (A) User inputs an image of a sample stained for ECM components (in this case, collagen is stained with picosirius red). Outputs consist of a mask, a csv file containing matrix metrics based on the mask, a processed image thresholded for HDM and a corresponding csv file giving HDM. (B) Example input images can be acquired from FDM in vitro, second harmonic imaging in vivo and in silico ECM (from left to right – VCAF8-derived matrix, NF2-derived matrix, mouse liver, mouse dermis, and in silico[5]).

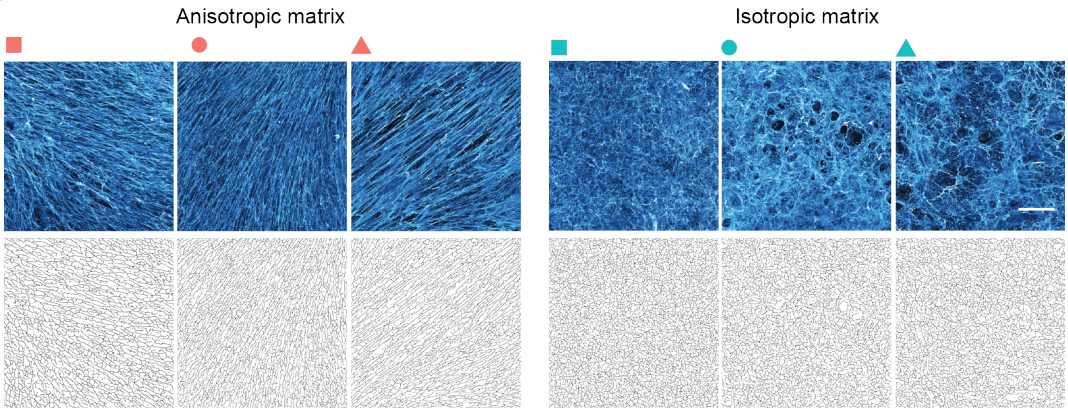
bioRxiv preprint doi: https://doi.org/10.1101/867507; this version posted December 6, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

A



	A	B	C	D	E	F	G	H	I	J
EndPoints	193 (0.010)	393 (0.023)	283 (0.031)	382 (0.022)	294 (0.022)	159 (0.022)	167 (0.019)	292 (0.014)	404 (0.024)	257 (0.011)
HighDensityMatrix	0.23	0.193	0.119	0.298	0.203	0.114	0.109	0.289	0.321	0.337
Curvature	39	53	54	48	27	82	39	48	61	56
FractalDimension	1.48	1.45	1.36	1.43	1.43	1.34	1.31	1.53	1.47	1.46
BranchPoints	324 (0.017)	266 (0.016)	77 (0.008)	129 (0.008)	120 (0.009)	198 (0.018)	59 (0.007)	449 (0.021)	166 (0.010)	500 (0.021)

B



C

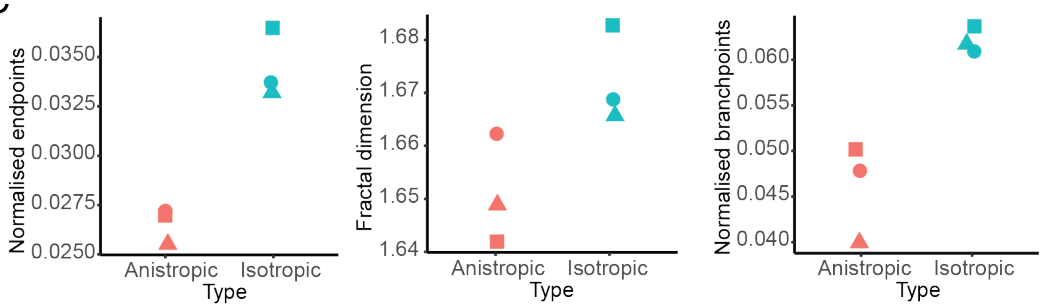


Figure 4: Using TWOMBLIi for quantification of ECM patterns: (A) Output metrics for the corresponding tumour biopsies shown in Figure 2. Pairs (A, B), (C, D) etc. show contrasting biopsies with low and high values of each metric as indicated in the table in blue/orange respectively. For endpoints and branchpoints, normalised values are given in brackets below the raw value. The normalised value is computed by dividing the raw value by the total length of the fibres in the mask. (B) FDM imaging of fibronectin produced by six different fibroblast lines in vitro these are (from left to right): hNLF, VCAF2B, VCAF8, NF2, CAF1, CAF2 [4]. Three patterns are anisotropic (left) and three patterns are isotropic (right). Corresponding masks derived in TWOMBLI are displayed underneath. (C) Quantification of matrix patterns from (B) across a number of metrics. The anisotropic matrices have lower number of normalised end points, HDM, curvature, fractal dimension and normalised branch points then the isotropic matrices. Normalisation is performed by dividing the raw value by the total length of fibres in the masks.