

# SCIENTIFIC REPORTS

OPEN

## $\mu$ MAPPS: a novel phasor approach to second harmonic analysis for in vitro-in vivo investigation of collagen microstructure

F. Radaelli<sup>1</sup>, L. D'Alfonso<sup>1</sup> , M. Collini<sup>1,3</sup>, F. Mingozzi<sup>2</sup>, L. Marongiu<sup>2</sup>, F. Granucci<sup>2</sup>, I. Zanoni<sup>2,4</sup>, G. Chirico<sup>1,3</sup> & L. Sironi<sup>1</sup>

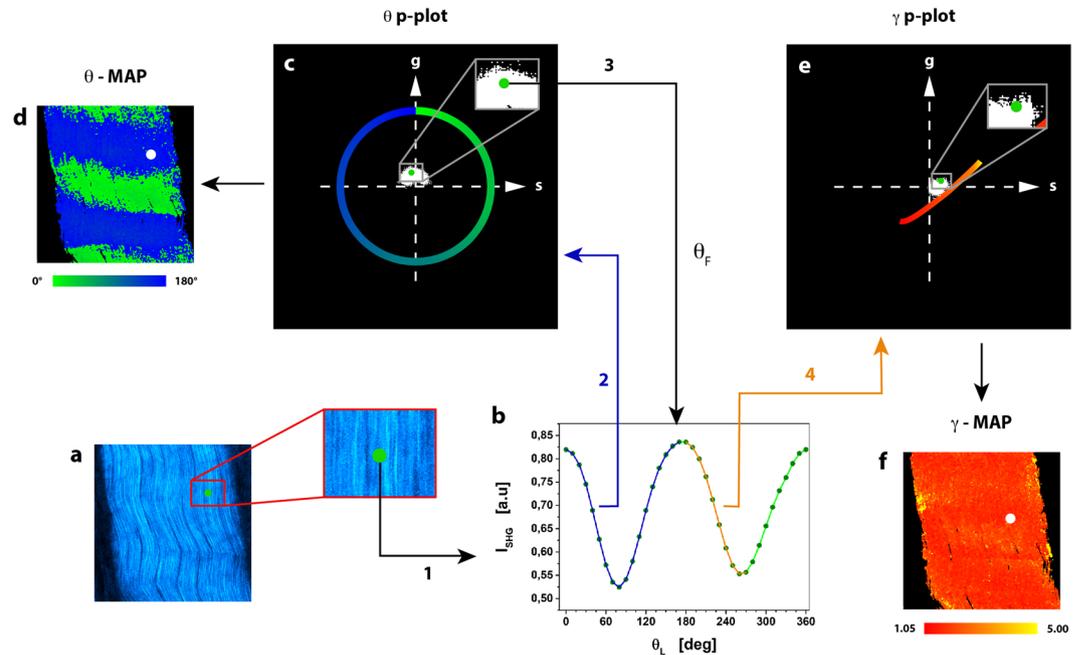
Second Harmonic Generation (SHG) is a label-free imaging method used to monitor collagen organization in tissues. Due to its sensitivity to the incident polarization, it provides microstructural information otherwise unreachable by other intensity based imaging methods. We develop and test a Microscopic Multiparametric Analysis by Phasor projection of Polarization-dependent SHG ( $\mu$ MAPPS) that maps the features of the collagen architecture in tissues at the micrometer scale.  $\mu$ MAPPS retrieves pixel-by-pixel the collagen fibrils anisotropy and orientation by operating directly on two coupled phasor spaces, avoiding direct fitting of the polarization dependent SHG signal. We apply  $\mu$ MAPPS to fixed tissue sections and to the study of the collagen microscopic organization in tumors *ex-vivo* and *in-vivo*. We develop a clustering algorithm to automatically group pixels with similar microstructural features.  $\mu$ MAPPS can perform fast analyses of tissues and opens to future applications for *in-situ* diagnosis of pathologies and diseases that could assist histo-pathological evaluation.

Collagen is the prominent constituent protein of tissues, providing them structural support and protection<sup>1,2</sup>. Since its organization changes in pathological processes<sup>3-9</sup>, it can be exploited as an early diagnostic marker in experimental and clinical medicine<sup>10,11</sup>. Collagen, myosin and microtubules in cells and tissues<sup>12,13</sup> can be non-invasively imaged also *in-vivo* by Second Harmonic Generation (SHG) microscopy, a nonlinear coherent optical process where two incident photons of frequency  $\omega$  are converted into a single photon of exactly twice the frequency,  $2\omega$ .

A part from the morphological information extracted from intensity-based SHG imaging<sup>14-19</sup>, the polarization dependence of the SHG response (P-SHG spectrum) can provide the molecular and supramolecular structure of different organs and tissues<sup>10,20-27</sup>. P-SHG microscopy is a promising and efficient diagnostic tool in the identification of cancer types<sup>14,17,23</sup> and other pathological conditions such as ischemia, skin pathologies, fibrosis and aging<sup>8,28-31</sup>. In fact, P-SHG imaging, in different experimental implementations<sup>27,32-39</sup> and with the support of various theoretical models of the fibrils organization, can be exploited to obtain pixel-by-pixel maps of the fibril orientation<sup>26,40,41</sup>, of the protein pitch angle<sup>12,42</sup> and of the nonlinear susceptibility tensor components<sup>12,15,20,43-46</sup>. Although usually a time consuming non-linear fitting of the P-SHG spectrum of each pixel is exploited, substantial speed-up in the analysis has been recently obtained through 1D Fast Fourier Transform of the P-SHG curves<sup>26</sup>.

Starting from a similar Fourier Transform (FT) approach, we build here a phasor analysis of the P-SHG images that acts at a global level and identifies automatically clusters of homogenous microstructure. The phasor approach has been extensively applied to graphically analyze spectra on a variety of techniques from fluorescence lifetime imaging to MRI<sup>47-50</sup>, allowing to extract quantitative information on a model parameter at a pixel-by-pixel level.

<sup>1</sup>Dipartimento di Fisica, Piazza della scienza 3, 20126, Milano, Italy. <sup>2</sup>Dipartimento di Bioscienze, Piazza della scienza 2, 20126, Milano, Italy. <sup>3</sup>CNR - ISASI, Institute of Applied Sciences & Intelligent Systems, Via Campi Flegrei 34, Pozzuoli, NA, Italy. <sup>4</sup>Harvard Medical School and Division of Gastroenterology, Boston Children's Hospital, Boston, MA, USA. Correspondence and requests for materials should be addressed to M.C. (email: [maddalena.collini@mib.infn.it](mailto:maddalena.collini@mib.infn.it)) or L.S. (email: [laura.sironi@unimib.it](mailto:laura.sironi@unimib.it))



**Figure 1.**  $\mu$ MAPPs applied to experimental SHG image stacks from mouse-tail tendon. (a) Maximum intensity projection of the image stack. Image size:  $225 \times 225 \mu\text{m}^2$ . (b) Graph of the P-SHG spectrum corresponding to the pixel highlighted by the green dot in (a). (c)  $\theta$ p-plot of the image stack according to Equation (1) over the  $[0, \pi]$  interval (blue portion of P-SHG curve in (b)). The green dot corresponds to the green dot in (a). The green-blue circle is the reference curve for  $\gamma \rightarrow \infty$ ; the color codes for the  $\theta_F$  angles. From the  $\theta$ p-plot, the angle  $\theta_F$  is retrieved by means of Equation (4) and used for the computation of the  $\gamma$  values. (d)  $\theta$ -map obtained pixel by pixel from the  $\theta$ p-plot (Equation (4)). (e)  $\gamma$ p-plot of the image stack according to Equation (2) over the range  $[\theta_F, \theta_F + \pi/2]$  (orange portion of P-SHG curve in (b)). The green dot corresponds to the green dot in (a). The red-yellow curve is the reference for the  $\gamma$ p-plot obtained for  $0 \leq \gamma \leq 10$  and  $\Delta\theta = 10^\circ$  assuming, as an example, the specific value  $\theta_F = 0^\circ$ . (f)  $\gamma$ -map obtained pixel by pixel from the  $\gamma$ p-plot by means of Equation (5).

We extend this global analysis by proposing a new method, named Microscopic Multiparametric Analysis by Phasor projection of P-SHG ( $\mu$ MAPPs), that provides micro-structural information (at pixel level) of the collagen architecture in tissues, both in terms of the mean orientation angle of the fibril ( $\theta_F$ ) and of the susceptibility anisotropy parameter ( $\gamma$ , the ratio of off-diagonal to diagonal elements of the susceptibility tensor  $\chi^{(2)}$ ).  $\mu$ MAPPs retrieves maps of both model parameters in few hundreds of ms to few seconds, by working on two coupled phasor spaces. Moreover, instead of manual and subjective clustering procedures exploited in other phasor approaches<sup>51–53</sup>, we develop an automated clustering algorithm, working simultaneously on the two phasor spaces, to back-project onto the original image the Regions of Interest (ROIs) characterized by similar  $\theta_F$  and  $\gamma$  values. In this way we are able to automatically single out patches sharing similar micro-structural properties.

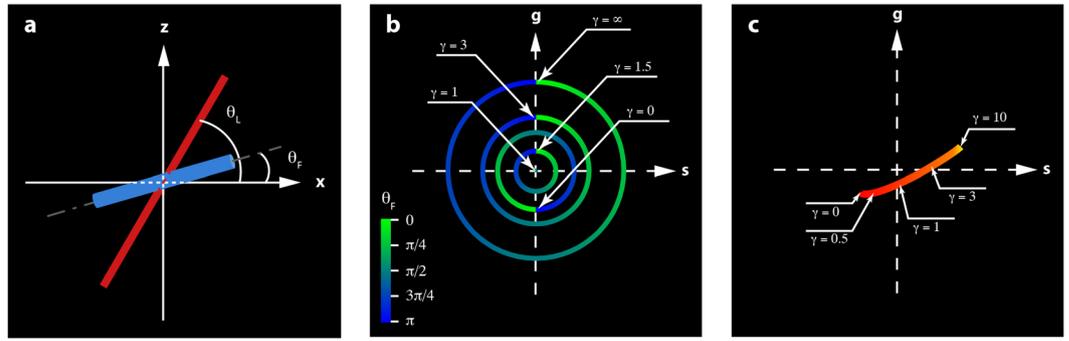
We prove that  $\mu$ MAPPs allows to take full advantage of the phasor approach. It provides a fast and intuitive representation of the data set, which is globally analyzed exploiting the dispersion of the data in the phasor space. In fact, differences in the microscopic behavior of the sample can be singled-out by segmenting patches of the phasor points distribution (related to the model parameters describing the system) directly in the phasor spaces and by measuring geometrical features of this distribution. Data denoising can be performed in the phasor space without affecting the spatial resolution of the original data. More important, by assuming a molecular model for the SHG response, the microstructural parameters retrieved graphically are not biased by the initial parameters values and constrains, contrary to the usual non-linear fitting procedure.

In the following a description of the  $\mu$ MAPPs core is presented together with its validation against simulations and *ex-vivo* experiments on mouse-tail tendons<sup>20,54</sup>.

The application of  $\mu$ MAPPs is then illustrated in the extraction of the described microscopic parameters in fixed and explanted tumor samples (biopsies). Finally, the robustness of our method has been tested in *in-vivo* experiments.

## Results

**$\mu$ MAPPs Analysis Method.** The  $\mu$ MAPPs projection method starts from the acquisition of a stack of images as a function of the laser polarization  $\theta_L$ . The maximum intensity projection of the stack provides the overall disposition of the collagen in the sample (Fig. 1a). The P-SHG spectrum (SHG as a function of  $\theta_L$ , Fig. 1b) of each pixel is projected onto a first complex plane (called  $\theta$ p-plot, Fig. 1c), via the first harmonic normalized Discrete Fourier Transform (DFT) applied to the  $0 \leq \theta_L < \pi$  range. Each pixel is transposed into a point, whose coordinates are the real and the imaginary parts of the signal DFT (see also Supplementary Notes 1–2 and Supplementary Figs. 1–3). The points in the  $\theta$ p-plot spread on a cloud around the origin (Fig. 1c) and their



**Figure 2.** Reference system for the P-SHG measurements. **(a)** Coordinate system scheme: the laser is incident along the  $y$ -direction, while the collagen fibrils lie in the  $xz$ -plane.  $\theta_L$  and  $\theta_F$  are, respectively, the laser excitation polarization angle and the fibril orientation angle with respect the  $x$ -axis direction. **(b)** Reference curves for the  $\theta_p$ -plot, obtained from Equations (1) and (3) for different values of  $\gamma$  and fixed value of  $\theta_F$ , as a function of  $\theta_L$  in the range  $[0, \pi)$ . For  $\gamma > 1$  the radius of the circle increases as  $\gamma$  increases; for  $0 \leq \gamma < 1$  the radius increases as  $\gamma$  decreases and it is simultaneously  $\pi$ -tilted. For  $\gamma = 1$ , the reference curve degenerates to a point in the origin. For  $0 < \gamma < 1$  and  $1 < \gamma < 2.1375$  (numerically obtained) the reference circles may be superimposed, although  $\pi$  shifted. The angle in the  $\theta_p$ -plot is double the  $\theta_F$  value. **(c)** Reference curve for the  $\gamma_p$ -plot. The mapping of the pixel on the  $\gamma_p$ -plot depends only on  $\gamma$  and falls, apart from data non-idealities, on the reference curve. For the simulations, a data sampling value  $\Delta\theta = 1^\circ$  was used.

absolute position depends on both the collagen fibrils angle  $\theta_F$  (Fig. 2a) and the microscopic order parameter. On the contrary, the angular position of these points in the  $\theta_p$ -plot depends only on the collagen fibrils angle  $\theta_F$ . A second phasor plot (called  $\gamma_p$ -plot, Fig. 1e), is then computed by DFT of the P-SHG in the domain  $\theta_F \leq \theta_L < \theta_F + \pi/2$  (Fig. 1b). The pixel projection in the  $\gamma_p$ -plot is related to the microscopic order parameter.

By mapping the P-SHG spectrum onto the two phasor plots we are therefore able to retrieve both the fibril orientation and the local order parameter pixel by pixel and build a  $\theta$ -map (Fig. 1d) and  $\gamma$ -map (Fig. 1f) of the sample from a stack of images. We have verified that  $\mu$ MAPPs (Equations (4–5)) provides results in agreement with the direct fitting of the P-SHG spectra to Equation (3).

More in details, the coordinates  $(g_\theta, s_\theta)$  of a point in the  $\theta_p$ -plot are the first DFT of the normalized P-SHG curve,  $\{I(\theta_L^n)\}_{n=0, N-1}$ , acquired as a function of  $\theta_L$  ( $0 \leq \theta_L < 3\pi/2$ ) with resolution  $\Delta\theta$ . We define the frequency  $K_\theta = 2\pi(N\Delta\theta)^{-1}$  and compute these coordinates as ( $N = \pi/\Delta\theta$ ):

$$g_\theta = \frac{\sum_{n=0}^{N-1} I(\theta_L^n) \cos(\theta_L^n K_\theta)}{\sum_{n=0}^{N-1} I(\theta_L^n)}; \quad s_\theta = \frac{\sum_{n=0}^{N-1} I(\theta_L^n) \sin(\theta_L^n K_\theta)}{\sum_{n=0}^{N-1} I(\theta_L^n)} \quad (1)$$

Each pixel is projected by Equation (1) onto the  $\theta_p$ -plot and from its angular position it is possible to estimate the fibril orientation  $\theta_F$  (Fig. 1c). This value corresponds to orientation of the fibril with respect to the polarization axis.

The anisotropy of the SHG susceptibility is measured by building a second phasor plot, called  $\gamma_p$ -plot, that depends on the estimated value  $\theta_F$  and whose coordinates are defined as ( $K_\gamma = 2\pi[N(\Delta\theta + \theta_F)]^{-1}$ ):

$$g_\gamma = \frac{\sum_{n=0}^{N/2-1} I(\theta_L^n + \theta_F) \cos((\theta_L^n + \theta_F) K_\gamma)}{\sum_{n=0}^{N-1} I(\theta_L^n + \theta_F)}; \quad s_\gamma = \frac{\sum_{n=0}^{N/2-1} I(\theta_L^n + \theta_F) \sin((\theta_L^n + \theta_F) K_\gamma)}{\sum_{n=0}^{N-1} I(\theta_L^n + \theta_F)} \quad (2)$$

In this case, we compute the DFT in the  $\theta_F \leq \theta_L < \theta_F + \pi/2$  range. In this way, we are maximally sensitive to the shape of the P-SHG spectrum (see Supplementary Fig. 2) with minimal influence from  $\theta_F$ . The position of the projection  $(g_\gamma, s_\gamma)$  in the  $\gamma_p$ -plot (Fig. 1e) is a measure of the P-SHG shape. The latter can be related, through a suitable model, to the local disorder of the fibrils.

Only P-SHG spectra above a threshold ( $T_{NL}$ ) are analyzed in order to discard the background contribution.

By exploiting the graphical analysis in the two linked phasor plots, we can map the P-SHG spectra of each pixel to the microscopic features of the sample. In order to quantify the degree of collagen disorder in the tissue, we need to choose a microscopic model for the SHG response and build in the  $\theta_p$ -plot and  $\gamma_p$ -plot two reference curves. We adopt here a widespread theoretical model<sup>15,20,43</sup> to describe the relation among the SHG signal,  $I(\theta_L^n)$ , the relative angle  $\theta_L^n - \theta_F$  between the laser polarization and the collagen fibrils orientations, and the anisotropy of the second order susceptibility  $\underline{\chi}^{(2)}$  tensor,  $\gamma = \chi_{zzz}^{(2)}/\chi_{zxx}^{(2)}$

$$I(\theta_L^n) = k \{ \sin^2[2(\theta_L^n - \theta_F)] + [\sin^2(\theta_L^n - \theta_F) + \gamma \cos^2(\theta_L^n - \theta_F)]^2 \} \quad (3)$$

The parameter  $\gamma > 0$  (ranging from  $\gamma \cong 0.5$  for myosin<sup>27,40</sup> to  $\gamma \cong 1.5$ –2 for collagen<sup>54</sup>) describes the molecular anisotropy and the scale factor  $k$  includes the absolute intensity of the SHG signal as affected by the setup parameters. Equation (3) is valid under the assumption of cylindrical ( $C_\infty$ ) or hexagonal ( $C_6$ ) plus Kleinman symmetries (see also Supplementary Note 1).

We build therefore reference curves in the  $\theta_F$ -plot and the  $\gamma$ -plot by applying the Equations (1–2) to the microscopic model described by Equation (3) as a function of  $\gamma$  and  $\theta_F$ . A closed circular curve (green-blue color coded curve in Fig. 2b) is obtained in the  $\theta_F$ -plot by varying the fibril angle  $\theta_F$  in the range  $[0, \pi]$ , while keeping  $\gamma$  constant.  $\gamma$  affects the  $\theta_F$ -plot reference curve by changing its radius (Fig. 2b). For  $\gamma > 1$  the circle radius increases with  $\gamma$ . For  $0 \leq \gamma < 1$  the radius increases as  $\gamma$  decreases and it is simultaneously  $\pi$ -tilted. For  $\gamma = 1$  the circle collapses at the phasor plot origin. Therefore, the angular position of the projection of the P-SHG spectrum on the  $\theta_F$ -plot determines univocally the fibril angle  $\theta_F$  according to the following algorithm:

$$\cos(2\theta_F) = \frac{g_\theta}{\sqrt{g_\theta^2 + s_\theta^2}} \quad (4)$$

$\theta_F$  is half the value of the angle that the vector  $(g_\theta, s_\theta)$ , pointing from the center of the  $\theta_F$ -plot to the pixel projection, makes with the phasor plot imaginary axis (see Supplementary Fig. 4). The reference curve of the  $\gamma$ -plot (Fig. 2c) is then built by DFT transforming (Equation (2)) the P-SHG spectrum (Equation (3)). We obtain an open reference curve (Fig. 2c, red-yellow color code) which is, unless for data sampling effects, independent of the  $\theta_F$  value. The parameter  $\gamma$  is measured by retrieving the minimum Euclidean distance projection onto the Reference Curve (RC, see Supplementary Fig. 4) defined as:

$$\begin{cases} d_{e-RC} = \sqrt{(g_e - g_{RC})^2 + (s_e - s_{RC})^2} \\ \gamma_e = \gamma_{RC} |_{\min(d_{e-RC})} \end{cases} \quad (5)$$

In Equation (5) the minimum is searched over the points of the  $\gamma$ -plot reference curve computed with the  $\theta_F$  value measured on the first phasor plot (Equation (4)) and with the experimental sampling angle  $\Delta\theta$ .

The algorithm for the  $\gamma$ -map retrieval (Equation (5)) was modified in order to take into account the onset, in tumor samples, of a non-negligible background value ( $\gamma_0$  term added to Equation (3)), probably due to an intra-pixel fibrils distribution. We studied the mutual dependence between  $\gamma$  and  $\gamma_0$  and modified the procedure to retrieve the  $\gamma$  values by minimizing the angle between the  $(g_\gamma, s_\gamma)$  vector and all vectors spanning the reference curve (Supplementary Note 3 and Supplementary Fig. 5).

It should be noted that our graphical method, by acting on a double phasor space, goes beyond the measure of the phase of the first Fourier component of the P-SHG spectrum as done in ref.<sup>26</sup>, which is equivalent to our Equation 4. Our double phasor space approach allows us to quantify both  $\theta_F$  and  $\gamma$  parameters, avoiding the algebraic solution proposed in ref.<sup>26</sup> to determine  $\gamma$ , and exploiting the denoising opportunities of the phasor approach.

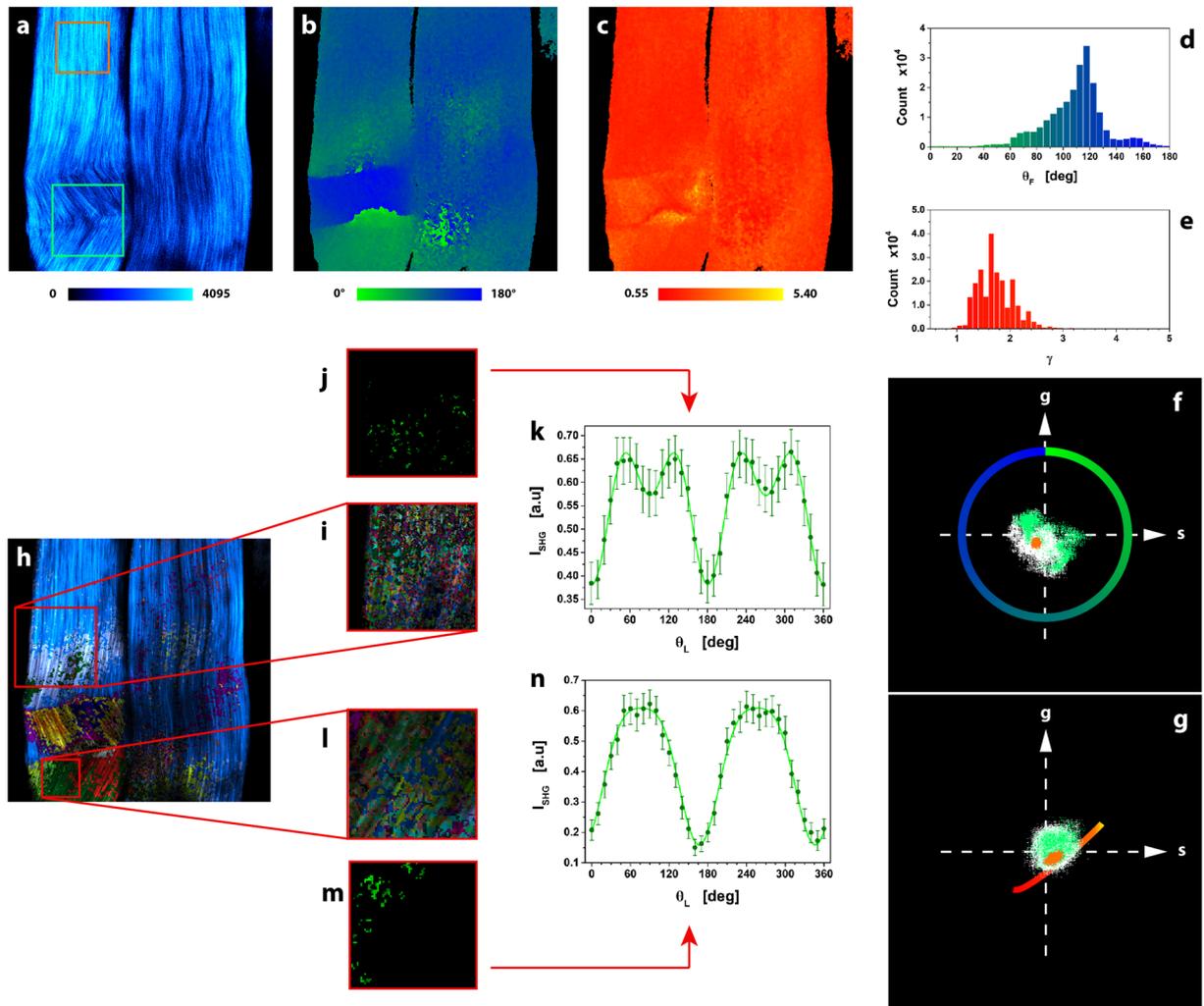
**Validation through numerical simulations.** The  $\mu$ MAPPs core has been thoroughly validated against numerical simulations.  $\mu$ MAPPs is able to tackle with nonideal experimental conditions such as the laser polarization sampling  $\Delta\theta$  and the noise (see Supplementary Fig. 6). For  $\Delta\theta \rightarrow 0^\circ$  the  $\gamma$ -plot reference curve is strictly independent of  $\theta_F$ . Non ideal data sampling (for example  $\Delta\theta = 10^\circ$  as in our experimental condition) results in the shift of the points from the ideal reference curve in the  $\gamma$ -plot. This is due to the fact that we compute the point coordinates  $(g_\gamma, s_\gamma)$  on the P-SHG spectrum shifted by an angle that is the nearest sampled  $\theta_F$  value estimated from the  $\theta_F$ -plot. Moreover, two P-SHG spectra characterized by  $\theta_F$  values differing by a multiple integer of  $\Delta\theta$  (and  $\gamma_1 = \gamma_2$ ), have similar projections on the  $\gamma$ -plot reference curve. In fact, exactly the same ( $\theta_F \leq \theta_L < \theta_F + \pi/2$ ) portion of identical P-SHG spectra are DFT analyzed in this case.

By working in the phasor plots, we are able to efficiently tackle with the experimental noise. P-SHG Gaussian noise scatters the projections in both phasor plots. This effect is enhanced if combined to a distribution of  $\gamma$  and  $\theta_F$  over the whole image stack (see Supplementary Fig. 7) or if a Gaussian distribution of fibrils angles within each pixel is considered (see Supplementary Fig. 8). A denoising filter was applied directly in the phasor space in order to reduce the dispersion of projected spectra without affecting the P-SHG intensity images<sup>35</sup>. In our case, a median filter with a  $3 \times 3$  pixels mask was directly and independently applied to the  $s$  and  $g$  coordinates (Supplementary Note 4 and Supplementary Fig. 9). In this way, it is possible to reduce the dispersion of the projections in the  $\theta_F$ -plot and  $\gamma$ -plot by  $4 \pm 0.6$  and  $3.4 \pm 0.3$  times, respectively.

Finally,  $\mu$ MAPPs allows us to solve efficiently also an intrinsic degeneracy of the model (Equation (3)). As shown in Supplementary Fig. 10, P-SHG spectra with  $\theta_1 = \theta_F$  and  $1 \leq \gamma_1 < 2.1375$  (numerically determined) and  $\theta_2 = \theta_F + \pi/2$  and  $0 < \gamma_2 < 1$  may fall on the same position on the  $\gamma$ -plot reference curve. In these cases, the experimental P-SHG spectrum is then directly compared to the model (Equation (3)) only for the couples  $\theta_F = \theta_1$ ,  $\gamma = \gamma_1$  (within the range  $[\theta_1, \theta_1 + \pi/2]$ ) and  $\theta_F = \theta_2$ ,  $\gamma = \gamma_2$  (within the range  $[\theta_2, \theta_2 + \pi/2]$ ). The solution with the minimum chi-squared value from the experimental P-SHG is selected. This procedure can be avoided if the type of collagen (e.g. Collagen I is characterized by values of  $\gamma > 1$ ) is a priori known.

**Data clustering.** For tissue analysis, it is essential to identify pixels characterized by similar microstructural parameters. Instead of manually (and somehow arbitrary) selecting different areas in the phasor plot and map the corresponding pixels back onto the morphological image, we develop here a data clustering approach (please refer to Supplementary Note 5).

Our clustering algorithm works pixel-by-pixel in the  $\theta_F - \gamma$  space. It first identifies the putative clusters' centers according to a maximum density approach<sup>56</sup> over both the  $\theta_F$  and  $\gamma$  parameters. Each element which is not a putative center is assigned, by an iterative procedure, to the cluster for which the element-center distance is the lowest. The validation of the clustering algorithm on P-SHG simulated noisy spectra (Supplementary Fig. 11) shows that the results are independent from the pixels location within the acquired image stack and the pixel average SHG



**Figure 3.**  $\mu$ MAPPs analysis of mouse-tail tendon. (a) Maximum intensity projection of the image stack ( $\Delta\theta = 10^\circ$ ). Image size:  $225 \times 225 \mu\text{m}^2$ . (b and c) represent respectively the pixel by pixel  $\theta$ - and  $\gamma$ - maps, where the color codes are shown in the legends. (d and e) are the counts histograms extracted from the  $\theta$ - and  $\gamma$ - maps. (f and g) show the  $\theta$ p-plot and  $\gamma$ p-plot respectively, from which the  $\theta$ - and  $\gamma$ -maps are derived according to Equations (4) and (5). The small compact cloud of points in the  $\theta$ - and  $\gamma$ p-plots (orange) corresponds to the uniform area highlighted in (a) with an orange box, while a non-homogeneous microscopic behavior (green box in a) causes points scattering (green) in both phasor plots. (h) Result of clustering procedure, highlighting the different microscopic behavior in the tissue. 23 clusters were obtained with the following cutoff values:  $\theta_C = 30^\circ$  and  $\gamma_C = 1$ . ET = 1%. The threshold noise level was  $T_{NL} = 250$ . Each monochromatic LUT encodes for a cluster with an intensity scaling as the pixel integrated P-SHG signal. (i and l) show two ROIs (i:  $154 \times 154$  pixels; l:  $69 \times 69$  pixels) selected from (h) and analyzed with tighter cutoff values ( $\theta_C = 2^\circ$ ,  $\gamma_C = 0.1$  and ET = 1%) to select clusters with more similar microscopic behavior. One cluster, selected from each ROI in (i) and (l), was reported in (j and m). The average P-SHG spectrum of the two selected clusters (reported in (k) and (n)) were fit to the model Equation (3). The best fit parameters  $\theta_F = 91.0^\circ \pm 0.4^\circ$  and  $\gamma = 1.43 \pm 0.01$  (cluster j),  $\theta_F = 76.5^\circ \pm 0.6^\circ$  and  $\gamma = 2.09 \pm 0.05$  (cluster m), are in excellent agreement with the  $\mu$ MAPPs estimates  $\theta_F = 92.2^\circ \pm 0.6^\circ$  and  $\gamma = 1.45 \pm 0.05$  (cluster j),  $\theta_F = 77.9^\circ \pm 0.6^\circ$  and  $\gamma = 2.16 \pm 0.05$  (cluster m).

intensity. Pixels characterized by parameters  $\theta_F$  and  $\gamma$  within the cutoffs are then superimposed on the morphological image with a color code.

**Validation through measurements.**  $\mu$ MAPPs was validated at first on the intense SHG signals produced by the highly organized and regular fibers of the mouse-tail tendon<sup>20,32,54</sup>. Figure 3a reports the maximum intensity projection images for a native mouse tendon. The  $\theta$ -map and  $\gamma$ -map (Figs. 3b,c) were obtained by applying  $\mu$ MAPPs to the  $\theta$ - and  $\gamma$ p-plots (Figs. 3f,g). The  $\gamma$  values were obtained by projecting the P-SHG spectra onto the reference curve of the  $\gamma$ p-plot (calculated with  $\Delta\theta = 10^\circ$  and  $\theta_F$  retrieved from the first phasor plot) as in Equation (5).

As highlighted in Fig. 3, qualitative information on the microstructure heterogeneity can be obtained also without assuming any theoretical model (RC): uniform regions (orange box in Fig. 3a) correspond to a small compact cloud of points (orange in Figs. 3f,g) in the  $\theta$ - and  $\gamma$ p-plots, while microscopic non-homogeneity (green

box in Fig. 3a) results in clouds of scattered points (green in Figs. 3f,g). A mean fibrils angle  $\langle\theta_f\rangle = 116^\circ \pm 5^\circ$  is obtained for the more uniform region in Fig. 3a (as from Gaussian fit of the  $\theta_f$  distribution in the orange box), with low anisotropy values ( $\gamma < 2$ ). Regions that correspond in the morphological image to crimped features (Fig. 3a, green box) have a double distribution of the fibril angles, at  $\langle\theta_f\rangle = 72^\circ \pm 4^\circ$  and  $\langle\theta_f\rangle = 154^\circ \pm 4^\circ$ , and high anisotropy values ( $\gamma > 2$ ). Over the whole image in Fig. 3 the distribution of the anisotropy value is wide with  $\langle\gamma\rangle = 1.7 \pm 0.6$  (as from the Gaussian fit of the distribution in Fig. 3e) in agreement with the literature<sup>20,32,54</sup> and the distribution of the fibrils angle is peaked at  $\theta_f \cong 116^\circ$  with two secondary components at  $\theta_f \cong 80^\circ$  and at  $\theta_f \cong 155^\circ$ . A more quantitative analysis of the spatial in-homogeneity of the tissue can be obtained starting from the application of our clustering algorithm to the two coupled phasor spaces. Figure 3h shows 23 clusters encoded with different monochromatic linear LUTs (coding for the integrated SHG intensity of the pixel): the blue color highlights the main cluster, containing the 57% of the total elements (122'998) and corresponding to a uniform regions with  $\langle\theta_f\rangle = 113^\circ \pm 7^\circ$  and  $\langle\gamma\rangle = 1.7 \pm 0.2$ .

To further validate the  $\mu$ MAPPs algorithm, we compared our results with those from fitting-based procedures. The fit of the P-SHG spectra, obtained from clusters retrieved with tight cutoff values on two ROIs ( $\theta_c = 2^\circ$  and  $\gamma_c = 0.1$ , Figs. 3i,j and l,m), are reported as an example in Figs. 3k and n, showing an excellent agreement between the two methods.  $\mu$ MAPPs was then also applied to samples of mouse-tail tendon denatured via collagenase digestion, retrieving again results in agreement with the literature<sup>57</sup> (see Supplementary Note 6 and Supplementary Fig. 12).

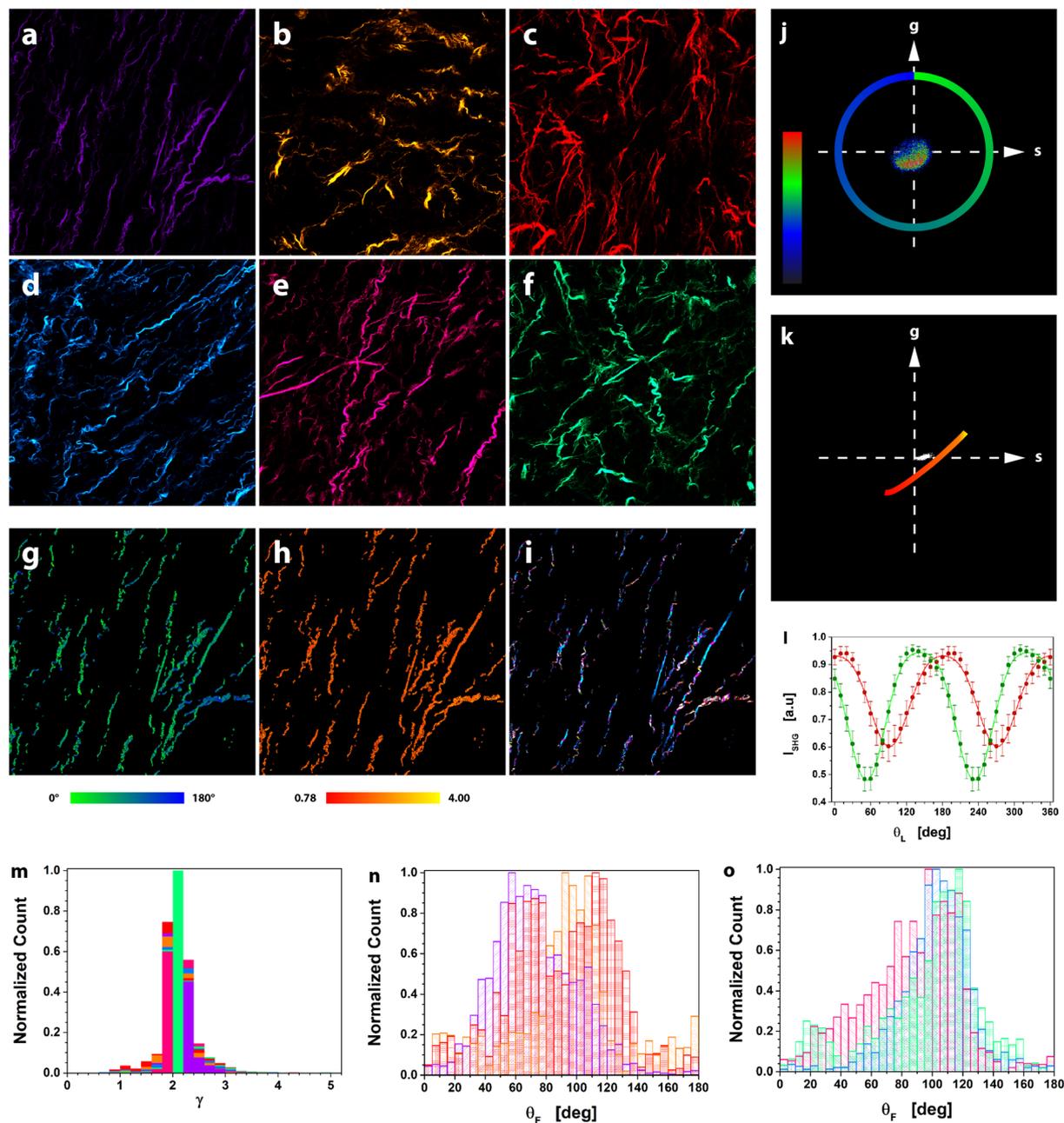
**$\mu$ MAPPs analysis of histology tumor sections.** Alterations in the ExtraCellular Matrix (ECM), whose prominent constituent is collagen, are correlated to tumor development: the increased collagen deposition in some tissues facilitates tumor initiation and progression<sup>3,4,10,14,58,59</sup>. Therefore, we evaluated the efficacy of  $\mu$ MAPPs to characterize the ECM of fixed sections of CT26 derived tumor from mice (5  $\mu$ m thickness). The average SHG signal from collagen reveals thin wavy fibers in the ECM of tumor samples explanted 7 days after inoculation, as shown in Figs. 4a–f for six fields of view (FOVs) from sections obtained from two different samples. Figures 4g,h show the  $\theta$ -map and  $\gamma$ -map of the FOV in Fig. 4a, obtained by analyzing data of the  $\theta$ p-plot and the  $\gamma$ p-plot (Figs. 4j,k) through Equations (4–5). The  $\gamma$  and  $\theta$  counts histograms related to the six FOVs are reported in Figs. 4m–o with the same color-code as in Figs. 4a–f. The retrieved anisotropy parameter spans the range [0.78–4] within each FOV (Fig. 4m), whose average  $\gamma$  values lie within [2.06  $\pm$  0.18–2.10  $\pm$  0.20], and the overall mean value is  $\langle\gamma\rangle = 2.08 \pm 0.04$ . The fibrils angle obtained within each FOV is characterized by quite large distributions with widths in the [8°–22°] range. Again, microscopic parameters can then be grouped and highlighted in the original image by using our cluster algorithm (Fig. 4i). In our analysis we found a mean of 8  $\pm$  2 clusters (cutoff values  $\theta_c = 5^\circ$  and  $\gamma_c = 0.2$ ; ET, minimum number of elements in a cluster with respect to the total number of analyzed pixels, =2%).

In the tumor samples the presence of non-negligible background values made us modify the  $\gamma$ -map retrieval algorithm (Supplementary Fig. 5). Therefore, we validated again  $\mu$ MAPPs against fitting-based procedures: Fig. 4l reports an example of the fit of the P-SHG spectra extracted from two different clusters ( $\theta_c = 2^\circ$  and  $\gamma_c = 0.1$ ), leading to parameters in excellent agreement with  $\mu$ MAPPs results. This successful validation not only demonstrates that  $\mu$ MAPPs is able to obtain microscopic parameters in agreement with fitting-based algorithms, but also that it is a versatile method that can afford deviations from the theoretical model described by Equation (3).

**$\mu$ MAPPs analysis of tumor biopsies.** Here we report our preliminary results on tumor ECM studies in biopsies excised at day 7 and 12 after B16 melanoma cells implantation in the left flank of mice. We found that the tumor structure is heterogeneous throughout the sample, depending on the depth from the tumor surface and the size of the selected FOV. At day 12 the collagen fibers appear straight below the tumor cap, whereas they are curly at the surface. At day 7 fibers are wavier, shorter and less dense with respect to day 12. Due to this heterogeneous environment, we chose to perform our experiments at a fixed z position range below the tumor surface (10–15  $\mu$ m). Two exemplary images of a tumor at day 7 and 12, out of 30 FOVs from 3 biopsies, are shown in Figs. 5a and e, respectively. By applying  $\mu$ MAPPs, we extracted the  $\theta_f$  and  $\gamma$  maps for the tumor at day 7 (Figs. 5b,c) and at day 12 (Figs. 5f,g). The  $\gamma$  distributions (Fig. 5i) computed over all the experiments at day 7 (purple) and the day 12 (green) show no change. In each FOV,  $\gamma$  is variable in the range [0.9–4], while the mean values obtained by averaging over the different FOVs, are:  $\langle\gamma\rangle = 2.11 \pm 0.04$  (day 7) and  $\langle\gamma\rangle = 2.09 \pm 0.04$  (day 12) (Fig. 5j). The width of the fibrils angular distribution within each FOV lies in the range [8°–19°] at day 7 (Fig. 5k), while it is larger at day 12 (Fig. 5l), variable within [13°–26°], as obtained by a multi-Gaussian fit.

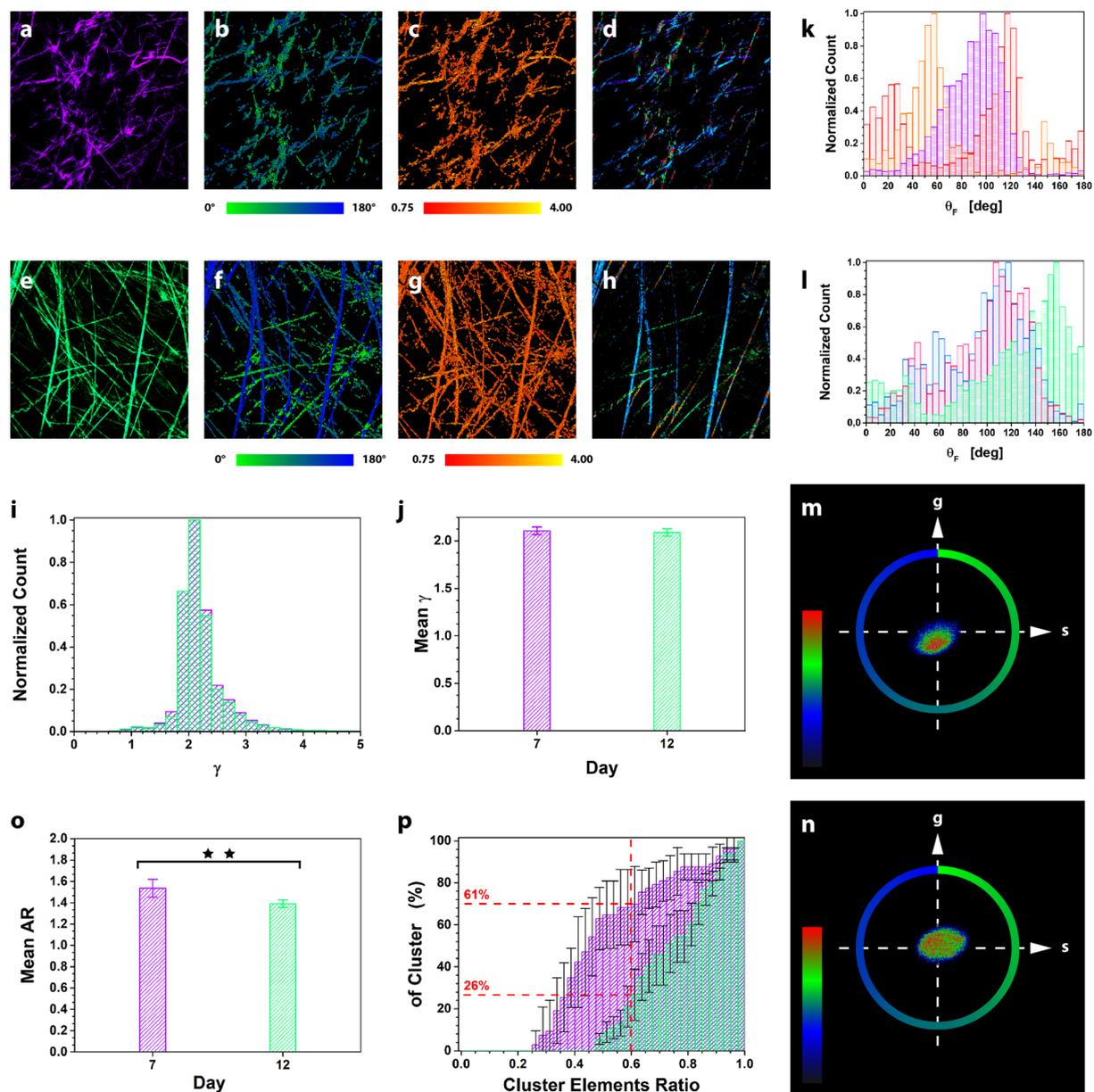
Differences in the microscopic organization of the investigated samples can be also extracted by a global analysis of the shape of the cloud of points within the  $\theta$ p-plot (Figs. 5m,n). Although the distribution of the  $\gamma$  values is not changing between day 7 and day 12 (Fig. 5i), different shapes of the  $\theta$ p-plots for the two time-points are retrieved and characterized by measuring (Fig. 5o) the ratio (AR) between the major and minor axis of the ellipse enclosing the points in the phasor plot. At day 12,  $\langle AR \rangle = 1.39 \pm 0.03$ , while at day 7  $\langle AR \rangle = 1.53 \pm 0.08$ , indicating a more isotropic fibrils angular distribution within each FOV at day 12. The phasor analysis is therefore able to obtain results in agreement with Figs. 5k,l, and can help in highlighting differences between the samples faster than histogram construction and peak fitting procedures.

We further characterize the tumor biopsies by applying the clustering algorithm. First, we assumed large clustering parameters ( $\theta_c = 30^\circ$  and  $\gamma_c = 1$  ET = 10%) in order to evaluate the mean Cluster Elements Ratio (CER, ratio between the number of elements in each retrieved cluster and in the most populated cluster) and to single out the presence of a main cluster. At day 12 the  $\langle CER \rangle$  is  $0.54 \pm 0.14$ , while  $\langle CER \rangle = 0.14 \pm 0.14$  at day 7, indicating the presence of a major cluster at day 7. Clusters obtained with these large cutoff parameters are reported in Figs. 5d and h.



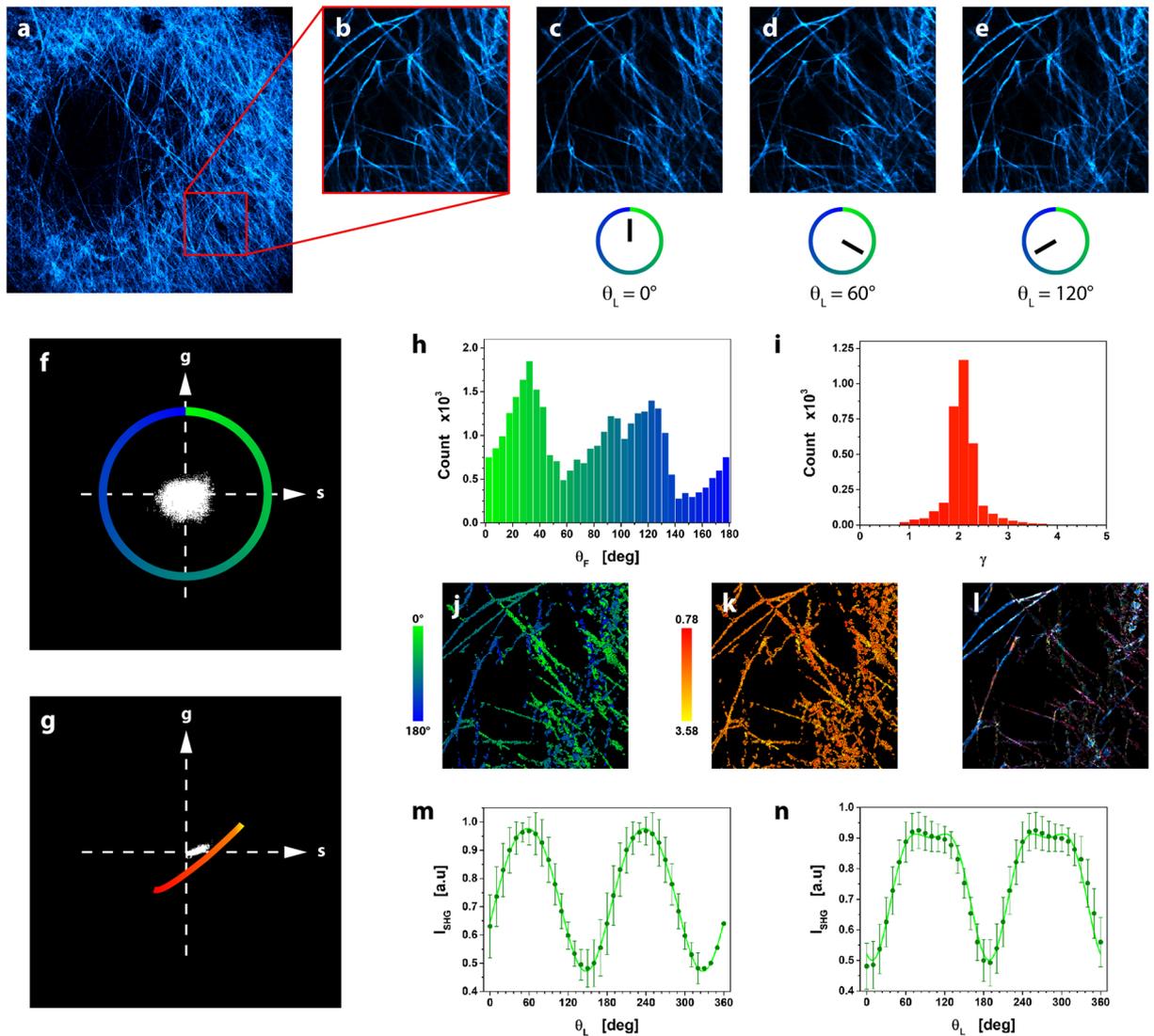
**Figure 4.**  $\mu$ MAPPs analysis of tumor biopsies. (a–f) Maximum intensity projection of the image stack related to six different FOVs ( $\Delta\theta = 10^\circ$ ). Panels (a–c) are related to three different FOVs of a tumor section, while panels (d–f) show three FOVs from a second tumor section. Image size:  $283 \times 283 \mu\text{m}^2$ . (g and h) show the  $\theta$ -map and the  $\gamma$ -map of the FOV in (a). (i) Result of the clustering procedure on the FOV in (a). 10 clusters were obtained with cutoff values  $\theta_C = 5^\circ$  and  $\gamma_C = 0.2$ ; ET = 2%. Each monochromatic LUT encodes for a cluster with an intensity scaling as the pixel integrated P-SHG signal. (j and k) represent, respectively, the  $\theta$ p-plot and  $\gamma$ p-plot of the FOV in (a). The color scale encodes for the counts per phasor plot pixel. (l) The average P-SHG spectra related to two clusters (cutoff values  $\theta_C = 2^\circ$  and  $\gamma_C = 0.1$ ) were fit to the model Equation (3), with the addition of the background  $\gamma_0$  term (lines). The best fit parameters are:  $\theta_F = 143.2^\circ \pm 0.4^\circ$ ,  $\gamma = 2.11 \pm 0.03$ ,  $\gamma_0 = 0.35 \pm 0.01$ ,  $k = 0.13 \pm 0.01$  (green) and  $\theta_F = 1.5^\circ \pm 0.6^\circ$ ,  $\gamma = 2.21 \pm 0.06$ ,  $\gamma_0 = 0.52 \pm 0.01$ ,  $k = 0.08 \pm 0.01$  (red). The  $\mu$ MAPPs results are:  $\theta_F = 143.2^\circ \pm 0.3^\circ$ ,  $\gamma = 2.18 \pm 0.03$ , (green) and  $\theta_F = 3.0^\circ \pm 0.6^\circ$ ,  $\gamma = 2.19 \pm 0.03$  (red). The threshold noise level for all analyses was  $T_{\text{NL}} = 200$ . (m) shows the  $\gamma$  counts histograms related to the six FOVs in (a–f), with the same color code. (n and o) show the  $\theta$ -histograms for the FOVs in (a–c) and (d–f), respectively. Each histogram is shown with the same color code of the related image.

Small scale heterogeneity in the tissue can be singled out by looking for clusters with few elements. To this purpose, we clustered the data set with more tight conditions ( $\theta_C = 5^\circ$  and  $\gamma_C = 0.2$ ; ET = 2%), obtaining the results summarized by the cumulative distributions of Fig. 5p. At day 7 the tumors are characterized by a higher



**Figure 5.**  $\mu$ MAPPs analysis of tumor biopsies. (a and e) show, respectively, the maximum intensity projection of the SHG image stack for a tumor biopsy at day 7 and at day 12 out of a total 30 FOVs (3 biopsies) analyzed. Image size:  $225 \times 225 \mu\text{m}^2$ . (b,c)  $\theta$ - and  $\gamma$ -maps of the image in (a). (d) Result of the clustering of the tumor biopsy at day 7 in (a) with cutoff parameters:  $\theta_C = 30^\circ$ ,  $\gamma_C = 1$ ,  $ET = 10\%$ ,  $T_{NL} = 200$ . The P-SHG intensity of the pixels of a cluster is encoded by a monochromatic LUT. (f,g)  $\theta$ - and  $\gamma$ -maps of the image in (e). (h) Clustering procedure ( $\theta_C = 30^\circ$ ,  $\gamma_C = 1$ ,  $ET = 10\%$ ,  $T_{NL} = 200$ ) for the tumor biopsy at day 12 in (e). (i) The  $\gamma$  counts histograms related to all our experiments performed at day 7 (purple) and day 12 (green) are superimposed. (j) shows the mean  $\gamma$  value for all the samples at day 7 (purple),  $\langle \gamma \rangle = 2.11 \pm 0.04$ , and at day 12 (green),  $\langle \gamma \rangle = 2.09 \pm 0.04$ . (k and l) report the  $\theta$ -count histograms for the biopsies excised at day 7 and day 12, respectively. Each color codes for a different FOV selected in a tumor biopsy at day 7 and at day 12. (m and n) show the two  $\theta$ p-plots for the FOVs reported respectively in (a and e). The color scale encodes for the counts per phasor plot pixel. (o) Ratio (AR) between the major and minor axis of the ellipse enclosing the points for tumors at day 7 (purple) and 12 (green). Note that two asterisks (\*\*\*) have the significant meaning  $0.001 < p < 0.01$ , based on two-tailed Student's t-test. (p) Cumulative distribution of the number of clusters (reported as percentage) as a function of the ratio between the number of elements in each retrieved cluster and the most populated cluster (CER).

number of clusters with small CER with respect to day 12. In fact, the 60% of clusters at day 7 (purple) have a CER value below 0.5, while this percentage is reduced to 26% at day 12 (green). These results suggest the presence of a higher number of clusters with few elements at day 7 with respect to day 12, characterized instead by clusters with a more uniform population.



**Figure 6.**  $\mu$ MAPPs analysis of tumor *in-vivo*. (a) A stack of images ( $564 \times 564 \mu\text{m}^2$ ) comprising 30 z-planes acquired at  $1 \mu\text{m}$  z spacing, from 10 to  $40 \mu\text{m}$  below the tumor ECM surface, have been z-projected in the image for  $\theta_L = \pi$ . (b) ROI for *in-vivo*  $\mu$ MAPPs analysis ( $225 \times 225 \mu\text{m}^2$ ). (c–e) Three images ( $\theta_L = 0^\circ$ ,  $\theta_L = 60^\circ$  and  $\theta_L = 120^\circ$ ) corrected for motion artifacts (see methods). (f and g) are the  $\theta$ p-plot and  $\gamma$ p-plot computed on ROI in (b). (h and i) report the corresponding counts histograms. (j and k) represent the  $\theta$  and  $\gamma$  maps obtained from (f and g), respectively. (l) Output of the clustering algorithm on the ROI in (b) with  $\theta_C = 5^\circ$ ,  $\gamma_C = 0.2$ ,  $ET = 2\%$ ,  $T_{NL} = 300$ . The P-SHG intensity of the pixels of a cluster is encoded by a monochromatic LUT. (m and n) Average P-SHG spectra for two clusters (obtained with  $\theta_C = 2^\circ$ ,  $\gamma_C = 0.1$  and  $T_{NL} = 200$ ), fit to the model Equation (3) with the addition of the background  $y_0$  term (lines). The best-fit parameters are:  $\gamma = 2.57 \pm 0.03$ ,  $\theta_F = 57.6^\circ \pm 0.2^\circ$ ,  $y_0 = 0.38 \pm 0.01$ ,  $k = 0.09 \pm 0.01$  (cluster m);  $\gamma = 1.83 \pm 0.04$ ,  $\theta_F = 98.3^\circ \pm 0.7^\circ$ ,  $y_0 = 0.33 \pm 0.02$ ,  $k = 0.17 \pm 0.01$  (cluster n), in agreement with the  $\mu$ MAPPs values that are  $\gamma = 2.39 \pm 0.03$ ,  $\theta_F = 56.4^\circ \pm 0.3^\circ$  (cluster m) and  $\gamma = 1.90 \pm 0.04$ ,  $\theta_F = 95.2^\circ \pm 1.1^\circ$  (cluster n).

**$\mu$ MAPPs *in-vivo* analysis of tumors.** In order to demonstrate that our method is robust also against possible artifacts affecting intravital experiments,  $\mu$ MAPPs was also applied to study the micro-structure of ECM in developing tumors directly in living-anesthetized animals. Figure 6a reports the maximum intensity z-projection ( $10$  to  $40 \mu\text{m}$  below the tumor surface) of a SHG image stack ( $564 \times 564 \mu\text{m}^2$ ,  $\theta_L = 2\pi$ ) in the tumor ECM *in-vivo*. A ROI (Fig. 6b) was then selected for  $\mu$ MAPPs analysis (see Figs. 6c–e for an example of motion-corrected images). The  $\theta$  and  $\gamma$  maps (Figs. 6j,k) and histograms (Figs. 6h,i) were obtained from the corresponding  $\theta$ p-plot (Fig. 6f) and  $\gamma$ p-plot (Fig. 6g) by taking into account also the background effect (finite  $y_0$  value). The Gaussian fit of the anisotropy distribution (Fig. 6i) in the selected FOV gave  $\langle \gamma \rangle = 2.07 \pm 0.19$ , in agreement with biopsies, indicating that both experimental conditions can be used for collagen characterization. Moreover, clusters retrieved with the cutoff parameters  $\theta_C = 5^\circ$  and  $\gamma_C = 0.2$  ( $ET = 2\%$ ), are shown in Fig. 6l. A close agreement between fitting and  $\mu$ MAPPs results demonstrates the validity of our method in the evaluation of microstructural parameters also *in-vivo* (Figs. 6m,n).

## Discussion

$\mu$ MAPPS is based on the Fourier projection of the P-SHG stacks of images onto two concatenated phasor plots from which, by a simple geometrical construction and the assumption of a molecular model, we can extract distributions of the collagen fibrils angle,  $\theta_F$ , and of the anisotropy parameter,  $\gamma$ . The robustness of  $\mu$ MAPPS in retrieving  $\theta_F$  and  $\gamma$  was numerically tested for the effect of finite data sampling, of SHG signal noise and of microscopic distributions of  $\theta_F$  and  $\gamma$ . Denoising filters working in the phasor space were also applied. Moreover, the results obtained by  $\mu$ MAPPS were validated against direct fitting of the P-SHG spectra in various samples. We adopted a widespread<sup>20,22,27,43,60</sup> molecular model to assist our phasor plots analysis. Indeed, our successful application of  $\mu$ MAPPS to tumor tissues, where non-idealities (background,  $y_0$ , up to 30% of the SHG signal) are present, suggests that our algorithm can be effectively extended to more complex situations taking into account axial polarization component, experimental errors and deviations from the present theoretical model<sup>40</sup>. With respect to existing P-SHG analysis routines<sup>10,22,26,27,29</sup>,  $\mu$ MAPPS provides a fast and intuitive representation of the structural content of the sample by working globally on the image and by grouping the pixels through non-spatial features they share with other pixels.

We applied here  $\mu$ MAPPS to characterize the microscopic behavior of both organized tissues, such as tendons, and of the more complex tumor ECM architecture. In this regard, three different tissue preparations (fixed sections, biopsies, *in-vivo*) have been selected for a first validation of the method. A more in depth study has been performed in tumor biopsies at two stage of progression by exploiting also geometrical features of the phasor plots and our clustering algorithm.

Our main goal was to develop and validate our phasor approach on a number of pathological relevant samples. A more refined analysis of the  $\mu$ MAPPS outcome in dependence of the tissue parameters, i.e. depth position, tumor species, and FOV dimensions, will be our next step in setting up a P-SHG based histopathological analysis. Here, we already pointed out the presence of a number of possible parameters for the micro-structure characterization of tissues: the geometrical features of the phasor plot (AR) and the CER factor in the cluster analysis, are two examples. For example, the biopsies at day 7 (less ordered) differ substantially from those taken at day 12 (more uniform population of clusters), notwithstanding the fact that we retrieved the same anisotropy distribution at the two time-points.

In summary, our clustering algorithm, working on a phasor-based analysis of the P-SHG images, allows to automatically back-project data in the original image highlighting areas that share the same microscopic behavior ( $\theta_F$  and  $\gamma$  values within a selected threshold) and to obtain other quantitative global characterization of the microscopic architecture of tissues that can be amenable to define scores for the automatic recognition of pathologies both *ex-vivo* and *in-vivo*. We believe therefore that our method could perform fast microstructural analyses that assist or replace the histo-pathological evaluation and could open future applications for *in-situ* diagnosis of pathologies and diseases. Efforts are ongoing to develop cluster-based machine learning algorithms with diagnostic capability of tumors.

## Methods

**Two-photon microscopy set-up.** The optical setup was built around a confocal scanning head (FV-300, Olympus, Japan) mounted on an upright optical microscope (BX51, Olympus, Japan), coupled to a tunable fs-pulsed Ti:Sa laser (690–980 nm, 120 fs full width at half maximum pulse duration and 80 MHz repetition frequency, Mai Tai HP, Spectra Physics, CA). A high working distance objective (NA = 0.95, WD = 2 mm, 20X, water immersion, XLUMPlan FI, Olympus, Japan) has been employed to both excite and collect the emitted signal. The backscattered second harmonic signal, primed by 800 nm wavelength, has been steered to a home-built non-descanned unit, filtered by a short-pass 670 nm filter (Chroma Inc., Brattleboro, VT, USA) and by a 400/20 nm band-pass filter (Chroma Inc., Brattleboro, VT, HQ400/20), and collected by a photomultiplier tube (HC125-02, Hamamatsu, Japan). The entire microscope is surrounded by a custom made thermostatic cabinet in which the temperature is kept at 37 °C (air thermostating by “The Cube”, Life Imaging Services, Basel, CH) for *in-vivo* measurements. For polarization-resolved SHG measurements, a half-wavelength waveplate has been placed along the optical path to control the laser polarization.

**Image acquisition and analysis.** Sequences of images of single z plane have been acquired by rotating the laser polarization from 0° to 180° in steps of 5° in order to obtain a rotation of the laser polarization from 0° to 360° in steps of 10°. Each image has a resolution of 512 × 512 pixels. The acquired field of view is described in the images captions (variable between 200 × 200  $\mu\text{m}^2$  and 600 × 600  $\mu\text{m}^2$ ). The following excitation laser power, measured before the scanning-head, has been exploited:  $P_{\text{exc}} = 10$  mW for normal mouse-tail tendon,  $P_{\text{exc}} = 20$  mW for mouse-tail tendon in case of collagenase-digestion (due to the reduced emitted SHG signal). Regarding tumors,  $P_{\text{exc}} = 50$  mW for histology sample analysis, while it was variable in the range 50–60 mW for biopsies and *in-vivo* experiments. Higher power (above 60 mW) have been used only during z-scan imaging deep within the tumor. P-SHG data have been acquired in the range 10–15  $\mu\text{m}$  below the tumor surface. Care must be taken in the proper laser power selection in order to avoid tissue damages. Each image is the result of 3–5 Kalman average scans (512 × 512 pixels) and has been acquired in 3.4–5.6 s. For *in-vivo* P-SHG experiments in tumors, a sequence of 6 z-planes has been acquired along the z axis at 1  $\mu\text{m}$  inter-distance for each incident laser polarization in order to be able to post-process the images in case of z movements. In this case 2 Kalman average scans have been acquired for each angular step. Specific corrections to the  $\mu$ MAPPS algorithm for birefringence, polarization cross-talk or diattenuation<sup>36</sup> were not applied to our P-SHG data, collected from tissue surface (10–15  $\mu\text{m}$  below the surface).

For tumor analysis, in order to further reduce the noise and retrieve the  $\theta_F$  and  $\gamma$  values with an increased precision, a gaussian filter of the image stack acquired as function of  $\theta_i$  has been implemented before  $\mu$ MAPPS analysis. This procedure, equivalent to a smoothing of the SHG intensity profile in dependence of  $\theta_i$ , allows us to analyze the acquired image stack pixel by pixel to retrieve  $\theta_F$  and  $\gamma$  without the need of signal averaging among adjacent pixels, which can lead to a worse  $\theta_F$  and  $\gamma$  maps resolution and to pixels-averaged behavior that could hide microstructural variations among them.

Moreover, following  $\mu$ MAPPs analysis, color-coded images are created to represent in each pixel the  $\theta_F$  and  $\gamma$  values retrieved from the  $\theta_F$  and  $\gamma$  p-plots, respectively.

Regarding the clustering procedure (see Supplementary Note 5), each cluster is encoded with different monochromatic LUT with an intensity scaling as the pixel integrated P-SHG signal.

The cutoff values for  $\gamma$  and  $\theta_F$  used for performing the clustering analysis of the experimental data are reported in the caption of the corresponding figures. Moreover only clusters with a number of elements higher with respect a threshold (ET), computed as a percentage of the total number of analyzed pixels, have been considered and reported in the figures captions. Only P-SHG spectra above a threshold ( $T_{NL}$ ) are analyzed in order to discard the background contribution.

In order to extract, for each cluster, the mean SHG signal as a function of  $\theta_L$  for the fit procedure, the signal related to each pixel of the P-SHG stack (belonging to the considered cluster) is divided for its maximum intensity and then the average of these signals is computed.

In the  $\theta$ -plot, the ratio between the major and minor axis of the ellipse enclosing the points, AR, has been extracted by means of the *Measure* tool of ImageJ (U.S. National Institute of Health, Bethesda, Maryland, USA).

**Software.** All the polarization-dependent analysis based on the phasor approach, the  $\theta$  and  $\gamma$  plots, maps and histograms have been obtained by means of a custom designed C++ based software. The program Origin (Origin 8.5, OriginLab Corporation) has been employed for the polarization-dependent SHG curves fitting. All the acquired images have been visualized and linearly contrast-adjusted using ImageJ (U.S. National Institute of Health, Bethesda, Maryland, USA). Images affected by xyz drift or to other movements during *in-vivo* image acquisition have been manually post-processed and realigned by means of ImageJ and Photoshop software.

**Mice.** All mice inoculated with the murine tumor cell line B16, housed under specific pathogen-free conditions, were female on a C57bl/6 background for at least 12 generations and were used at 7–9 weeks of age. All mice, inoculated with the tumor cell line CT26, were female on a BALB/c background for at least 12 generations and were used at 7–12 weeks of age. The mice were kept in a pathogen-free conventional animal house facility. The animal house is run by professional employees fully equipped with state-of-the-art instrumentation in order to maintain the standard of animal welfare at the maximum levels. All mice were housed in individual, ventilated cages with 12 h light/dark cycles with food and water ad libitum. Experiments were performed using protocols approved by the Institutional Animal Care and Use Committee of the University of Milano-Bicocca and by the Italian Ministry of Health or were performed under Boston Children's Hospital IACUC approved protocols.

**Mouse-tail tendon.** Mouse-tail tendons have been harvested from sacrificed C57bl/6 background mice at 8 weeks of age. The tendons have been cut into strips 3–4 cm in length and placed in a Phosphate Buffered Saline (PBS) solution at pH = 7.4 before visualization under the two-photon microscope. Then, the samples have been mounted in a chamber made of two cover-slips sealed with silicone and filled with PBS.

For collagenase-digestion, highly purified collagenase has been employed (Collagenase VIII Sigma-Aldrich C2139) in the buffer Hepes-NaOH 10 mM pH 7.4/NaCl 150 mM/KCl 5 mM/MgCl<sub>2</sub> 1 mM/CaCl<sub>2</sub> 5 mM, at a concentration of 5 mg/ml. Mouse-tail tendons have been incubated in the collagenase solution for 1 h at 37 °C; for SHG visualization the samples have been placed in a chamber filled with PBS solution. SHG polarization-dependent images have been acquired at the surface of the mouse-tail tendons in order to eliminate or reduce effects due to refractive-index-induced spherical aberrations.

10 FOVs ( $225 \times 225 \mu\text{m}^2$ ) related to 4 tendons of different mice have been analyzed.

**Cells.** The murine tumor cell line B16 and the CT26 cell line were cultured in IMDM-10 complete medium: IMDM, 10% heat-inactivated FBS (EuroClone), 2 mM l-glutamine, 100 U/ml penicillin, 100  $\mu\text{g}/\text{ml}$  streptomycin. Cells at the confluence of 70% were collected for the mice injection.

CT26 cell line (ATCC, CRL2638), a BALB/c mouse colon carcinoma, was obtained from Mario Colombo (Istituto Nazionale del Tumori, Milan, Italy).

**Tumor injection and analysis.** For the *in vivo* and *ex-vivo* B16 tumor analysis, C57bl/6 mice were inoculated in the deep derma in the left flank with the tumorigenic dose of B16 tumor cells ( $2 \times 10^6$ ) at Day 0. For the *in vivo* experiment, at day 7 mice were anesthetized I.P. with a Ketamine/xilazine cocktail (0.1 mL/20 g mouse), the derma was cut off and the tumor exposed and analyzed with two-photon microscopy. For the *ex-vivo* experiments at day 7 and day 12, mice were euthanized, the tumor collected and placed in PBS solution before polarized-SHG analysis. For the *ex-vivo* (biopsies) experiments, tumors have been placed in a chamber made of two cover-slips sealed with silicone and filled with PBS solution.

For the CT26 tumor analysis, BALB/c mice were inoculated in the deep derma in the left flank with the minimal tumorigenic dose of CT26 tumor cells ( $5 \times 10^4$ ) at Day 0. Explanted tumors at Day 7 were embedded in OCT freezing media (Biooptica). Sections ( $5 \mu\text{m}$ ) were cut on a Cryostat, adhered to Superfrost Plus slide (Thermo Scientific), and then imaged under the two-photon excitation microscope.

Mice subjected to tumor injection were monitored on a daily basis for signs of discomfort, including hunched posture, ruffled fur and lack of movement within the cage. The body condition score index (a qualitative assessment of an animal's overall appearance based on its weight, muscle mass and bone prominence) was used to evaluate the welfare of the mice. Generally, mice did not present signs of distress given the short period of time between tumor injection and killing.

Regarding experiments in histology fixed sections, during our analysis we acquired five field of views (FOVs,  $283 \times 283 \mu\text{m}^2$ ) for each analyzed sample (3 tumors, 3 sections per tumor).

Regarding biopsies we acquired 10 FOVs ( $225 \times 225 \mu\text{m}^2$ ) for each analyzed tumors ( $n = 3$ ) both at day 7 and at day 12 after tumor inoculation.

Results from *in-vivo* microscopy are related to 2 FOVs in 2 different animals.

**Statistical method.** Unless otherwise stated, results are expressed as mean  $\pm$  SEM. All statistical analyses were performed in Prism 5 (GraphPad Software). Means between two groups were compared with two-tailed Student's t-test. The value of  $P \leq 0.05$  was considered statistically significant. The degree of significance was assigned as follows: \*for  $P \leq 0.05$ , \*\*for  $P \leq 0.01$ , \*\*\*for  $P \leq 0.001$ , and \*\*\*\*for  $P \leq 0.0001$ .

**Data availability.** The data that support the findings of this study are available from the corresponding authors on reasonable request.

**Code availability.** The custom-written C++ software will be made available online at <http://fisica.mib.infn.it/media/homepages/biofisica/bicocca-hepht/index.html>.

## References

- Buehler, M. J. Nature designs tough collagen: explaining the nanostructure of collagen fibrils. *Proc. Natl. Acad. Sci. USA* **103**, 12285–12290 (2006).
- Cicchi, R. *et al.* From molecular structure to tissue architecture: collagen organization probed by SHG microscopy. *J. Biophotonics* **6**, 129–142 (2013).
- Nadiarykh, O., LaComb, R. B., Brewer, M. A. & Campagnola, P. J. Alterations of the extracellular matrix in ovarian cancer studied by Second Harmonic Generation imaging microscopy. *BMC Cancer* **10**, 94 (2010).
- Raja, A. M. *et al.* Differential remodeling of extracellular matrices by breast cancer initiating cells. *J. Biophotonics* **8**, 804–815 (2015).
- Keikhosravi, A., Bredfeldt, J. S., Sagar, A. K. & Eliceiri, K. W. Second-harmonic generation imaging of cancer. *Methods Cell Biol.* **123**, 531–46 (2014).
- Liao, Y. H. *et al.* Determination of chronological aging parameters in epidermal keratinocytes by *in vivo* harmonic generation microscopy. *Biomed. Opt. Express* **4**, 77–88 (2013).
- Kirkpatrick, N. D., Andreou, S., Hoying, J. B. & Utzinger, U. Live imaging of collagen remodeling using angiogenesis. *Am. J. Physiol. Heart Circ. Physiol.* **292**, H3198–206 (2007).
- Strupler, M. *et al.* Second harmonic imaging and scoring of collagen in fibrotic tissues. *Opt. Express* **15**, 4054–65 (2007).
- Wang, Q. *et al.* Second-harmonic generation microscopy for assessment of mesenchymal stem cell-seeded acellular dermal matrix in wound-healing. *Biomaterials* **53**, 659–668 (2015).
- Kumar, R. *et al.* Polarization second harmonic generation microscopy provides quantitative enhanced molecular specificity for tissue diagnostics. *J. Biophotonics* **8**, 730–739 (2015).
- Drifka, C. R. *et al.* Comparison of Picrosirius Red Staining With Second Harmonic Generation Imaging for the Quantification of Clinically Relevant Collagen Fiber Features in Histopathology Samples. *J. Histochem. Cytochem.* **64**, 519–29 (2016).
- Plotnikov, S. V., Millard, A. C., Campagnola, P. J. & Mohler, W. A. Characterization of the myosin-based source for second-harmonic generation from muscle sarcomeres. *Biophys. J.* **90**, 693–703 (2006).
- Campagnola, P. J. & Loew, L. M. Second-harmonic imaging microscopy for visualizing biomolecular arrays in cells, tissues and organisms. *Nat. Biotechnol.* **21**, 1356–1360 (2003).
- Tilbury, K., Lien, C. H., Chen, S. J. & Campagnola, P. J. Differentiation of Col I and Col III isoforms in stromal models of ovarian cancer by analysis of second harmonic generation polarization and emission directionality. *Biophys. J.* **106**, 354–65 (2014).
- Williams, R. M., Zipfel, W. R. & Webb, W. W. Interpreting Second-Harmonic Generation Images of Collagen I Fibrils. *Biophys. J.* **88**, 1377–1386 (2005).
- Raub, C. B. *et al.* Image correlation spectroscopy of multiphoton images correlates with collagen mechanical properties. *Biophys. J.* **94**, 2361–73 (2008).
- Ambekar, R. *et al.* Quantifying collagen structure in breast biopsies using second-harmonic generation imaging. *Biomed. Opt. Express* **3**, 2021–35 (2012).
- Rao, R. A., Mehta, M. R. & Toussaint, K. C. Jr. Fourier transform-second-harmonic generation imaging of biological tissues. *Opt. Express* **17**, 14534–42 (2009).
- Chen, X., Nadiarykh, O., Plotnikov, S. & Campagnola, P. J. Second harmonic generation microscopy for quantitative analysis of collagen fibrillar structure. *Nat. Protoc.* **7**, 654–69 (2012).
- Stoller, P., Kim, B. M., Rubenchik, A. M., Reiser, K. M. & Da Silva, L. B. Polarization-dependent optical second-harmonic imaging of a rat-tail tendon. *J. Biomed. Opt.* **7**, 205–14 (2002).
- Roth, S. & Freund, I. Second harmonic generation in collagen. *J. Chem. Phys.* **70**, 1637–1643 (1979).
- Latour, G., Gusachenko, I., Kowalczyk, L., Lamarre, I. & Schanne-Klein, M. C. *In vivo* structural imaging of the cornea by polarization-resolved second harmonic microscopy. *Biomed. Opt. Express* **3**, 1–15 (2012).
- Hristu, R., Stanciu, S. G., Tranca, D. E. & Stanciu, G. A. Improved quantification of collagen anisotropy with polarization-resolved second harmonic generation microscopy. *J. Biophotonics* **10**, 1171–1179 (2017).
- Rouède, D. *et al.* Determination of extracellular matrix collagen fibril architectures and pathological remodeling by polarization dependent second harmonic microscopy. *Sci Rep.* **7**, 12197 (2017).
- Schürmann, S., von Wegner, F., Fink, R. H., Friedrich, O. & Vogel, M. Second harmonic generation microscopy probes different states of motor protein interaction in myofibrils. *Biophys. J.* **99**, 1842–51 (2010).
- Amat-Roldan, I., Psilodimitrakopoulos, S., Loza-Alvarez, P. & Artigas, D. Fast image analysis in polarization SHG microscopy. *Opt. Express* **18**, 17209–19 (2010).
- Nucciotti, V. *et al.* Probing myosin structural conformation *in vivo* by second-harmonic generation microscopy. *PNAS* **107**, 7763–7768 (2010).
- Campagnola, P. J. & Dong, C. Y. Second harmonic generation microscopy: principles and applications to disease diagnosis. *Laser & Photon. Rev.* **5**, 13–26 (2011).
- Su, P. J. *et al.* Discrimination of collagen in normal and pathological skin dermis through second-order susceptibility microscopy. *Opt. Express* **17**, 11161–11171 (2009).
- Psilodimitrakopoulos, S. *et al.* Quantitative imaging of microtubule alteration as an early marker of axonal degeneration after ischemia in neurons. *Biophys. J.* **104**, 968–975 (2013).
- Ait-Belkacem, D. *et al.* Microscopic structural study of collagen aging in isolated fibrils using polarized second harmonic generation. *J. Biomed. Opt.* **17**, 080506 (2012).
- Gusachenko, I., Latour, G. & Schanne-Klein, M. C. Polarization-resolved Second Harmonic microscopy in anisotropic thick tissues. *Opt. Exp.* **18**, 19339–19352 (2010).
- Chu, S. W. *et al.* Studies of  $\chi(2)/\chi(3)$  tensors in submicron-scaled bio-tissues by polarization harmonics optical microscopy. *Biophys. J.* **86**, 3914–3922 (2004).
- Brasselet, S. Polarization resolved nonlinear microscopy: application to structural molecular and biological imaging. *Adv. Opt. Photon.* **3**, 205–271 (2011).
- Tuer, A. E. *et al.* Nonlinear optical properties of type I collagen fibers studied by polarization dependent second harmonic generation microscopy. *J. Phys. Chem. B* **115**, 12759–12769 (2011).

36. Tokarz, D. *et al.* Ultrastructural features of collagen in thyroid carcinoma tissue observed by polarization second harmonic generation microscopy. *Biomed. Opt. Express* **6**, 3475–3481 (2015).
37. Lien, C. H., Tilbury, K., Chen, S. J. & Campagnola, P. J. Precise, motion-free polarization control in Second Harmonic Generation microscopy using a liquid crystal modulator in the infinity space. *Biomed. Opt. Express* **4**, 1991–2002 (2013).
38. Psilodimitrakopoulos, S., Loza-Alvarez, P. & Artigas, D. Fast monitoring of *in-vivo* conformational changes in myosin using single scan polarization-SHG microscopy. *Biomed. Opt. Express* **5**, 4362–73 (2014).
39. Turcotte, R., Mattson, J. M., Wu, J. W., Zhang, Y. & Lin, C. P. Molecular Order of Arterial Collagen Using Circular Polarization Second-Harmonic Generation Imaging. *Biophys. J.* **110**, 530–533 (2016).
40. Psilodimitrakopoulos, S. *et al.* *In vivo*, pixel-resolution mapping of thick filaments orientation in nonfibrillar muscle using polarization-sensitive second harmonic generation microscopy. *J. Biomed. Opt.* **14**, 014001 (2009).
41. Stoller, P., Reiser, K. M., Celliers, P. M. & Rubenchik, A. M. Polarization-modulated second harmonic generation in collagen. *Biophys. J.* **82**, 3330–42 (2002).
42. Tiaho, F., Recher, G. & Rouede, D. Estimation of helical angles of myosin and collagen by second harmonic generation imaging microscopy. *Opt. Express* **15**, 12286–12295 (2007).
43. Freund, I. & Deutsch, M. Second-harmonic microscopy of biological tissue. *Opt. Lett.* **11**, 94–98 (1986).
44. Odin, C., Le Grand, Y., Renault, A., Gailhouse, L. & Baffet, G. Orientation fields of nonlinear biological fibrils by second harmonic generation microscopy. *J. Microsc.* **229**, 32–38 (2008).
45. Odin, C. *et al.* Collagen and myosin characterization by orientation field second harmonic microscopy. *Opt. Express* **16**, 16151–65 (2008).
46. Chen, W. L. *et al.* Second harmonic generation  $\chi^2$  tensor microscopy for tissue imaging. *Appl. Phys. Lett.* **94**, 183902 (2009).
47. Digman, M. A., Caiolfa, V. R., Zamai, M. & Gratton, E. The phasor approach to fluorescence lifetime imaging analysis. *Biophys. J.* **94**, L14–6 (2008).
48. Fereidouni, F., Bader, A. N. & Gerritsen, H. C. Spectral phasor analysis allows rapid and reliable unmixing of fluorescence microscopy spectral images. *Opt. Express* **20**, 12729–41 (2012).
49. Ranjit, S., Lanzano, L. & Gratton, E. Mapping diffusion in a living cell via the phasor approach. *Biophys. J.* **107**, 2775–85 (2014).
50. Vergeldt, F. J. *et al.* Multi-component quantitative magnetic resonance imaging by phasor representation. *Sci. Rep.* **7**, 861 (2017).
51. Stringari, C., Nourse, J. L., Flanagan, L. A. & Gratton, E. Phasor Fluorescence Lifetime Microscopy of Free and Protein-Bound NADH Reveals Neural Stem Cell Differentiation Potential. *PLoS One* **7**, e48014 (2012).
52. Stringari, C. *et al.* *In vivo* single-cell detection of metabolic oscillations in stem cells. *Cell Reports* **10**, 1–7 (2015).
53. Wright, B. K. *et al.* Phasor-FLIM analysis of NADH distribution and localization in the nucleus of live progenitor myoblast cells. *Microsc. Res. Tech.* **75**, 1717–1722 (2012).
54. Gusachenko, I., Tran, V., Goulam Houssen, Y., Allain, J. M. & Schanne-Klein, M. C. Polarization-resolved second-harmonic generation in tendon upon mechanical stretching. *Biophys. J.* **102**, 2220–9 (2012).
55. Cutrale, F. *et al.* Hyperspectral phasor analysis enables multiplexed 5D *in vivo* imaging. *Nat. Methods* **14**, 149–152 (2017).
56. Rodriguez, A. & Laio, A. Machine learning. *Clustering by fast search and find of density peaks*. *Science* **344**, 1492–6 (2014).
57. Stoller, P. C., Reiser, K. M., Celliers, P. M. & Rubenchik, A. M. Effect of structural modification on second harmonic generation in collagen. *Proc. SPIE* **4963**, 41 (2003).
58. Kaushik, S., Pickup, M. W. & Weaver, V. M. From transformation to metastasis: deconstructing the extracellular matrix in breast cancer. *Cancer Metastasis Rev.* **35**, 655–667 (2016).
59. Bonnans, C., Chou, J. & Werb, Z. Remodelling the extracellular matrix in development and disease. *Nature Rev. Mol. Cell Biol.* **15**, 786–801 (2014).
60. Psilodimitrakopoulos, S. *et al.* Quantitative discrimination between endogenous SHG sources in mammalian tissue, based on their polarization response. *Opt. Express* **17**, 10168–76 (2009).

## Acknowledgements

The research was funded by Università degli Studi di Milano-Bicocca Academic fundings 2015 to L.S. and G.C. F.G. is supported by Associazione Italiana per la Ricerca sul Cancro (IG 2016Id.18842), Cariplo Foundation (Grant 2014-0655), and Fondazione Regionale per la Ricerca Biomedica (FRRB). I.Z. is supported by NIH grants 1R01AI121066-01A1, 1R01DK115217, CCFA Senior Research Awards, the Eleanor and Miles Shore 50th Anniversary Fellowship Program, and the Cariplo Foundation.

## Author Contributions

F.R. and L.S. conceived the  $\mu$ MAPPs method and the experiments, acquired and analyzed data. L.S., F.R., G.C. and M.C. wrote the manuscript and supervised the project. F.R. wrote the C++ code. L.D. analyzed data. F.M. and L.M. performed mouse surgery and helped in tumor-related experiments. L.S., G.C., I.Z., F.G. provided financial support for the experiments.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-017-17726-y>.

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017