

# SCIENTIFIC REPORTS



OPEN

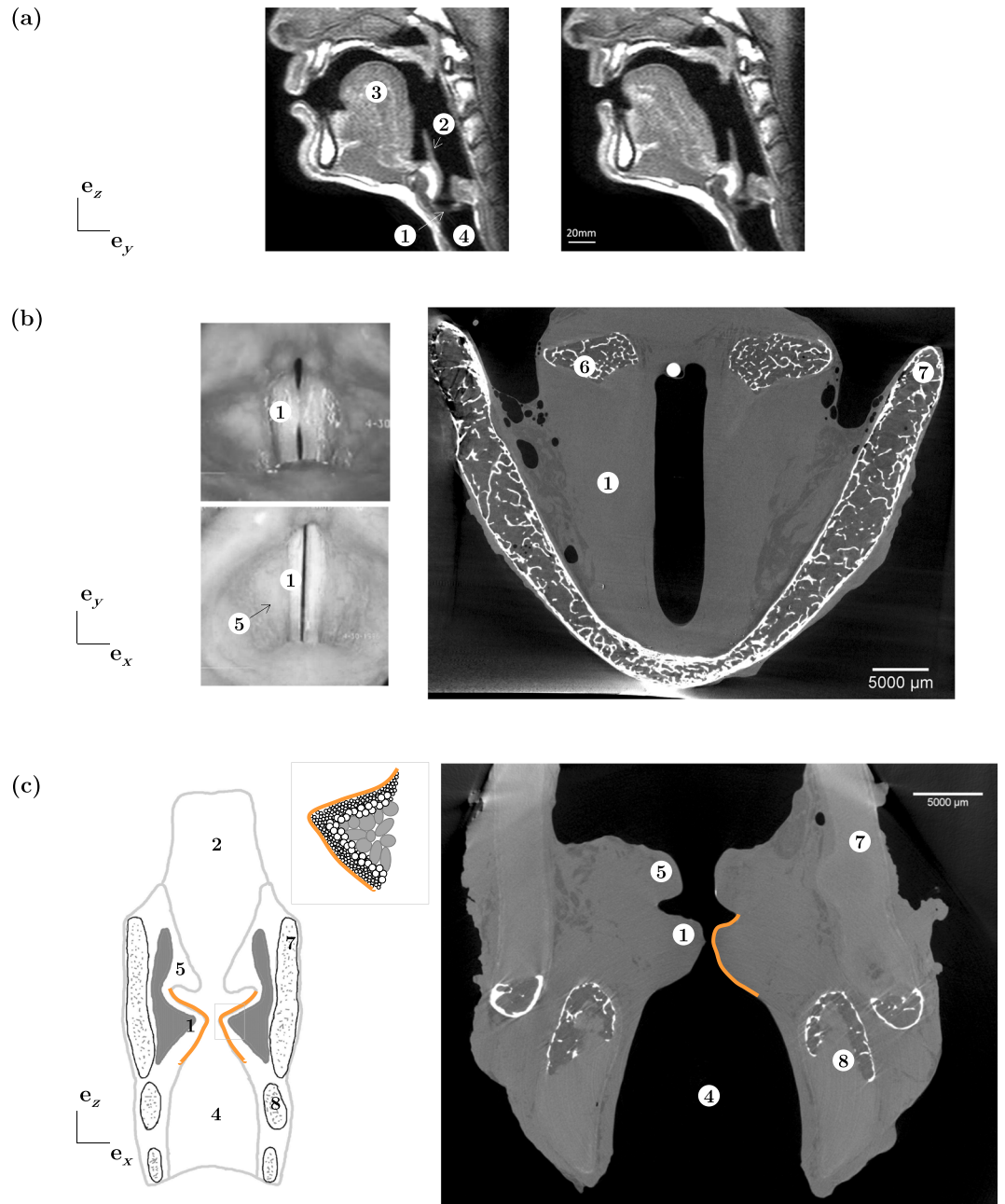
## 3D multiscale imaging of human vocal folds using synchrotron X-ray microtomography in phase retrieval mode

Lucie Bailly<sup>1</sup>, Thibaud Cochereau<sup>1,2</sup>, Laurent Orgéas<sup>1</sup>, Nathalie Henrich Bernardoni<sup>2</sup>, Sabine Rolland du Roscoat<sup>1</sup>, Anne McLeer-Florin<sup>3</sup>, Yohann Robert<sup>4</sup>, Xavier Laval<sup>2</sup>, Tanguy Laurencin<sup>1</sup>, Philippe Chaffanjon<sup>4,2</sup>, Barbara Fayard<sup>5</sup> & Elodie Boller<sup>6</sup>

Human vocal folds possess outstanding abilities to endure large, reversible deformations and to vibrate up to more than thousand cycles per second. This unique performance mainly results from their complex specific 3D and multiscale structure, which is very difficult to investigate experimentally and still presents challenges using either confocal microscopy, MRI or X-ray microtomography in absorption mode. To circumvent these difficulties, we used high-resolution synchrotron X-ray microtomography with phase retrieval and report the first *ex vivo* 3D images of human vocal-fold tissues at multiple scales. Various relevant descriptors of structure were extracted from the images: geometry of vocal folds at rest or in a stretched phonatory-like position, shape and size of their layered fibrous architectures, orientation, shape and size of the muscle fibres as well as the set of collagen and elastin fibre bundles constituting these layers. The developed methodology opens a promising insight into voice biomechanics, which will allow further assessment of the micromechanics of the vocal folds and their vibratory properties. This will then provide valuable guidelines for the design of new mimetic biomaterials for the next generation of artificial larynges.

Human vocal folds exhibit remarkable vibro-mechanical properties, allowing them to generate an outstanding range of sounds. These laryngeal soft tissues are in the order of 1 cm long, anteriorly and posteriorly connected to the thyroid cartilage and the two arytenoid cartilages, respectively (Fig. 1). Their self-sustained vibration is induced by pulmonary airflow and leads to phonation, *i.e.*, the production of audible air pulse trains. This acoustic wave is filtered by the resonances of the vocal tract, the geometry of which is shaped by speech articulators (*e.g.*, jaw, tongue, lips) to generate distinguishable voiced sounds, such as vowels and sonorous consonants. Furthermore, during this process, laryngeal intrinsic muscles drive the vocal-fold adduction and abduction, stretching, and bulging. Vocal folds behave as non-linear and coupled oscillators, able to bifurcate towards complex regimes of vibration, responding to gradual variation of control parameters<sup>1–3</sup>: geometrical ones (*e.g.*, their distance or mean glottal width), mechanical ones (*e.g.*, their stiffness) and aerodynamical ones (*e.g.*, air pressure), to name a few. Their outstanding vibratory abilities are due to two major properties of the vocal-fold tissue<sup>2,4</sup>: (i) the ability to endure large reversible 3D deformations during phonation in the presence of numerous collisions and mechanical stresses<sup>5</sup> (typically between 10–50 % stretching during a glide or intonational variations, Fig. 1(b)); (ii) the ability to vibrate with a fundamental frequency ranging from less than 50 Hz to more than 1500 Hz, which is much higher than other biological oscillators such as the heart. These properties are mainly inherited from the complex and hierarchical structure of the vocal folds and of surrounding laryngeal muscles. Adult human vocal folds are known to possess a specific lamellar structure made of several layers<sup>6–10</sup>, from superficial to deep:

<sup>1</sup>Univ. Grenoble Alpes, CNRS, Grenoble INP, 3SR, Grenoble, F-38000, France. <sup>2</sup>Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, Grenoble, F-38000, France. <sup>3</sup>Univ. Grenoble Alpes, CHU Grenoble Alpes, CNRS, Grenoble INP, IAB, Grenoble, F-38000, France. <sup>4</sup>Univ. Grenoble Alpes, CHU Grenoble Alpes, LADAF, Grenoble, F-38000, France. <sup>5</sup>Novitom, Grenoble, F-38000, France. <sup>6</sup>ID19 beamline, ESRF - European Synchrotron Radiation Facility, CS40220, Grenoble, 38043, France. Correspondence and requests for materials should be addressed to L.B. (email: [lucie.bailly@3sr-grenoble.fr](mailto:lucie.bailly@3sr-grenoble.fr))



**Figure 1.** Human phonatory system. (a) Mid-sagittal view of the upper airways: typical IRM images obtained during production of sounds [u] and [e] respectively (male subject, source: GIPSA-lab). (b) Transverse view of the vocal folds: (left) *in vivo* videolaryngoscopic images obtained during a sound-pitch variation (male subject); (right) *ex vivo* X-ray microtomographic image ( $L_2$ ,  $V_{\text{vox}} = 25^3 \mu\text{m}^3$ ). (c) Mid-coronal view of the larynx: (left) idealised scheme and zoom on the vocal-fold fibrous microstructure; (right) *ex vivo* X-ray tomographic image ( $L_3$ ,  $V_{\text{vox}} = 25^3 \mu\text{m}^3$ ). ① Vocal fold, ② Epiglottis, ③ Tongue, ④ Trachea, ⑤ Ventricular fold, ⑥ Arytenoid cartilage, ⑦ Thyroid cartilage, ⑧ Cricoid cartilage.

- A stratified squamous non-keratinised *epithelium* (EP, thickness  $\approx 50\text{--}100 \mu\text{m}$ ), covered by mucus and involved in the underlying tissue protection and renewal.
- A loose connective tissue called *lamina propria* (LP, thickness  $\approx 1\text{--}2.5 \text{ mm}$ ), made of cells and extracellular matrix (ECM) with amorphous ground substances (e.g., hyaluronic acid) and fibrous networks (e.g., collagen Type I-III and elastin, both arranged in fibre bundles of diameter ranging from 0.1 to 20  $\mu\text{m}$ ). The LP layer is further divided into three sublayers with distinct fibre types, densities and arrangements: the superficial one (SL), also called Reinke's space, composed of loose fibrous components comparable to soft gelatin, the intermediate one (IL) primarily composed of elastin fibres, and the deep layer (DL) primarily composed of collagen fibres. As a whole, the LP contributes to the tissue's "passive" biomechanical properties and to the regulation of its water content.

- The inferior thyroarytenoid muscle (M) or *vocalis* (thickness  $\approx 7\text{--}8$  mm), innervated by the inferior laryngeal nerve and responsible for the tissue's "active" contractile properties. The inertia and stiffness of the *vocalis* vary according to the degree of contraction, to reach a given phonatory position.

Current understanding of the multiscale histological features of the vocal folds is still insufficient to make the link to their vibromechanical performance. This motivates the use of full-field 3D imaging to capture these features in human vocal folds at rest or under mechanical load:

- By contrast with other soft tissues (arteries, skin), vocal folds are not easily reachable *in vivo* by any medical 3D imaging technique (including ultrasound imaging), due to the surrounding laryngeal cartilages (Fig. 1(b,c)). Gold standard techniques used in clinics for the direct visualisation of their vibrations (high-speed cinematography, videostroboscopy)<sup>11</sup> only provide partial 2D views of the vocal fold's superior plane (Fig. 1(b)).

Recent imaging developments of major interest include optical coherence tomography which probes a tissue sample with infrared light and uses interferometric methods to detect light reflected from up to 3 mm within the tissue<sup>10,12,13</sup>. Although very promising to capture the lamellar structure of the vocal folds during *in vivo* vibration, this technique is severely limited for the characterisation of the several fibrous networks in the *lamina propria* layers and the *vocalis*, due to the limited spatial resolution of about  $10\ \mu\text{m}$  and to the strong attenuation of light beneath the *lamina propria* layer<sup>10,14</sup>.

- The 3D *ex vivo* observation of excised vocal folds at the (sub-)micron scale remains a challenge. Various imaging techniques have been tested, such as micro Magnetic Resonance Imaging<sup>15–17</sup>, multiphoton nonlinear laser scanning microscopy (NLSM)<sup>18–21</sup> and X-ray microtomography in standard absorption mode (see pre-tests in Fig. 1).

Micro Magnetic Resonance Imaging is limited by its spatial resolution (even with an ultra-high magnetic field above 7T), limiting useful voxel sizes to  $40^3\ \mu\text{m}^3$  implying that the various sublayers of the *lamina propria* cannot be observed<sup>15</sup>.

NLSM offers a smaller voxel size of  $\approx 1^3\text{--}20^3\ \mu\text{m}^3$ , allowing observations of the *lamina propria* sublayers.

In particular, NLSM allows the fibrous networks of elastin and collagen fibres to be distinguished without exogenous staining, by the combination of two-photon autofluorescence (TPAF) and second-harmonic generation (SHG) microscopies, respectively. However, NLSM techniques exhibit three major drawbacks:

(i) The TPAF and SHG signals arising from other inherently fluorescent molecules or highly ordered structures present in the vocal folds (*e.g.*, SHG sources in myosin filaments, TPAF sources in epithelial cells, fibroblasts, muscle cells...)<sup>21,22</sup>; (ii) The long scanning time preventing 3D *in situ* observation during the deformation of samples; (iii) The depth of field which is typically limited to under  $150\ \mu\text{m}$  with sub-cellular resolution in such highly scattering tissues. Recently, using a longer excitation wavelength which reduces scattering probability, third-harmonic generation microscopy of porcine vocal folds has allowed images up to  $420\ \mu\text{m}$  deep in the superficial *lamina propria* to be acquired, albeit with a limited contrast to distinguish collagen and elastin fibres<sup>20</sup>. Optical clearing of vocal-fold tissue using a glycerol-based reagent is also a promising approach to reduce light scattering and thus to increase imaging depth<sup>23,24</sup>.

X-ray microtomography with absorption mode and conical laboratory X-ray sources can reach sub-micron spatial resolutions (voxel size  $\approx 300^3\ \text{nm}^3$ ), but suffers from two main problems: long scanning times (typically above 1 h) implying high radiation doses in the sample, and low contrast between soft biological materials (*e.g.*, muscles, ECM fibres, epithelial cells which have similar X-ray attenuation coefficients) – see Fig. 1(c). These two problems can be overcome using synchrotron X-ray sources and facilities. Indeed, short scanning times can be reached<sup>25</sup> ( $\approx 0.5$  s per scan) with a sub-micron spatial resolution, enabling 3D *in situ* observations of samples during deformation. In addition, using phase retrieval imaging (PRI) modes, *e.g.*, with the Paganin method<sup>26</sup>, highly contrasted 3D images of soft biological tissues can be acquired<sup>27–29</sup>.

Thus, following the recent studies of Vågberg *et al.*<sup>27</sup> and Dudak *et al.*<sup>28</sup>, this work first addresses a purely technical question: (A) Does synchrotron X-ray microtomography with single phase retrieval yield sufficient contrast and spatial resolution to characterise the hierarchical structure of human vocal folds? To answer this question, 10 human larynges were scanned at various voxel sizes ( $\approx 13^3$  down to  $0.65^3\ \mu\text{m}^3$ ) enabling the first *ex vivo* 3D images of human vocal folds to be obtained with this technique. The relevance and the limitations of this imaging technique (particularly at the smallest scales) are discussed by comparing the obtained 3D images with 2D optical micrographs of vocal folds prepared with standard histological stainings.

Despite the above limitations, the overall success in answering (A) allows quantitative 3D analysis of the layered structure of vocal folds, together with the geometries, orientation and the arrangement of muscular and ECM fibres inside the *vocalis* and the *lamina propria*. This, in turn, immediately poses a number of fundamental measurement questions in voice biomechanics:

- (B) What are the size and spatial variation of the aforementioned layers?
- (C) What is the 3D structural anisotropy of the various fibrous networks the layers are made of, as well as its spatial variation?
- (D) What is the morphology of muscle and ECM fibres in their network?
- (E) How do these 3D and multiscale descriptors evolve during a mechanical loading?

Larynx name	Donated body		Alcohol immersion	Microtomographic imaging parameters				
	Gender	Age [y]	[C <sub>2</sub> H <sub>6</sub> O] [%]	V <sub>vox</sub> [ $\mu\text{m}^3$ ]	x <sub>c</sub> [mm]	U [kV]	I [ $\mu\text{A}$ ]	n <sub>p</sub>
L <sub>1</sub>	F	86	0	45 <sup>3</sup>	—	101	297	1 440
L <sub>2</sub>	M	94	0	25 <sup>3</sup>	—	100	300	3 000
L <sub>3</sub>	F	82	0	25 <sup>3</sup>	—	100	300	3 000
Larynx name	Gender	Age [y]	[C <sub>2</sub> H <sub>6</sub> O] [%]	V <sub>vox</sub> [ $\mu\text{m}^3$ ]	x <sub>c</sub> [mm]	$\delta$ : $\beta$	E [keV]	n <sub>p</sub>
L <sub>4</sub>	F	90	0	13 <sup>3</sup>	1 200	1 800	65	4 902
L <sub>5</sub>	F	85	0	13 <sup>3</sup>	1 200	1 800	65	4 902
L <sub>6</sub>	F	89	0	13 <sup>3</sup>	1 200	1 800	65	4 902
L <sub>7</sub>	M	81	0	13 <sup>3</sup>	1 200	1 100	65	4 902
L <sub>8</sub>	M	89	0	13 <sup>3</sup>	1 200	1 100	65	4 902
L <sub>9</sub>	F	75	[0; 30; 100]	13 <sup>3</sup>	11 000	1 200	60	4 900
L <sub>10</sub> <sup>†</sup>	F	92	direct vocal-fold dissection - see Table 2					

**Table 1.** Scanned larynges. Italic lines refer to samples imaged with a laboratory conical X-ray source with absorption imaging mode, the others being imaged with a synchrotron source. F: female, M: male, V<sub>vox</sub>: voxel size, x<sub>c</sub>: distance between sample and camera, U: generator voltage, I: scanning current intensity,  $\delta$ : $\beta$ : ratio of the dispersive and absorptive aspects of the wave-matter interaction, E: beam energy, n<sub>p</sub>: number of X-ray 2D projections. <sup>†</sup>Sample only dedicated to the “high” resolution imaging of the vocal-fold structure *i.e.*, at a voxel size of 0.65<sup>3</sup>  $\mu\text{m}^3$ .

Vocal-fold sample					Microtomographic imaging parameters					
Name	Type	Conservation			V <sub>vox</sub> [ $\mu\text{m}^3$ ]	x <sub>c</sub> [mm]	U [kV]	I [ $\mu\text{A}$ ]	n <sub>p</sub>	N
L <sub>2</sub> -S <sub>1</sub>	M + LP + EP	cryopreserved at -80 °C			15 <sup>3</sup>	—	100	100	1 440	2
L <sub>3</sub> -S <sub>1</sub>	M + LP + EP	defrosted from -20 °C			12 <sup>3</sup>	—	100	100	2 496	2
L <sub>3</sub> -S <sub>2</sub>	M + LP + EP	cryopreserved at -80 °C			12 <sup>3</sup>	—	100	100	2 496	1
Name	Type	[C <sub>2</sub> H <sub>6</sub> O] [%]	[CH <sub>2</sub> O] [%]	$\delta t$ [min]	V <sub>vox</sub> [ $\mu\text{m}^3$ ]	x <sub>c</sub> [mm]	$\delta$ : $\beta$	E [keV]	n <sub>p</sub>	N
L <sub>4</sub> -S <sub>1</sub>	M	30 <sup>†</sup>	0	5	0.65 <sup>3</sup>	40	500	19	1 995	1
L <sub>4</sub> -S <sub>2</sub>	M	30	0	5	0.65 <sup>3</sup>	40	500	19	1 995	5
L <sub>4</sub> -S <sub>3</sub>	LP + EP	30	0	5	0.65 <sup>3</sup>	40	500	19	1 995	4
L <sub>6</sub> -S <sub>1</sub>	M + LP + EP	30	0	4 140	0.65 <sup>3</sup>	40	300	19	1 499	16
L <sub>7</sub> -S <sub>1</sub>	M	0	10 <sup>†</sup>	650	0.65 <sup>3</sup>	40	500	19	1 995	1
L <sub>7</sub> -S <sub>2</sub>	LP + EP	0	10 <sup>‡</sup>	600	0.65 <sup>3</sup>	22	500	19	1 995	2
L <sub>10</sub> -S <sub>1</sub>	M + LP + EP	<b>0</b>	<b>0</b>	—	<b>0.65<sup>3</sup></b>	<b>40–400</b>	<b>300</b>	<b>19</b>	<b>1 499</b>	<b>7</b>
L <sub>10</sub> -S <sub>2</sub>	M + LP + EP	<b>50</b>	<b>0</b>	<b>7 620</b>	<b>0.65<sup>3</sup></b>	<b>40</b>	<b>300</b>	<b>19</b>	<b>1 499</b>	<b>37</b>
L <sub>10</sub> -S <sub>3</sub>	M + LP + EP	<b>70</b>	<b>0</b>	<b>5 940</b>	<b>0.65<sup>3</sup></b>	<b>40</b>	<b>300</b>	<b>19</b>	<b>1 499</b>	<b>34</b>
L <sub>10</sub> -S <sub>4</sub>	M + LP + EP	<b>100</b>	<b>0</b>	<b>[1; 7; 13]</b>	<b>0.65<sup>3</sup></b>	<b>40</b>	<b>300</b>	<b>19</b>	<b>1 499</b>	<b>3</b>

**Table 2.** Scanned vocal-fold samples S<sub>j</sub> extracted from larynx L<sub>j</sub>. Italic lines refer to samples imaged with a laboratory conical X-ray source. Bold lines refer to the imaged samples used for comparison with histological measurements. V<sub>vox</sub>: voxel size, x<sub>c</sub>: distance between sample and camera,  $\delta$ : $\beta$ : ratio of the dispersive and absorptive aspects of the wave-matter interaction, E: beam energy, n<sub>p</sub>: number of X-ray 2D projections, N: number of scans realised on the sample. <sup>†/‡</sup>Sample kept immersed in alcohol/glue during the scan.

## Optimisation of X-ray Imaging Conditions

**Standard absorption imaging mode.** A first series of scans was performed on 3 larynges (L<sub>1</sub> to L<sub>3</sub>) using a laboratory X-ray microtomograph in standard absorption imaging mode. The spatial resolution was first set to a voxel size of 45<sup>3</sup> and 25<sup>3</sup>  $\mu\text{m}^3$ , which is at least twice to ten times larger than the one used in standard hospital CT-scans. *A priori*, this resolution should allow the identification of the various layers of the vocal folds as well as a fibrous texture corresponding to muscle fibres or collagen/elastin fibre bundles. The other X-ray scanning parameters are listed in Table 1, but are similar to those already used on medical scanners for the radiology of *in vivo* and *ex vivo* larynges<sup>30,31</sup>. Typical orthogonal slices obtained with these imaging conditions are reported in Fig. 1(b,c). The external surface of the larynges is clearly identified. In addition, calcified portions of the thyroid, cricoid and arytenoid cartilages, coming from a partial ossification that starts typically at the age of 20<sup>32,33</sup>, are visible as lighter zones corresponding to higher attenuation – calcium exhibiting a much higher X-ray absorption coefficient than the other soft tissues of the larynges. Unfortunately, given that the chemical components of soft tissues exhibit similar absorption coefficients, within the vocal fold, there is insufficient contrast to separate fat tissues, non-ossified cartilages, muscles, *epithelium* and the inner structure of the *lamina propria*. This was not significantly improved while imaging the larynges with a synchrotron tomograph in absorption mode (Supplementary Fig. S1(a)), despite the use of a monochromatic beam that should improve contrast between chemically close materials. The 3D hierarchical vocal-fold structure could not be detected even with a voxel size

of  $13^3 \mu\text{m}^3$ , nor even when imaging cryo-preserved samples or excised vocal folds without any surrounding cartilage (Supplementary Fig. S2(a)).

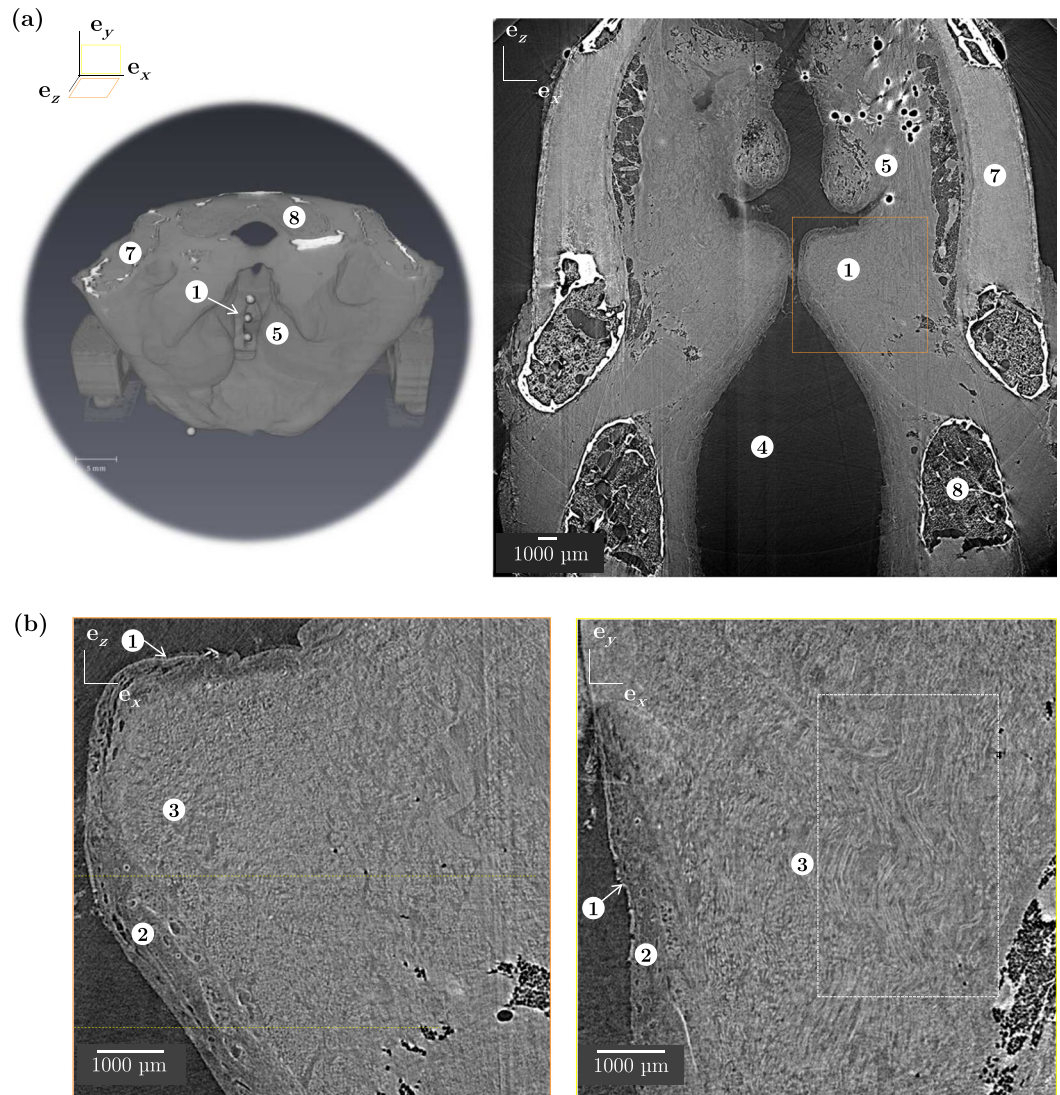
**Phase retrieval imaging mode without contrast agent.** To increase the contrast between the different constituents of the vocal folds, the samples were imaged on a synchrotron X-ray source with the objective of using phase retrieval (Paganin method), at “medium” (larynx samples) or “high” (excised vocal folds) spatial resolutions. Corresponding scanning and reconstruction parameters are reported in Tables 1 and 2. The effect of the sample-to-detector propagation distance  $x_c$  on the phase contrast enhancement, with mid-field and far-field optical configurations was investigated. Some representative 3D images of the larynx and excised vocal folds obtained with this technique are shown in Supplementary Fig. S1(b–d).

On the larynx at “medium” spatial resolution (hereforth defined at voxel size of  $13^3 \mu\text{m}^3$ ), phase retrieval allows adipose elements to be distinguished from other soft tissues (Supplementary Fig. S1(b,c)), which is in line with results obtained with phase retrieval imaging of human breasts<sup>34</sup>. However, neither the mid-field ( $x_c = 1.16$  m) nor the far-field ( $x_c = 11$  m) configurations were sufficient to allow clear visualisation of any multi-layered arrangement or fibrous network within the vocal folds. In the mid-field configuration, some hints of texture are discernible within the vocal tissue (see arrows in zoomed view). The far-field configuration further improves contrast (revealing several vocal-fold structural features), at the cost of detrimental reconstruction artefacts at the air-tissue interface, that spread into the sample and severely alter the image (Supplementary Fig. S1(c)).

On the excised vocal fold at “high” spatial resolution (hereforth defined at voxel size of  $0.65^3 \mu\text{m}^3$ ), a slight network-like texture is revealed within the muscular layer with the mid-field configuration ( $x_c = 40$  mm), as displayed in Supplementary Fig. S1(d). However, the *lamina propria* still appears as uniform, *i.e.*, without any textures related to collagen and elastin fibres. Doubling the propagation distance  $x_c$  to 80 mm enhances the phase contrast for the muscular layer, but also interferences at air-tissue interface. In the far-field configuration ( $x_c = 400$  mm), contrast within the vocal fold is not drastically improved and details of the texture are lost due to blur.

**Phase retrieval imaging mode with contrast agents.** To further increase the contrast between the constituents of soft tissues, a series of excised vocal folds were first immersed into aqueous solutions of formaldehyde or ethanol and then imaged according to similar scanning parameters as above (see Table 2). As evident from the corresponding images shown in Supplementary Fig. S2, formaldehyde reduces contrast and was therefore discarded: the muscular fibrous networks of the *vocalis* now appear as a fully homogeneous material. On the contrary, the use of ethanol solutions improve contrast, in agreement with the 3D images obtained in previous studies<sup>28,35,36</sup>. After a 3-day immersion in an aqueous solution with 30% of ethanol (Supplementary Fig. S2(b)), the different vocal-fold sublayers can now be identified, *i.e.*, the *vocalis* with well-defined cross sections of muscle fibres, the *lamina propria* with a similar but finer texture, and the *epithelium* with some saturated brightness due to sharp phase contrast at air-tissue interface. These qualitative observations are confirmed and even reinforced for an increased concentration of ethanol, as illustrated in Supplementary Fig. S2(c) in case of a 4-day immersion in an aqueous solution with 70% of ethanol. As displayed in Supplementary Fig. S2(e), immersion and scanning in pure ethanol was also attempted with the objective of measuring diffusion times. A diffusion time  $\delta t$  of 13 min was found as a minimum to identify the vocal-fold sublayers and the fibrous texture within the *lamina propria*. Given the success of pure ethanol immersion combined with phase retrieval, images of the entire larynx confined in a box filled with pure ethanol were also acquired at the “medium” spatial resolution (voxel size of  $13^3 \mu\text{m}^3$ ) and far-field configuration: Fig. 2 shows representative slices obtained in coronal and transverse planes. The wavy arrangement of muscular-fibre bundles and their orientation along the anteroposterior direction are clearly visible. In addition, although the fibrous texture within the *lamina propria* is still hardly detectable at this spatial resolution, the delineation of this layer with the *vocalis* can be well-differentiated (see Fig. 5(a)).

**Comparison with standard histological optical micrographs.** Standard histological stained samples were extracted from a vocal fold. The resulting optical micrographs were compared with slices obtained on the same larynx using high resolution synchrotron X-ray tomography with phase retrieval and ethanol immersion. This comparison is shown in Figs 3 and 4 and qualitatively validates the imaging procedure. In particular, the skeletal “muscle fibres” (also called “muscle striated cells” or “rhabdomyocytes”) are usually recognised by their distinctive transverse banding patterns due to the arrangement of myofibrils. This transverse banding is detectable with both imaging techniques (Fig. 3). As shown on both images, the muscle fibres are grouped together into bundles (or *fasciculi*) and surrounded by loose connective tissue. Furthermore, multiple epithelial cell layers, closely packed, stratified, and nucleated<sup>2,37</sup>, are also apparent on both X-ray images and histological micrographs, thereby allowing the *epithelium* delineation (Fig. 4(a–c,e)). Moreover, the tortuous arrangement of smaller fibres which are noticeable in tomographic images between the muscle fibres of the *vocalis* and the *epithelium*, is similar to the organisation of ECM fibres in the *lamina propria* layer, as highlighted in histological views in Fig. 4 and Supplementary Fig. S3. However, such histological micrographs show that, in the investigated regions, collagen Type I, collagen Type III, and elastin fibres in the ECM are closely entangled, which makes their distinction into separate networks difficult, as is the case in nonlinear microscopy<sup>18</sup>. Synchrotron X-ray microtomography does not solve the issue either: we were not able to distinguish elastin fibres from collagen ones, since their X-ray absorption and phase coefficients are very close. Finally, the microstructural differences between arytenoid hyaline and elastic cartilages can also be identified with both imaging techniques (Supplementary Fig. S4). In particular, the high content of elastic fibres surrounding the chondrocytes (*nuclei*) in *lacunae* is detectable in X-ray scans of elastic cartilage.



**Figure 2.** Characterisation of human laryngeal tissue's architecture using X-ray synchrotron microtomographic PRI imaging mode and medium spatial resolution ( $V_{\text{vox}} = 13^3 \mu\text{m}^3$ ). (a) *left*: 3D reconstruction of sample  $L_5$ ; *right*: tomographic image of a larynx vertical 2D slice, coronal plane ( $L_9$ ,  $[\text{C}_2\text{H}_6\text{O}] = 100\%$ ). ① Vocal fold, ④ Trachea, ⑤ Ventricular fold, ⑦ Thyroid cartilage, ⑧ Cricoid cartilage. (b) Zoom on one vocal fold and its inner fibrous architecture in coronal (orange) and perpendicular transverse (yellow) planes ( $L_9$ ,  $[\text{C}_2\text{H}_6\text{O}] = 100\%$ ). ① EP, ② LP, ③ M.

### 3D Hierarchical Structure of the Vocal Folds

**Larynx and vocal fold geometries.** Using large voxel sizes of  $45^3$  and  $25^3 \mu\text{m}^3$ , the obtained 3D images allow the definition of the outer surface of the larynx and the vocal folds (Figs 1 and 2). Using phase retrieval imaging combined with ethanol immersion, important measurements of the inner and layered structure of vocal folds can be made (see next subsections). In addition, the database of samples scanned in this work comprises four laryngeal samples ( $L_5$  to  $L_8$ ) scanned at rest and while being subjected to a 3D deformation resulting from cricothyroid tilt (Supplementary Figs S7 and S8). Thanks to the glass beads stuck inside the samples onto the vocal folds (see Fig. 2(a)), the longitudinal elongation  $\lambda$  of vocal folds after the larynx deformations can be estimated: depending on the tested samples,  $\lambda$  ranges from 1.05 to 1.27 with a mean value 1.13.

**Structure of the *vocalis*.** In the vicinity of the *lamina propria*, the *vocalis* is constituted of well-delimited fibres, mainly oriented along the vocal-fold longitudinal direction (Figs 2 and 3). Further from the *lamina propria* (typically beyond 1 mm), other fibrous networks with perpendicular orientation are also highlighted in several muscle regions (see Fig. 3(c)). They appear at the transition between the *vocalis* and the external part of the thyroarytenoid muscle, or between the *vocalis* and the lateral cricoarytenoid muscle. In addition, some interesting results were obtained on the ethanol-preserved samples at both “medium” (voxel size of  $13^3 \mu\text{m}^3$ , Fig. 5) and “high” (voxel size of  $0.65^3 \mu\text{m}^3$ , Fig. 6) spatial resolutions:

**Medium spatial resolution.** Figure 5(b) shows a 3D fibre orientation map of the muscular fibrous network in the reference frame ( $\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z$ ), where  $\mathbf{e}_x$  coincides with the mediolateral direction,  $\mathbf{e}_y$ , with the anteroposterior direction and  $\mathbf{e}_z$ , with the inferosuperior direction. The map was obtained from a subvolume of about  $3.25 \times 5.01 \times 2.55 \text{ mm}^3$  that was extracted near the *lamina propria* of larynx  $L_9$ . The selected 3D region was characterised by a single network of muscle fibres mainly oriented along the longitudinal direction  $\mathbf{e}_y$  of the fold – Fig. 2(b) shows a representative slice plotted in white within the cropped subvolume. Indeed, the mean values of the angles  $\theta_i$  and  $\varphi_i$  are both close to  $82^\circ$ , the non-diagonal components of the fibre orientation tensor  $\mathbf{A}$  are very small, and its component  $a_{yy}$  is much larger than  $a_{xx}$  and  $a_{zz}$ . However, muscle fibres are not perfectly aligned: the standard deviations for  $\theta_i$  and  $\varphi_i$  reach non zero values, of the order of  $25^\circ$  and  $28^\circ$ , and so the components  $a_{xx}$  and  $a_{zz}$ . It is worth noting that these two components are practically equal, which means that the network of muscle fibres exhibits transverse isotropy at first order. In other words, the muscle fibres can be seen as a network with rotational symmetry with respect to the anteroposterior direction,  $\mathbf{e}_y$ . More precisely, three peaks emerge from the 3D orientation map (noted with white crosses on Fig. 5(b)):  $(\theta_i, \varphi_i) = (77^\circ, 55^\circ); (59^\circ, 107^\circ); (108^\circ, 114^\circ)$ . Such peaks are also highlighted once reported in cumulative histograms of  $\theta_i$  and  $\varphi_i$  angles (noted with blue arrows). They are probably related to the 3D wavy arrangement of muscle fibres, as illustrated on representative orthogonal slices in Fig. 5(b). Please note that the variations of local  $\varphi_i$ -orientations in the  $(\mathbf{e}_u, \mathbf{e}_v)$  plane as defined in Fig. 5(b) are delimited by peaks and valleys of a quasi-sinusoidal waveform, which is a characteristic of the fibrous network waviness: the corresponding spatial period and magnitude were estimated to be around  $\lambda_{uv} = 1.50 \text{ mm} \pm 0.15 \text{ mm}$  and  $R_{0uv} = 220 \mu\text{m} \pm 60 \mu\text{m}$ , respectively. In the orthogonal plane  $(\mathbf{e}_u, \mathbf{e}_v)$ , a quasi-sinusoidal waveform is also observed, with a spatial period  $\lambda_{uv} = 1.51 \text{ mm} \pm 0.2 \text{ mm}$  and an amplitude  $R_{0uv} = 198 \mu\text{m} \pm 60 \mu\text{m}$ .

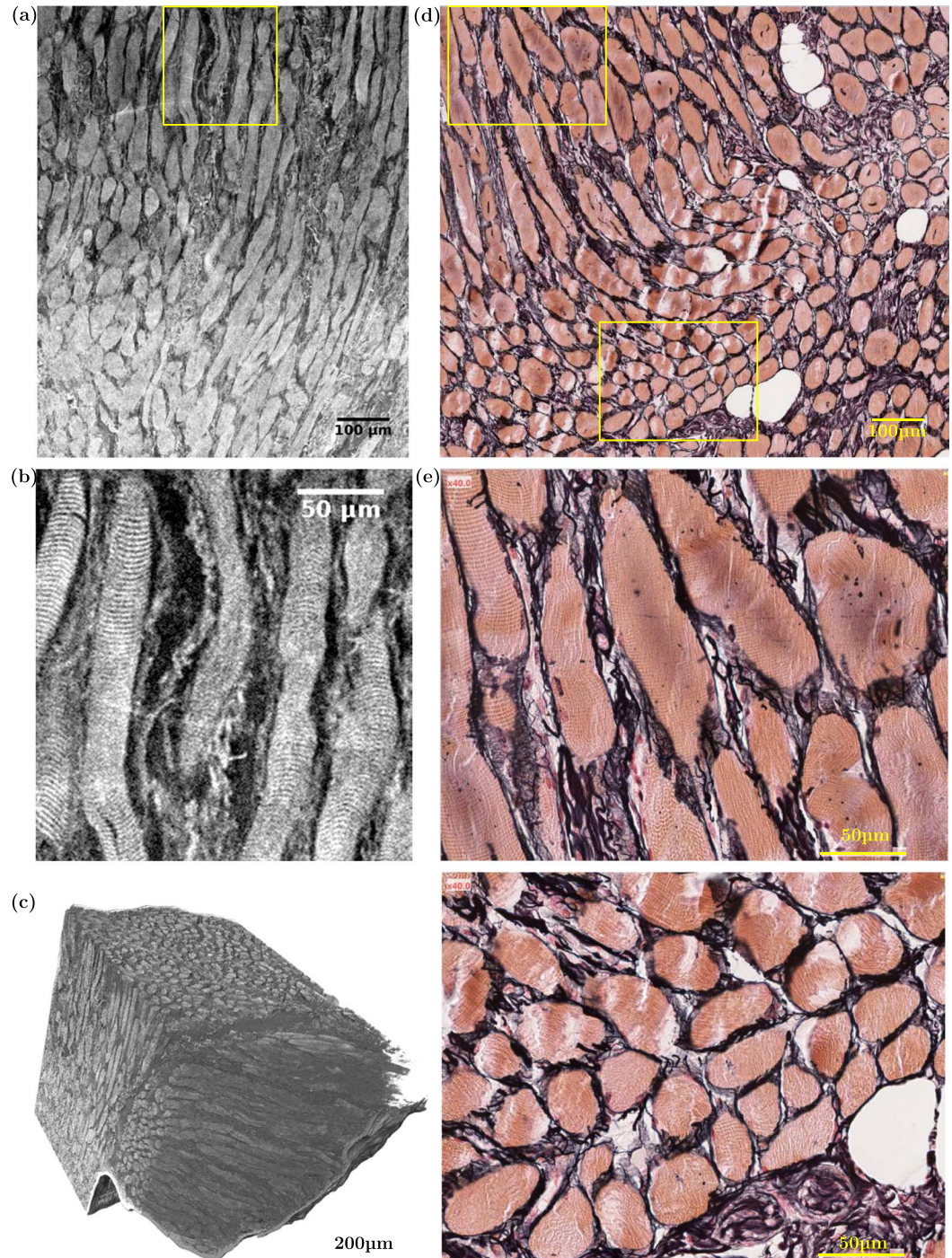
**High spatial resolution.** The same (but local) analysis was carried out with a subvolume of  $209 \times 403 \times 81 \mu\text{m}^3$  (Fig. 6(a)) extracted from sample  $L_{10}$ - $S_3$  at 1.3 mm from the *lamina propria*. The corresponding fibre orientation tensor is practically identical to that found at “medium” spatial resolution (note that the  $a_{xy}$  component reaches a non negligible value due to a slight misalignment of the microstructure with respect to the reference frame). The local fibre orientation distribution is also dispersed (see the 3D orientation map in Fig. 6(a)), with standard deviations for  $\theta_i$  and  $\varphi_i$  equal to  $24^\circ$  and  $28^\circ$ , respectively. The whole angular distribution looks rather homogeneous and centred around a single point  $(\theta_i, \varphi_i) = (93^\circ, 76^\circ)$ . Yet, two peaks emerge in the 3D orientation distribution, for  $(\theta_i, \varphi_i) = (93^\circ, 61^\circ); (100^\circ, 80^\circ)$ . At this spatial resolution, the analysis is sensitive to the variation of local orientation along fibres, so that both peaks can be related to their waviness, as illustrated by the segmented fibres shown in Fig. 6(a). In addition, other descriptors of muscle fibres were extracted within a subvolume of size  $615 \times 450 \times 98 \mu\text{m}^3$ , close to the previous one. For instance, the average surface number of muscle fibres was found to be around  $380 \text{ mm}^{-2}$ , by analysing slices perpendicular to the main fibre orientation. The equivalent diameter of muscle fibres  $d_e$ , the size of their minor and major inertia axes were computed: the corresponding distributions are plotted in Fig. 6(b). The peak value of  $d_e$  is  $29.3 \mu\text{m}$  (median value  $23.8 \mu\text{m}$ ). The shape of fibres was quantified by a peak roundness metric  $\xi$  of 0.70 (median value 0.63), which shows that muscle fibres exhibit quasi-circular cross-sections. These results are in agreement with those obtained from a single muscle fibre of length  $700 \mu\text{m}$  (Fig. 6(b)), which was isolated from the fibrous network. The mean diameter  $d_e$  of this fibre is  $54 \mu\text{m}$  (values ranging from  $44.3$  to  $64.3 \mu\text{m}$ ) and its mean roundness metric is close to 0.71 (values ranging from 0.58 to 0.85).

**Structure of the *lamina propria* and the *epithelium*.** For the *lamina propria* and the *epithelium*, the following results were obtained on ethanol-preserved samples:

**Medium spatial resolution.** The thickness of the *lamina propria* together with the *epithelium* was measured with both 2D and 3D views. In Fig. 4, regions of tissues characterised by X-ray microtomography and histological staining were positioned in the vocal-fold membranous part, at about 2–3 mm from the vocal process (arytenoid’s anterior end). Within this vocal fold (female, 92 years), the thickness in the coronal plane was measured between  $530$  and  $930 \mu\text{m}$  from the 3D images (mean value  $750 \mu\text{m} \pm 95 \mu\text{m}$  over 50 measurements in a 3D neighbourhood region), and between  $430$  and  $650 \mu\text{m}$  from the 2D histological micrographs. The 3D local thickness maps of both vocal folds were also derived on sample  $L_9$  (Fig. 5(a)), revealing an asymmetry between each fold, where thickness values globally varied from  $780$  to  $2054 \mu\text{m}$  depending on the anatomical location.

**High spatial resolution.** The set of 2D and 3D micrographs related to sample  $L_{10}$ - $S_3$  in the mid-membranous region<sup>38</sup> and displayed in Fig. 4 allows an interesting inspection of the inner and multilayered structure of vocal folds in the strongest collision zone. The slice (a), which is perpendicular to the anteroposterior direction, and its two companion and orthogonal slices (b) and (c) emphasize the *epithelium* ①, the *lamina propria* ② and the *vocalis* ③. From them, the spatial variation of the *epithelium* thickness along the vocal-fold outer surface (free edge) was estimated. For this sample, the 3D measurements yielded a thickness ranging from  $30$  to  $80 \mu\text{m}$ , which is in line with standard 2D histological assessments (see Fig. 4(e) and Supplementary Fig. S3), and previous values reported in the literature<sup>2,39,40</sup>. More interestingly, X-ray microtomographic slices (a-c) also reveal sublayers of complex shapes and characteristic textures within the *lamina propria*. When these sublayers can be distinctly separated, they are delineated in Fig. 4(a-c), and interpreted according to those detectable on the 2D photomicrograph (e):

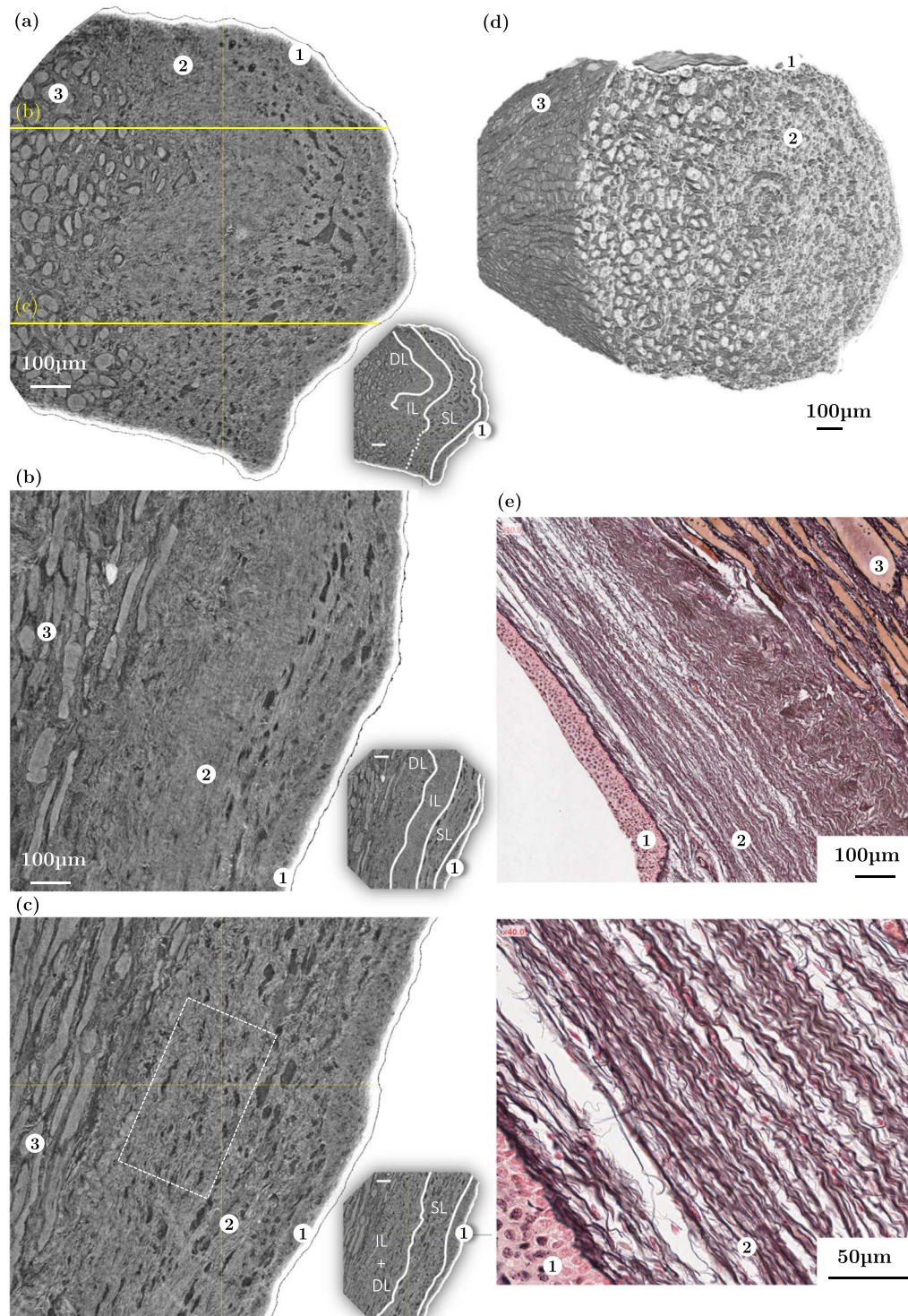
- Underneath the *epithelium*, a first sublayer is recognised (a-c, e), where the ECM fibre content is lower than in the rest of the *lamina propria*: it is characterised by large darker zones of  $\approx 20 \mu\text{m}$  in size, aligned with the fibrous networks. This region corresponds to the superficial layer of the *lamina propria* (SL), or Reinke’s space. In the observed volumes, large variations of the SL thickness were observed, ranging between  $100$  and  $200 \mu\text{m}$ .



**Figure 3.** Comparison between (left) synchrotron high-resolution X-ray microtomographic images (PRI mode,  $V_{vox} = 0.65^3 \mu\text{m}^3$ ) of the left *vocalis* fibrous network and (right) 2D histological photomicrographs of the right *vocalis* excised from the same larynx  $L_{10}$ . (a) 2D coronal view of  $L_{10}$ - $S_3$ ,  $[C_2H_6O] = 70\%$ . (b) Zoom on the banding patterns of the muscle fibres located within the yellow frame in (a). (c) 3D reconstruction of  $L_4$ - $S_2$ ,  $[C_2H_6O] = 30\%$ . (d) 2D coronal view of  $L_{10}$ - $S_3$ , prepared with Reticulin stain: Type III collagen fibres (black); muscular striated-cells or “fibres” (orange); adipocytes (white). (e) Zooms on the banding patterns of the muscle fibres located within the yellow frames in (d).

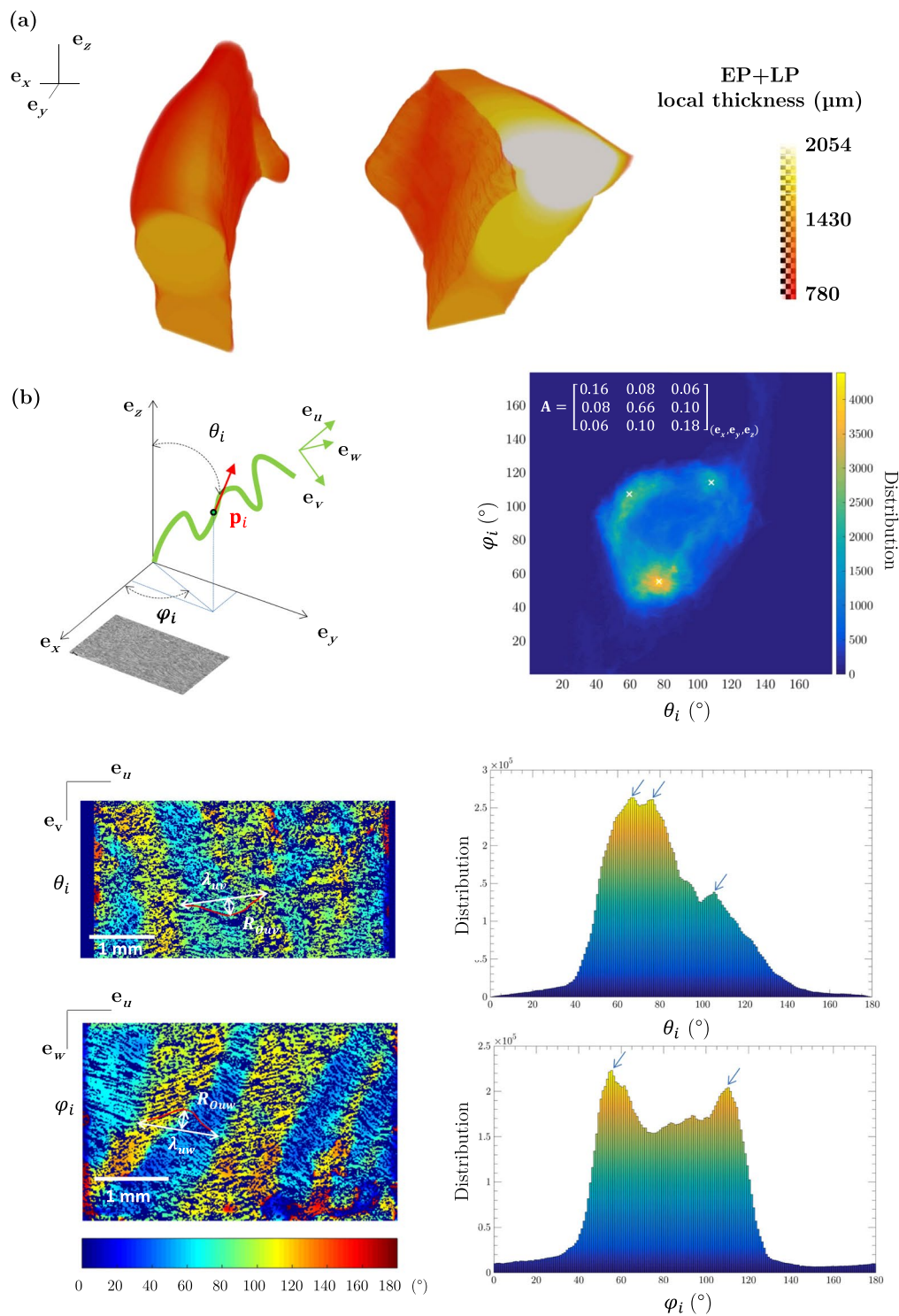
- Underlying the superficial sublayer, the intermediate sublayer (IL) is also detected (a-b, e), where the ECM fibre content is higher, and the fibre waviness is far less pronounced than in the third deep sublayer (DL), located between the IL and the *vocalis*.
- The three aforementioned sublayers are clearly visible in micrographs (b, e). However, it is worth noting that this is not the case everywhere in the volume: in slices (a) and (c) for example, both IL and DL layers can hardly be separated by the naked eye.





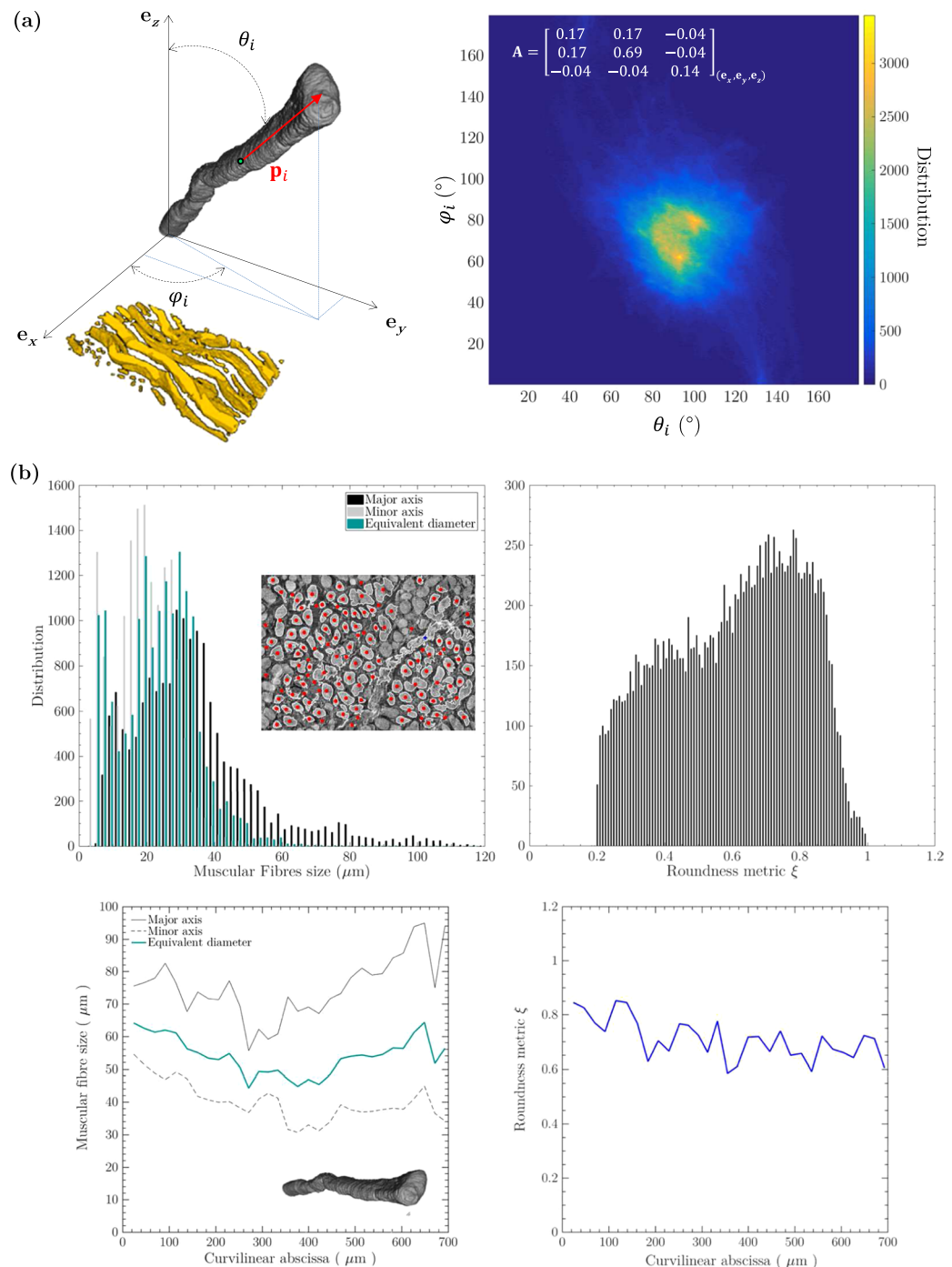
**Figure 4.** Comparison between (a–d) synchrotron high-resolution X-ray microtomographic images (PRI mode,  $V_{\text{vox}} = 0.65^3 \mu\text{m}^3$ ) of the left *lamina propria* fibrous network and (e) 2D histological micrographs of the right *lamina propria* excised from the same larynx  $L_{10}$ . (a) 2D coronal view of  $L_{10}$ - $S_3$ ,  $[C_2H_6O] = 70\%$ . (b,c) Companion and orthogonal 2D transversal views. (d) Corresponding 3D reconstruction; (e) (top) Histological photomicrograph of  $L_{10}$ - $S_3$  prepared with Reticulin stain: Type III collagen fibres (black); muscle fibres (orange); stratified squamous *epithelium* (pink); (bottom) Zoom on the wavy arrangement of elastin and collagen bundles of fibres. © EP, © LP, © M.

We further analysed the fibrous structure within the so-called “vocal ligament”, which contains both intermediate and deep sublayers of the *lamina propria*<sup>2,41</sup>, and is considered as the primary load-bearing portion of the vocal fold, especially at high longitudinal stretches<sup>19</sup>. The 3D orientation map of the network of collagen-elastin



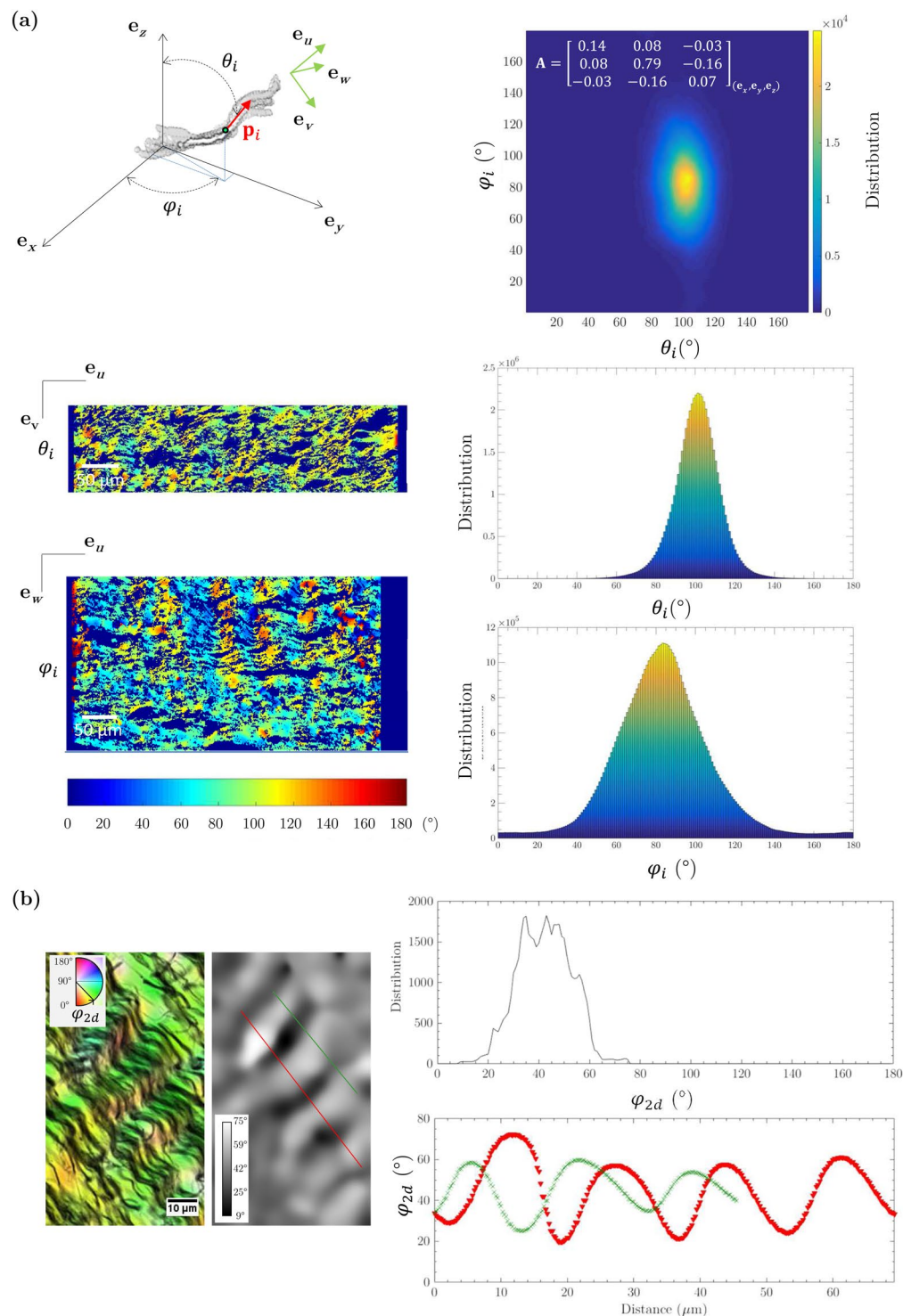
**Figure 5.** Quantification of several structural descriptors of the vocal-fold EP-, LP- and M-sublayers derived from images acquired at medium spatial resolution (voxel size of  $13^3 \mu\text{m}^3$ ) ( $L_0$ ). (a) 3D local thickness map of an EP+LP-layer subvolume. (b) 3D orientation map of the muscle fibres network in an M-layer  $41.52 \text{ mm}^3$  subvolume; corresponding 2D distribution of  $\theta_i$  and  $\varphi_i$  values; illustration of their distribution on two orthogonal 2D slices of the subvolume.

fibres together with its corresponding fibre orientation tensor  $A$  are given in Fig. 7(a). These data were obtained from a subvolume of  $267 \times 515 \times 131 \mu\text{m}^3$  extracted from sample  $L_{10}\text{-S}_3$  – Fig. 4(c) shows a representative slice plotted in white within the cropped subvolume. Despite the slight misalignment of the microstructure with respect to the reference frame ( $e_x, e_y, e_z$ ), it is fair to conclude that fibre bundles are mainly oriented along the



**Figure 6.** Quantification of several structural descriptors of the vocal-fold muscular layer ( $L_{10}$ - $S_3$ ) derived from images acquired at high spatial resolution (voxel size of  $0.65^3 \mu\text{m}^3$ ). **(a)** 3D orientation map of the muscle fibres network in a M-layer  $0.007 \text{ mm}^3$  subvolume. **(b)** Statistical distributions and shapes of the muscular fibrous network (*top panel*) and that of an individual muscle fibre extracted from the 3D image (*middle panel*).

$e_y$  direction, as for muscle fibres near the *lamina propria*: the fibre bundle orientation distribution is centred around  $(\theta_i, \varphi_i) = (102^\circ, 85^\circ)$ , so that the component  $a_{yy}$  of  $A$  is much higher than  $a_{xx}$  and  $a_{zz}$ . It is also interesting to mention that  $a_{xx}$  is twice as large as  $a_{zz}$  (the dispersion of the angular distribution is characterised by standard deviations for  $\theta_i$  and  $\varphi_i$  equal to  $12^\circ$  and  $25^\circ$ , respectively), proving that the fibrous network of the *lamina propria* does not exhibit transverse isotropy but orthotropy: whereas muscle fibres exhibit transverse isotropy with respect to the anteroposterior direction, the fibre alignment along this direction and inside the *lamina propria* is more pronounced through the thickness of the *lamina propria* (within the plane  $(e_x, e_y)$ ) than perpendicularly (within the plane  $(e_x, e_z)$ ). In addition, as for muscle fibres, a quasi-periodic variation of the local orientation of fibres along the anteroposterior axis  $e_y$  is clearly observed on 2D slices. The spatial periods and amplitudes of the



**Figure 7.** Quantification of several structural descriptors of the vocal-fold LP-layer ( $L_{10}$ - $S_3$ ). (a) 3D orientation map of the collagen and elastin fibrous networks in a LP-layer  $0.018 \text{ mm}^3$  subvolume derived from microtomographic measurements acquired at high spatial resolution (voxel size of  $0.653 \mu\text{m}^3$ ); corresponding 2D distribution of  $\theta_i$  and  $\varphi_i$  values; illustration of their distribution on two orthogonal 2D slices of the subvolume. (b) In-plane waviness of Type III collagen fibres network, derived from histological measurements.

fibre waviness were thus measured in orthogonal planes ( $\mathbf{e}_u, \mathbf{e}_w$ ) and ( $\mathbf{e}_u, \mathbf{e}_v$ ) leading to  $\lambda_{uw} = 56.3 \mu\text{m} \pm 16 \mu\text{m}$ ,  $R_{0uw} = 4.6 \mu\text{m} \pm 1.5 \mu\text{m}$ ,  $\lambda_{uv} = 44.3 \mu\text{m} \pm 15 \mu\text{m}$  and  $R_{0uv} = 11.3 \mu\text{m} \pm 3 \mu\text{m}$ . Despite being slightly higher, the 3D spatial periods are consistent with the value deduced from the 2D histological micrograph shown in Fig. 7(b), *i.e.*,  $\lambda_{2D} = 16 \mu\text{m} \pm 1.6 \mu\text{m}$ . Finally, the 3D contourlines of several fibrous bundles of collagen and/or elastin were

also detected, as illustrated in Fig. 7(a). Such bundles have a mean diameter and length around 50  $\mu\text{m}$  and 400  $\mu\text{m}$ , respectively. However, due to the limited contrast obtained in this zone (see Fig. 4), no threshold-based segmentation was successfully achieved to provide a statistical quantification of their dimensions and shape.

## Discussion and Concluding Remarks

Since the reference histological findings and 3D sketches on the vocal-fold multilayered arrangement provided by Hirano *et al.* in the 1970–80s<sup>6,41,42</sup>, very few experimental quantitative data have been collected to characterise the complex architecture of human vocal-folds, in particular at the micrometer scale. Little information is available in the literature for the structure of *vocalis* at this scale. Regarding the multi-layered structure of the *lamina propria* and the *epithelium*, the existing information is mainly derived from standard 2D histological stainings<sup>8,37,40,41</sup>, micro-Magnetic Resonance Imaging techniques<sup>17</sup> with a restrained spatial resolution, or, in a growing number of studies, from non-linear laser scanning microscopy where some relevant quantitative structural descriptors were assessed for the ECM fibres and their networks<sup>18–21</sup>. In the present work, for the first time, the 3D hierarchical architecture of human vocal folds is revealed *ex vivo* by means of fast synchrotron X-ray microtomography with phase retrieval imaging mode, together with the use of a suitable contrast agent. This 3D characterisation was performed at two spatial resolutions: (i) “medium” (voxel size of 13<sup>3</sup>  $\mu\text{m}^3$ ), with the visualisation of the vocal fold geometries, their general placement within the larynx and the identification of their layers, *e.g.*, *vocalis*, *lamina propria* (and its sublayers) and *epithelium*; (ii) “high” (voxel size of 0.65<sup>3</sup>  $\mu\text{m}^3$ ), with a fine examination of muscular and ECM fibres as well as fibrous networks. Comparing results of the aforementioned studies with those obtained here, some answers to the technical and fundamental questions (A–E) listed in the Introduction can now be discussed:

**Relevance of synchrotron X-ray microtomography.** To answer question (A), a methodology was elaborated to prepare vocal-fold samples with suitable conditions of tissue conservation, and to optimise the optical settings for scanning. The 3D imaging protocol was validated by comparing 3D images with standard 2D histological micrographs. A high degree of accordance was found between both techniques from a qualitative standpoint. Quantitatively, however, some differences were found, *e.g.*, for the thickness of the *lamina propria* or for the waviness of its fibres. Although small, these discrepancies may be mainly ascribed to the manipulations to which the tissue is subjected during the protocol for histological staining, *e.g.*, pre-deformation of samples, dehydration and possible induced shrinkages, warming, paraffin-embedding, microtomy artefacts... This point undoubtedly constitutes a strength of 3D X-ray imaging. Furthermore, compared with histological staining or other advanced 3D imaging techniques such as multiphoton microscopy<sup>18–22,43</sup>, the other interesting advantage of X-ray 3D imaging is the possibility (i) to get high resolution and in depth 3D images of soft tissues (ii) with very fast scanning times (here the scans were 1–2 min long). However, the technique exhibits three main limitations that should be overcome for future observations. First, the use of ethanol was necessary in order to enhance phase contrast: its impacts on the mechanical behaviour of vocal fold tissues must be assessed before performing mechanical tests with 3D *in situ* observations. Note that the relatively short duration of fixation of 13 min – the minimum for the fibrous texture of the *lamina propria* to be observed – should limit the impact on the mechanical behaviour of the tissue. Second, we experienced difficulties in discriminating collagen fibres from elastin fibres, since both their absorption and phase coefficients are very close. A first possible solution to overcome this problem consists in selectively digesting a target protein from the vocal-fold tissue, as previously done in the case of arterial walls<sup>44</sup>. Another solution could be to use chemical markers able to track one specific fibrous protein<sup>29,45</sup>. Last, we were not able to properly analyse the morphology of individual ECM fibres: this challenging task could be achieved by further increasing the spatial resolution of the images<sup>46</sup>.

**Size and spatial variation of the vocal fold (sub)layers.** Regarding question (B), the thickness of the *lamina propria* and the *epithelium* is known to vary with gender and age<sup>37,41,47,48</sup>. The local 3D thickness map displayed in Fig. 5(a) demonstrates that it also varies with the anatomical location within the vocal-fold tissue, being the thickest at the midportion and becoming thinner towards the anterior and posterior ends, in agreement with previous 2D studies<sup>41</sup>. Besides, the qualitative division of the *lamina propria* into a superficial sublayer (SL) and a “vocal ligament” (IL + DL), as illustrated in Fig. 4(a–c), is in line with the typical distribution reported by Gray *et al.*<sup>37</sup>, *i.e.*, with a SL-layer ranging between 25% to 35% of the initial depth of the *lamina propria*. However, even in the mid-membranous region of the vocal folds, the 3D images show that this trilamellar structure is not always observable. More precisely, we found that the intermediate sublayer (IL), which is marked by a dense network of straighter ECM fibres, may be confined to a region likely to endure the highest collision stresses during phonation<sup>38,49</sup>. This heterogeneous distribution in the vocal ligament is consistent with the structural measurements previously achieved by Klepceck *et al.*<sup>17</sup> using micro-MRI and plastination methods. It also supports the hypothesis that regions exposed to higher stresses in the vocal tissues are characterised by a higher density of fibrous proteins produced to strengthen the ECM<sup>37,50</sup>.

**3D structural anisotropy of fibrous networks in the *lamina propria* and the *vocalis*.** Within the vocal ligament, Kelleher *et al.*<sup>19</sup> estimated the 2D orientation distribution function of collagen fibres which was centred along the anteroposterior axis. We computed the corresponding 2D second-order fibre orientation tensor:

$$\mathbf{A} = \begin{bmatrix} 0.15 & 0 \\ 0 & 0.85 \end{bmatrix}_{(e_x, e_y)} \quad (1)$$

the components of which are very close to those we obtained in  $(\mathbf{e}_x, \mathbf{e}_y)$  for the 3D fibre orientation tensor reported in Fig. 7(a). Thanks to the 3D images, it is now possible to extract 3D properties of the fibre orientation in the vocal, and thereby to answer question (C): at the first order, it neither exhibits a 2D layered symmetry, nor transverse isotropy, but orthotropy with a pronounced major orientation along the anteroposterior direction, and minor and intermediate orientations along the inferosuperior and mediolateral directions, respectively. It is worth noting that this result does not fit with the 3D schematic fibrous arrangement proposed by Madrugá de Melo *et al.*<sup>8</sup> for the *lamina propria*: the fibrous architecture these authors suggested exhibits orthotropy without any privileged direction, while it is clearly highlighted here along the anteroposterior direction. This information is also important to build relevant micromechanical models for the *lamina propria*<sup>19,47,50,51</sup>, that are able to accurately reproduce the biomechanical properties of vocal tissues in multiple directions close to those encountered *in vivo*. For example, Kelleher *et al.*<sup>19,50</sup> showed that both the degree and the magnitude of the vocal ligament anisotropy strongly impact the predictions of its fundamental frequency of vibration. Also, new information was obtained regarding the structural anisotropy of muscle fibres in the *vocalis*. The resulting 3D fibre orientation distributions demonstrate that muscle fibres are mainly aligned along the anteroposterior axis, with a transverse isotropy texture. Here again, this information is important to fully characterise and model the mechanics of the *vocalis* under realistic physiological loadings<sup>52,53</sup>. In particular, Böl *et al.*<sup>53</sup> showed that such an anisotropic property implies different stress responses of the tissue under the applied compression modes, which is critical for the better understanding of vocal-fold collision.

**Morphology of ECM and muscle fibres.** Regarding question (D), Miri *et al.*<sup>18</sup> quantified the wavy shape of collagen fibres from three *lamina propria* (65 year-old female larynx, 55- and 68-year-old male larynges). The average period  $\lambda_0$  was found to be around 35  $\mu\text{m}$  (values ranging from 18 to 48  $\mu\text{m}$ ), while the average amplitude  $R_0$  was found to be equal to 6  $\mu\text{m}$  (values ranging from 2.6 to 8.4  $\mu\text{m}$ ). The waviness parameters we estimated from X-ray microtomographic images are consistent with the aforementioned values, although lying in the highest range. Furthermore, regarding muscle fibres in the *vocalis*, prior works in the 1980s reported a typical equivalent diameter of 32  $\mu\text{m}$  for (white) fast-twitch fibres and 28  $\mu\text{m}$  for (red) slow-twitch fibres, based on 2D microscopic differentiation<sup>54</sup>. This is in line with the equivalent circular diameter found in the present study at the scale of the 3D network. These structural data are very helpful to understand and model the mechanics of vocal folds.

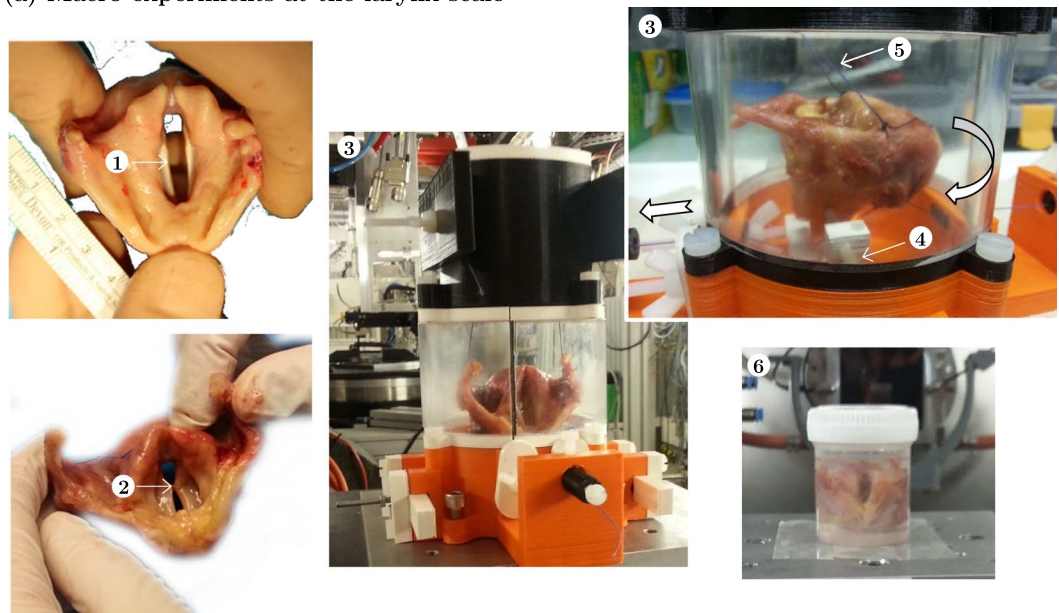
**Towards mechanical tests with 3D *in situ* observations.** At “medium” spatial resolution, a preliminary collection of 3D images was acquired for larynges subjected to anatomical placements close to those occurring during a typical phonatory cycle. This constitutes an important database which can be used either to study the larynx and vocal fold anatomy<sup>31,55</sup> or to simulate the mechanics of vocal folds<sup>56,57</sup>. For example, the vocal fold elongation  $\lambda$  which was measured after the laryngeal tilt is consistent with the recent measurements obtained by Lagier *et al.*<sup>55</sup>. Using 10 excised human larynges prepared and deformed in close conditions (body ages from 71 to 92, freezing in 0.9% saline, double adduction), these authors reported  $\lambda$ -values ranging from 1.05 to 1.28 (mean value 1.12). Finally, with a scan duration of 1–2 min, it is worth mentioning that the adopted methodology is very promising to further track, at “high” spatial resolution, the deformation and the rearrangement of muscle and ECM fibres during mechanical load, thereby answering question (E). Such information is crucial to gain an in-depth understanding of the link between the micromechanics of vocal-fold tissues and their unique vibratory performances. Although left static in the present study, a part of the scanned samples were clamped inside a micropress mounted in the imaging set-up, proving the feasibility of future *in situ* tension-compression tests.

## Methods

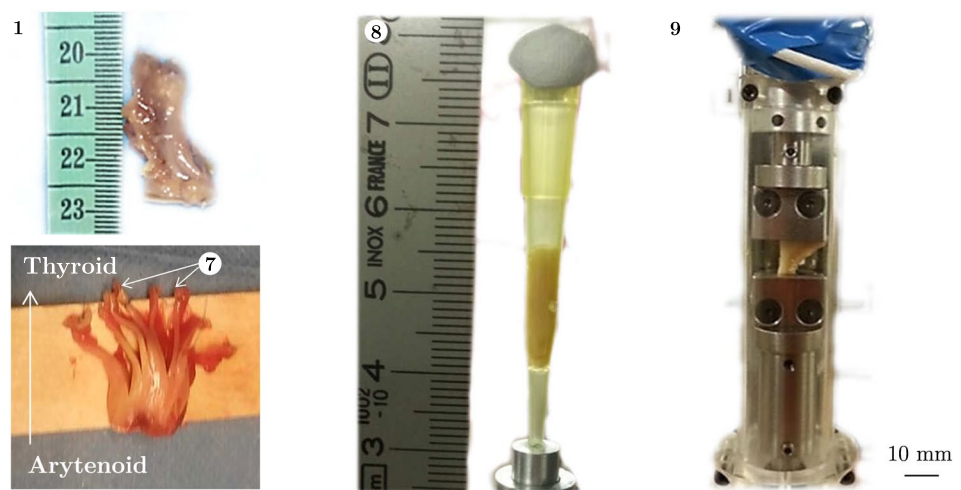
**Samples preparation.** The experiments were carried out with 10 *ex vivo* and healthy human larynges, referred to as  $L_i, i \in [1, \dots, 10]$ . Samples details are given in Table 1. Anatomical pieces were excised from donated bodies within 48 h *post-mortem* at the Laboratory of Anatomy of the French Alps (LADAF - UGA, Grenoble Hospital Univ.). All but one larynx (*fresh larynx*  $L_{10}$ ) were preserved by freezing ( $-20^\circ\text{C}$ ) up to X-ray scanning protocol and after. Experimental protocols were approved by the Bioethics Committee of the General Directorate for Research and Innovation (DGRI - French Research Ministry) and the LADAF. All experiments were conducted as regulated by the French ethical and safety laws in the frame of Body Donation (voluntary donation, made by the donor during his lifetime, after written and informed consent), in accordance with the ESRF Safety Office for Biology and Biochemistry and 3SR Lab policy. Samples were prepared following several protocols developed for testing and improving the microstructure exploration of the vocal-fold tissue, first with entire larynges, then on dissected vocal-fold samples:

**Experiments on larynges.** Larynx samples  $L_1$  to  $L_9$  were used with experimental details given in Table 1. Before any manipulation, each sample was defrosted 20 min in water ( $T \approx 20^\circ\text{C}$ ), then placed in a confined and fully-hydrated set-up, to reproduce realistic laryngeal phonatory placement and to control *vocalis* muscle stretching<sup>55</sup> (Fig. 8a). The set-up is made of 3D-printed PLA pieces and two PMMA half-cylinders forming a sealed transparent chamber (inner diameter of 8 cm, wall thickness of 5 mm). It was designed to mimic the rocking movement of the thyroid cartilage forward and backward on the cricoid cartilage, which elongates the vocal folds *in vivo*. Mounting a larynx in it comprised several steps: (i) a stitch was made between the vocal processes of the arytenoid cartilages using a 5/0 surgical wire (Polysorb®), so as to adduct the arytenoid cartilages and held the vocal folds close together in a phonatory-like position<sup>55</sup> (Fig. 8(a.2)); (ii) a stiffer 3/0 surgical wire was also fixed around the cricoid cartilage anterior part, passing through the thyrocricoid membrane, and linked to a manual activator of the cartilages’ rocking motion; (iii) a wooden pike was introduced through the thyroid cartilage and fixed into the containment device to ensure a reference static position (Fig. 8(a.4)); (iv) Glass beads (1 mm in diameter) were stuck on larynx cartilages using tissue glue (PeriAcryl®90, Surgibond®) so as to add 3D positions of reference and

## (a) Macro-experiments at the larynx scale



## (b) Micro-experiments at the vocal-fold scale



**Figure 8.** Experimental set-ups developed for X-ray microtomography characterisation of the vocal-fold tissues (a) within the preserved laryngeal structure and (b) once dissected. ① Vocal fold, ② Stitch between vocal processes, ③ Macro-mechanical set-up, ④ Wooden pike, ⑤ Wires blocking thyroid's branches, ⑥ Set-up for sample  $L_9$ , ⑦ Dissection procedures for samples  $L_i-S_j$ , ⑧ Pilot set-up, ⑨ Micro-mechanical tension device.

help to identify the laryngeal structures during X-ray data collection. Such markers also allowed to measure the vocal fold elongation during the rocking movement of the larynx; (v) the larynx was then placed into the containment device, blocking the thyroid cartilage thanks to the wooden pike and to additional wires stitched through its both ended branches (Fig. 8(a.5)); (vi) the whole device was kept opened in an airtight chamber which was regulated at proper hygrometric conditions (100% RH, achieved with a humidifier Fisher & Paykel HC150). Samples  $L_1$  to  $L_8$  were subjected to steps (i-v). For sample  $L_9$ , the larynx was directly confined in a sealed box filled with a dilute aqueous solution of ethanol (Fig. 8(a.6)) the concentration of which,  $[C_2H_6O]$ , was set to 0, 30 and 100%.

*Experiments on dissected vocal folds.* To enable 3D imaging at micrometer-scale resolution, vocal-fold samples were dissected from six laryngeal specimens  $L_i$  ( $i = 2, 3, 4, 6, 7, 10$  – see Table 2). All but one larynx (fresh larynx  $L_{10}$ ) were already exposed to X-ray radiation prior to dissection. They were cut along the anteroposterior direction with portion of thyroid and arytenoid cartilages, respectively (Fig. 8(b.1)). The dissection procedure yielded to a number  $n_i$  of vocal-fold samples derived from each larynx  $L_i$ , labelled as  $L_i-S_j$  ( $j \in [1, \dots, n_i]$ , Fig. 8(b.7)). As indicated in Table 2, samples were constituted either with only the muscular layer (M), or with several layers (LP+EP, M+LP+EP). It is worth noting that for the specific case of fresh larynx  $L_{10}$ , sample  $L_{10}-S_3$  was cut into the region

of highest collision, *i.e.*, close to the narrowest glottis<sup>38</sup>. Several conditions of tissue conservation were used: (i) defrosting after preservation at  $-20\text{ }^{\circ}\text{C}$ ; (ii) cryopreservation at  $-80\text{ }^{\circ}\text{C}$  using dry ice; (iii) pre-immersion in a dilute aqueous solution of ethanol with concentrations  $[\text{C}_2\text{H}_6\text{O}]$  ranging from 30% to 100% and diffusion time  $\delta t$  ranging from 1 min to more than 5 days (for diffusion times larger than 24 h, the immersion was achieved in successive baths of increasing alcoholic concentrations<sup>36</sup>, *i.e.*, 30%, 50% and 70%); (iv) pre-immersion in a 10%-formaldehyde ( $\text{CH}_2\text{O}$ ) solution. Two different containment procedures were tested. Vocal folds dissected from  $L_4$  and  $L_7$  were released from their cartilaginous ends and confined into a conic container (maximal inner diameter 5 mm and wall thickness 1 mm) sealed at its boundaries and optionally filled with glue or alcohol (Fig. 8(b.8)). Vocal folds dissected from  $L_6$  and  $L_{10}$  were mounted in a dedicated tension-compression micro-press<sup>58</sup> (Fig. 8(b.9)), with a chamber regulated at proper hygrometric conditions (100% RH), to explore the feasibility of future *in situ* tensile tests. Typical dimensions of the scanned samples were within  $10\text{ mm} \times 5\text{ mm}$  (Fig. 8(b.9)).

**X-rays microtomography.** *Laboratory X-ray microtomography.* Pre-tests were performed with a standard laboratory X-ray source (RX Solutions, 3SR Lab, Grenoble, France) equipped with a conical polychromatic and divergent beam (Hamamatsu L12161-07 source), allowing absorption imaging mode. Corresponding imaging parameters are reported in Tables 1 and 2 (italic lines)<sup>30,31</sup>. The scans were obtained with a number  $n_p$  of X-ray 2D radiographs onto a  $1914 \times 1580$  pixel<sup>2</sup> Varian flat panel detector, leading to a voxel size  $V_{\text{vox}}$  (varying from  $12^3$  to  $45^3\ \mu\text{m}^3$ ). Samples were exposed to a  $360^\circ$  rotation with respect to the X-ray source (step angle  $360^\circ/n_p$ , varying from  $0.07^\circ$  to  $0.25^\circ$ ), with an exposure time of 125 ms (respectively 400 ms) per radiograph for experiments on the whole laryngeal structures (respectively on dissected vocal tissues). To restrain the noise, an average of 6 radiographs per 2D image was used.

*Synchrotron X-ray microtomography.* Most of samples were imaged on the ID19 beamline of the European Synchrotron Radiation Facilities (ESRF, Grenoble, France), allowing advanced imaging possibilities thanks to: (i) a high photon flux in a homogeneous, parallel, monochromatic and highly coherent beam; (ii) the recording of phase images obtained by adjusting the sample-to-detector propagation distance,  $x_c$ , using single phase retrieval imaging mode (Paganin)<sup>26</sup>. The interaction of X-ray with matter is generally described by the complex refractive index of the sample,  $n = 1 - \delta + i\beta$ , where  $\delta$  is related to the phase-shift effects of the X-ray waves induced by the sample, and  $\beta$  determines their attenuation<sup>59,60</sup>. The term  $\beta$  is retrieved by recording the absorption images, *i.e.* for  $x_c = 0$ , while  $\delta$  can be retrieved using a single propagation according to Paganin's work, provided that the ratio of the dispersive and absorptive aspects of the wave-matter interaction,  $\delta:\beta$ , is known. The chosen imaging parameters are reported in Tables 1 and 2 for experiments achieved at “medium” and “high” spatial resolutions, respectively (*i.e.*, at voxel size of  $13^3$  and  $0.65^3\ \mu\text{m}^3$ , respectively)<sup>36,61–63</sup>. Corresponding optical set-ups are detailed thereafter:

*Experiments on larynges.* For these experiments, two optical set-ups were tested:

- For laryngeal samples  $L_4$  to  $L_8$ , a first series of scans was achieved using the mechanical set-up shown in Fig. 8. Two extreme phonatory positions were acquired sequentially for each sample: at rest in a first time, and at maximal macroscopic stretch of the vocal folds achieved by cricothyroid approximation in a second time. The X-ray beam was adjusted using filters (2.8 mm Al, 0.14 mm Cu), an average energy  $E$  of 65 keV ( $I = 200$  mA) and a wiggler gap of 95 mm. The transmitted beam was converted into visible light by a LuAg scintillator (thickness 500  $\mu\text{m}$ ), and recorded using a CCD camera (FReLoN-2K,  $2048 \times 2048$  pixel<sup>2</sup> chip, 14-bit dynamic range, FTM mode)<sup>64,65</sup>. The combination of the optics, using a sample-to-detector distance  $x_c = 1.2$  m in average, with the pixel size of the CCD ( $14^2\ \mu\text{m}^2$ ) allowed to work at “medium” spatial resolution, *i.e.* with an effective voxel size of  $13^3\ \mu\text{m}^3$ . 2D projections were collected according to the half-acquisition mode to obtain a 3D field-of-view of maximal size  $3587 \times 3587 \times 600$  voxels for each rotational acquisition. 4900 projections were acquired over the  $360^\circ$  rotation of samples. The beam exposure time was 100 ms per projection. Three rotational acquisitions were necessary to fully cover the whole vocal-fold tissue along its longitudinal axis (anteroposterior direction). These acquisitions were taken sequentially with an overlap of 127 slices (1.60 mm) from the angle of the thyroid cartilage. In the end, the scan duration was of 10 min and the total acquisition time about 1 h to scan both phonatory positions per larynx.
- For sample  $L_9$ , the optical set-up was modified so as to raise the sample-to-detector distance up to  $x_c = 11$  m and thereby maximise the phase-contrast effect<sup>61,62</sup>, by placing the sample in the ID19's monochromator hutch (allowing a propagation distance up to 14 m). Filters were updated (2.8 mm Al, 0.7 mm Cu, 0.28 mm Au), the average energy was kept at around 60 keV, and the wiggler gap fixed at 61 mm. Placed in a single static geometry at rest, sample  $L_9$  was scanned in this optical “far-field” configuration yielding to a volume of  $3706 \times 3706 \times 600$  voxels (other collection parameters being as above).

*Experiments on dissected vocal folds.* For these experiments, the optical set-up was changed to work at “high” spatial resolution, *i.e.* at a voxel size of  $0.65^3\ \mu\text{m}^3$ . In addition, a GGG10 scintillator and a PCO Edge 4.2 camera (SN 62000031,  $2048 \times 2048$  pixel<sup>2</sup> chip, no binning, FFM mode) were used, allowing fast and highly contrasted imaging. The acquisition parameters were characterised by a 19 keV beam energy, an undulator gap of 21 mm, a sample-to-detector distance  $x_c = 40$  mm in average, and a beam exposure time of 20 ms per radiograph. Several acquisitions were taken sequentially to cover the full height and width of the vocal-fold sample, with a field of view displaced by 1 mm-step in longitudinal and transversal directions, yielding to an overlap of 300  $\mu\text{m}$ . In the end, the scan duration was within 1–2 min and the total acquisition time about 1 h 35 min per sample.



**Image processing.** After reconstruction of the scans using Paganin's method for phase retrieval coupled with filtered back-projection<sup>26</sup> and removal of tomographic artefacts such as rings, the resulting 3D and greyscale images were analysed using Fiji<sup>®</sup>, Avizo<sup>®</sup> and Matlab<sup>®</sup> routines. Some of them were automatically segmented using standard smoothing and thresholding algorithms implemented in these software, to get meaningful 3D views (Figs 2(a), 3(c) and 4(d)) and to analyse the vocal-fold structure in detail (Figs 5(b), 6 and 7). In some cases, this was not possible. Thus, to pursue the quantitative analysis, manual image thresholding was carried out using Fiji<sup>®</sup> and a graphical tab (Wacom Cintiq 22HD touch). For example, the geometries of two vocal folds and the sublayers of their *lamina propria* were tracked on images acquired at “medium” spatial resolution (voxel size of  $13^3 \mu\text{m}^3$ ), as shown in Fig. 5(a). Similarly, some individual muscle fibres and ECM fibre bundles were extracted from the “high” resolution images (voxel size of  $0.65^3 \mu\text{m}^3$ ), to analyse their geometry (Figs 6(a) and 7(a)). Then, various operations were realised, to extract from these 3D images some relevant qualitative information and quantitative descriptors:

**Vocal-fold elongation.** From the segmented images of the larynges acquired before and after the rocking movement of the thyroid cartilage into the dedicated set-up (Fig. 8(a)), the local elongation of the vocal folds  $\lambda$  was measured thanks to the glass beads stuck on them. For that, beads were easily isolated from the rest of the larynx using thresholding algorithms (Fiji<sup>®</sup>), as illustrated in Supplementary Figs S7 and S8. The positions of the bead centres of mass were then detected with the 3D Particle Analyser plug-in of Fiji<sup>®</sup>. Their relative distance  $l_0$  and  $l$  was calculated in the non-deformed and deformed configurations respectively, to estimate the vocal-fold elongation  $\lambda = l/l_0$ .

**Thickness of the lamina propria.** A crop of the 3D image of larynx  $L_9$  was achieved in order to focus on the *lamina propria* structure at the millimetre scale. After manual segmentation (see above), the spatial thickness of the extracted *lamina propria* was computed using the Local Thickness plug-in of Fiji<sup>®66,67</sup>.

**Multiscale orientation of fibres.** The global 3D structural anisotropy of the muscle fibres was estimated from the greyscale image of larynx  $L_9$  acquired at “medium” spatial resolution (voxel size of  $13^3 \mu\text{m}^3$ ), thanks to the fibrous textures detected in the *vocalis*. Notice that this operation was not possible in the *lamina propria*, where no well-defined fibrous texture was observed at this scale. Thanks to the 3D images recorded at “high” spatial resolution (voxel size of  $0.65^3 \mu\text{m}^3$ ), the local 3D structural anisotropy of muscle fibres in the *vocalis* was also estimated, as well as the global 3D structural anisotropy of ECM fibres in the *lamina propria*. Whatever the tissue and spatial resolution, the following processing route was used for each considered Region Of Interest (ROI) – see Supplementary Figs S5 and S6: First, to enhance the grey-level gradients of the fibrous phases, 3D images were subjected to a coarse thresholding process, and a series of 3D smoothing and morphological operations (median filter, hole filling, opening/closure sequences). Second, the 3D Euclidean distance map was calculated in the considered ROI. Therewith, a home-made Matlab<sup>®</sup> code was used (i) to compute the gradients of grey levels in the as-treated ROI with a centred finite difference scheme, and (ii) to build from these gradients 3D structure tensors<sup>68</sup> for the  $N$  voxels  $i$  of the ROI. For that purpose, the lateral size of the Gaussian-shaped windows used to compute structure tensors was adapted to the size of the objects to be characterised: 21 voxels ( $273^3 \mu\text{m}^3$ ) for the muscle fibres of Fig. 5(b), 17 voxels ( $11^3 \mu\text{m}^3$ ) for those shown in Fig. 6(a), 21 voxels ( $13^3 \mu\text{m}^3$ ) for the ECM fibres of Figs 7(a) and 7 pixels ( $2 \mu\text{m}$ ) for those shown in Fig. 7(b). Then, the discrete 3D Orientation Distribution Function (ODF) of the minor eigen- and unit vectors  $\mathbf{p}_i$  of the structure tensors was built. Such vectors, locally parallel to the fibres, can be expressed in the reference frame of the ROI ( $\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z$ ), as follows:

$$\mathbf{p}_i = \sin \theta_i \cos \varphi_i \mathbf{e}_x + \sin \theta_i \sin \varphi_i \mathbf{e}_y + \cos \theta_i \mathbf{e}_z, \quad (2)$$

where  $0 \leq \theta_i \leq 180^\circ$  is the angle between  $\mathbf{p}_i$  and  $\mathbf{e}_z$ , and where  $0 \leq \varphi_i \leq 180^\circ$  is the angle between the projection of  $\mathbf{p}_i$  in the  $(\mathbf{e}_x, \mathbf{e}_y)$  plane and  $\mathbf{e}_x$ . The first non-zero moment of this ODF, *i.e.* the 3D second-order fibre orientation tensor  $\mathbf{A}$ <sup>69</sup>, was derived as a compact and meaningful descriptor of the fibre orientation in the ROI:

$$\mathbf{A} = \frac{1}{N} \sum_{i=1}^N \mathbf{p}_i \otimes \mathbf{p}_i. \quad (3)$$

**Waviness of fibres.** As shown in the 3D images (Figs 3(c), 4(d) and 6(a)), muscle and collagen/elastin fibres exhibit a quasi-periodic wavy shape. Thus, in a given ROI, we scrutinised two sets of orthogonal slices comprising the main fibre direction  $\mathbf{e}_m$ , *i.e.* slices within the planes  $(\mathbf{e}_m, \mathbf{e}_n)$  and  $(\mathbf{e}_m, \mathbf{e}_w)$  as sketched in Fig. 5(b):  $\mathbf{e}_m$  is defined as the main eigenvector associated to the major eigenvalue of the fibre orientation tensor  $\mathbf{A}$ ,  $\mathbf{e}_n$  and  $\mathbf{e}_w$  being eigenvectors related to the second and third eigenvalues respectively. Thereby, the spatial periods  $\lambda_{nw}$  (respectively  $\lambda_{mw}$ ) as well as the magnitudes  $R_{0nw}$  (respectively  $R_{0mw}$ ) of the quasi-periodic shape of fibres were estimated. This was done by manually clicking at least 50 data per descriptor.

**Sections of muscle fibres.** Two methods were used to analyse the size and shape of the sections of muscle fibres. Both of them used “high” resolution 3D images of the *vocalis* (voxel size of  $0.65^3 \mu\text{m}^3$ ). The first method consisted in analysing slices perpendicular to the mean orientation of muscle fibres. As depicted in the inset of Fig. 6(b), the edges of muscle fibres were detected using an edge detection subroutine (Matlab<sup>®</sup>). Then, the length  $P$  of each fibre's edges was estimated, together with the surface  $A$  of the closed area they defined, the equivalent fibre diameter  $d_e = \sqrt{4A/\pi}$ , the roundness of fibre  $\xi = 4\pi A/P^2$  and the position of their centre of mass  $x_G$ . During this process, muscle fibres that were misaligned, *i.e.*, with the tangent trajectory of their centre of mass  $x_G$  too far from

the main orientation of muscle fibres, were discarded. The same quantities were measured with the second method, which was applied along the whole cross sections perpendicular to the centreline of the muscle fibre that was manually extracted from its network (see bottom panels of Fig. 6(b)).

**Histological analyses.** A standard histological campaign with staining and 2D optical microscopy was also conducted to compare and validate the 3D images obtained with high-resolution X-ray synchrotron tomography. To this end, the vocal-fold tissue of *fresh* larynx L<sub>10</sub> was characterised with both imaging techniques, using two excised samples from the right and left vocal folds, respectively for tomography (L<sub>10</sub>-S<sub>3</sub> in Table 2) and for standard 2D histology. The latter sample, noted L<sub>10</sub>-S<sub>5</sub>, was first fixed directly after body excision by two successive baths containing a standard solution of 4% neutral buffered formalin for 1 h each (no freezing phase). Water was then removed using a series of ethanol baths (progressive concentrations of 70%, 80%, 95%, 100%, 100%, 100%; bath duration 1 h) and the sample was cleared with xylene, miscible with paraffin (two baths of 1 h). The tissue was infiltrated with liquid paraffin at 50 °C (two baths of 1 h) and then kept in a cold atmosphere (4 °C). It was then cut into 3 µm sections using a microtome, floated on a warm water bath to remove wrinkles, picked up on a glass microscopic slide and then sequentially plunged into xylene, alcohol and water baths, thereby reversing the embedding process so as to get the paraffin wax out of the tissue, and allow water-soluble dyes to penetrate the sections. Finally, different stains were chosen to enhance contrast under optical microscopy (resolution 0.2 µm): Hematoxylin-Eosin-Saffron (HES), Reticulin (modified Gordon and Sweets stain, Reticulum II staining kit, Roche) and Masson Trichrome were used to reveal both Type I and Type III collagen fibres (the latter being also called “reticular” fibres), whereas an Elastic stain allowed elastin fibres to be emphasised (see Supplementary Fig. S3).

### Data Availability

The datasets generated and/or analysed during the current study ( $\approx 1.5$  To) are available from the corresponding author on request. An open access upon the ESRF data portal is also possible, according to the ESRF data policy (<https://www.esrf.eu/datapolicy>).

### References

1. Herzel, H., Berry, D., Titze, I. & Steinecke, I. Nonlinear Dynamics of the Voice: Signal Analysis and Biomechanical Modeling. *Chaos (American Institute of Physics)* **5** (1), 30–34 (1995).
2. Titze, I. R. *Principles of voice production* (National Center for Voice and Speech, Iowa, USA, 2000), 2nd edn.
3. Lucero, J. Oscillation hysteresis in a two-mass model of the vocal folds. *Journal of Sound and Vibration* **282**, 1247–1254 (2005).
4. Hunter, E. J., Titze, I. R. & Alipour, F. A three-dimensional model of vocal fold abduction/adduction. *Journal of the Acoustical Society of America* **115**(4), 1747–1759 (2004).
5. Remacle, A., Morsomme, D. & Finck, C. Comparison of vocal loading parameters in kindergarten and elementary school teachers. *Journal of Speech, Language, and Hearing Research* **57**, 406–415 (2014).
6. Hirano, M. Morphological structure of the vocal cord as a vibrator and its variations. *Folia Phoniatrica* **26**, 89–94 (1974).
7. Garrett, C. G., Coleman, J. R. & Reinisch, L. Comparative Histology and Vibration of the Vocal Folds: Implications for Experimental Studies in Microlaryngeal Surgery. *The Laryngoscope* **110**, 814–824 (2000).
8. Madruga de Melo, E. C. *et al.* Distribution of collagen in the lamina propria of the human vocal fold. *The Laryngoscope* **113**, 2187–2191 (2003).
9. Miri, A. K., Heris, H. K., Mongeau, L. & Javid, F. Nanoscale viscoelasticity of extracellular matrix proteins in soft tissues: A multiscale approach. *Journal of the Mechanical Behavior of Biomedical Materials* **30**, 196–204 (2014).
10. Coughlan, C. *et al.* *In vivo* cross-sectional imaging of the phonating larynx using long-range Doppler optical coherence tomography. *Scientific Reports* **6**, 1–9 (2016).
11. Deliyski, D. & Hillman, R. E. State of the Art Laryngeal Imaging: Research and Clinical Implications. *Current Opinion in Otolaryngology & Head and Neck Surgery* **18**, 147–152 (2010).
12. Kobler, J. B., Zeitels, S. M., Chang, E. W. & Yun, S. H. Dynamic imaging of vocal fold oscillation with four-dimensional optical coherence tomography. *Laryngoscope* **120**, 1354–1362 (2010).
13. Chang, E. W., Kobler, J. B. & Yun, S. H. Triggered optical coherence tomography for capturing rapid periodic motion. *Scientific Reports* **1**, 1–7 (2011).
14. Maturò, S., Benboujja, F., Boudoux, C. & Hartnick, C. Quantitative distinction of unique vocal fold subepithelial architectures using optical coherence tomography. *Annals of Otolaryngology, Rhinology & Laryngology* **121**, 754–760 (2012).
15. Herrera, V. L. M. *et al.* 11.7 Tesla magnetic resonance microimaging of laryngeal tissue architecture. *Laryngoscope* **119**, 2187–2194 (2009).
16. Chen, T. *et al.* A New Method of Reconstructing the Human Laryngeal Architecture Using Micro-MRI. *Journal of Voice* **27**, 555–562 (2012).
17. Klepacek, I., Jirak, D., Duskova Smrkova, M., Janouskova, O. & Vampola, T. The Human Vocal Fold Layers. Their Delineation Inside Vocal Fold as a Background to Create 3D Digital and Synthetic Glottal Model. *Journal of Voice* **30**, 529–537 (2016).
18. Miri, A. K., Tripathy, U., Mongeau, L. & Wiseman, P. W. Nonlinear laser scanning microscopy of human vocal folds. *The Laryngoscope* **122**, 356–63 (2012).
19. Kelleher, J. E., Siegmund, T., Du, M., Naseri, E. & Chan, R. W. The anisotropic hyperelastic biomechanical response of the vocal ligament and implications for frequency regulation: A case study. *The Journal of the Acoustical Society of America* **133**, 1625–36 (2013).
20. Yildirim, M., Durr, N. & Ben-Yakar, A. Tripling the maximum imaging depth with third-harmonic generation microscopy. *Journal of Biomedical Optics* **20**, 096013–1–10 (2015).
21. Strupler, M. *et al.* *Biomedical Optics in Otorhinolaryngology: Head and Neck Surgery*, chap. Nonlinear Microscopy of the Vocal Folds, 511–528 (Springer Scienc +Business Media New York, Boston, Massachusetts, 2016).
22. Kazarine, A. *et al.* Multimodal imaging of vocal fold scarring in a rabbit model by multiphoton microscopy. In *Proc. SPIE 10069, Multiphoton Microscopy in the Biomedical Sciences XVII*, 100692E (2017).
23. Gabay, I. *et al.* Increasing the penetration depth for ultrafast laser tissue ablation using glycerol based optical clearing. In *Proc. SPIE 9707, Dynamics and Fluctuations in Biomedical Photonics XIII*, 97070X (2016).
24. Richardson, D. S. & Lichtman, J. W. Clarifying Tissue Clearing. *Cell* **162**, 246–257 (2015).
25. Laurencin, T. *et al.* 3D real-time and *in situ* characterisation of fibre kinematics in dilute non-Newtonian fibre suspensions during confined and lubricated compression flow. *Composites Science and Technology* **258**–66 (2016).
26. Paganin, D., Mayo, S. C., Gureyev, T. E., Miller, P. R. & Wilkins, S. W. Simultaneous phase and amplitude extraction from a single defocused image of a homogeneous object. *Journal of Microscopy* **206**, 33–40 (2002).

27. Vågberg, W., Larsson, D. H., Li, M., Arner, A. & Hertz, H. M. X-ray phase-contrast tomography for high-spatial-resolution zebrafish muscle imaging. *Scientific reports* **5**, 16625 (2015).
28. Dudak, J. *et al.* High-contrast X-ray micro-radiography and micro-CT of *ex-vivo* soft tissue murine organs utilizing ethanol fixation and large area photon-counting detector. *Nature Publishing Group* 1–9 (2016).
29. Shearer, T., Bradley, R. S., Araida Hidalgo-Bastida, L., Sherratt, M. J. & Cartmell, S. H. Three-dimensional visualisation of soft biological structures by X-ray computed micro-tomography. *Journal of Cell Science* **129**, 2483–2492 (2016).
30. Robert, Y. *et al.* Helical CT of the Larynx: A Comparative Study With Conventional CT Scan. *Clinical Radiology* **51**, 882–885 (1996).
31. Storck, C. *et al.* Three-dimensional imaging of the larynx for pre-operative planning of laryngeal framework surgery. *European Archives of Oto-Rhino-Laryngology* **267**, 557–563 (2010).
32. Pufe, T. *et al.* VEGF expression in adult permanent thyroid cartilage: implications for lack of cartilage ossification. *Bone* **35**, 543–552 (2004).
33. Joshi, V. M., Wadhwa, V. & Mukherji, S. K. Imaging in laryngeal cancers. *Indian J Radiol Imaging* **22**, 209–226 (2012).
34. Keyrilainen, J. *et al.* Phase-contrast X-ray imaging of breast. *Acta radiologica (Stockholm, Sweden: 1987)* **51**, 866–884 (2010).
35. Takeda, T. *et al.* Ethanol fixed brain imaging by phase-contrast X-ray technique. *Journal of Physics: Conference Series* **425**, 22004 (2013).
36. Mirea, I. *et al.* Very High-Resolution Imaging of Post-Mortem Human Cardiac Tissue Using X-Ray Phase Contrast Tomography. *Functional Imaging and Modeling of the Heart (FIMH)* **9126**, 65–73 (2015).
37. Gray, S. D., Titze, I. R., Alipour, F. & Hammond, T. H. Biomechanical and histologic observations of vocal fold fibrous proteins. *Annals of Otolaryngology, Rhinology & Laryngology* **109**, 77–85 (2000).
38. Tao, C., Jiang, J. & Zhang, Y. Simulation of vocal fold impact pressures with a self-oscillating finite-element model. *J. Acoust. Soc. Am.* **119**, 3987–3994 (2006).
39. Hirano, M. *Structure and vibratory behavior of the vocal folds*, In: *Dynamic aspects of speech production* (Tokyo University Press, 1977).
40. Roberts, T., Morton, R. & Al-Ali, S. Microstructure of the vocal fold in elderly humans. *Clinical Anatomy* **24**, 544–551, <https://doi.org/10.1002/ca.21114> (2011).
41. Hirano, M., Kukita, Y., Ohmaru, K. & Kurita, S. *Speech and language: advances in basic research and practice*, vol. 7, chap. Structure and Mechanical Properties of the Vocal Fold, 271–97 (Academic Press, 1982).
42. Hirano, M. *Vox Humana*, chap. The role of the layer structure of the vocal fold in register control, 50–62 (University of Jyväskylä, Finland, 1982).
43. Koch, R. G. *et al.* A custom image-based analysis tool for quantifying elastin and collagen micro-architecture in the wall of the human aorta from multi-photon microscopy. *Journal of Biomechanics* **47**, 935–43 (2014).
44. Holzapfel, G. Collagen in Arterial Walls: Biomechanical Aspects. In: Fratzl, P. (ed.) *Collagen Structure and Mechanics. Collagen: Structure and Mechanics* (Springer V, Heidelberg, Germany, 2008).
45. Disney, C., Lee, P. D., Hoyland, J. A., Sherratt, M. J. & Bay, B. K. A review of techniques for visualising soft tissue microstructure deformation and quantifying strain *Ex Vivo*. *Journal of Microscopy* **35**, 1–15 (2018).
46. Gramaccioni, C. *et al.* Nanoscale quantification of intracellular element concentration by X-ray fluorescence microscopy combined with X-ray phase contrast nanotomography. *Applied Physics Letters* **112**, 053701 (2018).
47. Zhang, K. & Siegmund, T. A constitutive model of the human vocal fold cover for fundamental frequency regulation. *J. Acoust. Soc. Am.* **119**, 1050–1062 (2006).
48. Miri, A. K. Mechanical characterization of vocal fold tissue: A review study. *Journal of Voice* **28**, 657–667 (2014).
49. Gunter, H. Modeling mechanical stresses as a factor in the etiology of benign vocal fold lesions. *J. Biomechanics* **37**, 1119–1124 (2004).
50. Kelleher, J. E., Zhang, K., Siegmund, T. & Chan, R. W. Spatially varying properties of the vocal ligament contribute to its eigenfrequency response. *Journal of the Mechanical Behavior of Biomedical Materials* **3**, 600–609 (2010).
51. Miri, A. K., Heris, H. K., Tripathy, U., Wiseman, P. W. & Mongeau, L. Microstructural characterization of vocal folds toward a strain-energy model of collagen remodeling. *Acta Biomaterialia* **9**, 7957–7967 (2013).
52. Morrow, D. A., Haut Donahue, T. L., Odegard, G. M. & Kaufman, K. R. Transversely isotropic tensile material properties of skeletal muscle tissue. *Journal of the Mechanical Behavior of Biomedical Materials* **3**, 124–129 (2010).
53. Bül, M., Ehret, A. E., Leichsenring, K., Weichert, C. & Kruse, R. On the anisotropy of skeletal muscle tissue under compression. *Acta Biomaterialia* **10**, 3225–34 (2014).
54. Sato, T. & Tauchi, H. Age changes in human vocal muscle. *Mechanisms of Ageing and Development* **18**, 67–74 (1982).
55. Lagier, A. *et al.* Control of the glottal configuration in *ex vivo* human models: quantitative anatomy for clinical and experimental practices. *Surgical and Radiologic Anatomy* **39**, 257–262 (2017).
56. Šidlof, P. *et al.* Geometry of human vocal folds and glottal channel for mathematical and biomechanical modeling of voice production. *Journal of Biomechanics* **41**, 985–995, <https://doi.org/10.1016/j.jbiomech.2007.12.016> (2008).
57. Vampola, T., Horacek, J. & Klepáček, J. Computer simulation of mucosal waves on vibrating human vocal folds. *Biocybernetics and biomedical engineering* **36**, 451–465 (2016).
58. Latil, P., Orgéas, L., Geindreau, C. & Dumont, P. J. J. & Rolland du Roscoat, S. Towards the 3D *in situ* characterisation of deformation micro-mechanisms within a compressed bundle of fibres. *Composites Science and Technology* **71**, 480–488 (2011).
59. Langer, M., Boistel, R., Pagot, E., Cloetens, P. & Peyrin, F. *Microscopy: Science, Technology, Applications and Education*, chap. X-ray in-line phase microtomography for biomedical applications, 391–402 (Formatex Research Center, Spain, 2010).
60. Myers, G. R., Siu, K. K. W. & Morgan, S. K. *Emerging imaging technologies in medicine*, chap. Propagation-Based Imaging, 75–91 (CRC Press, Taylor & Francis Group, 2013).
61. Khonsari, R. H. *et al.* The buccohypophyseal canal is an ancestral vertebrate trait maintained by modulation in sonic hedgehog signaling. *BMC Biology* **51**, 11–27 (2013).
62. Cupello, C. *et al.* Allometric growth in the extant coelacanth lung during ontogenetic development. *Nature Communications* **6** (2015).
63. Varray, F. *et al.* Extraction of the 3D local orientation of myocytes in human cardiac tissue using X-ray phase-contrast micro-tomography and multi-scale analysis. *Medical Image Analysis* **38**, 117–132 (2017).
64. Coan, P. *et al.* Evaluation of imaging performance of a taper optics CCD “FRELoN” camera designed for medical imaging. *Journal of Synchrotron Radiation* **13**, 260–270 (2006).
65. Labiche, J.-C. *et al.* The fast readout low noise camera as a versatile x-ray detector for time resolved dispersive extended x-ray absorption fine structure and diffraction studies of dynamic problems in materials science, chemistry, and catalysis. *Review of Scientific Instruments* **78**, 091301 (2007).
66. Hildebrand, T. & Rüesgsegger, P. A new method for the model-independent assessment of thickness in three-dimensional images. *Journal of Microscopy* **185**, 67–75 (1996).
67. Dougherty, R. & Kunzelmann, K.-H. Computing Local Thickness of 3D Structures with Image. *J. Microscopy and Microanalysis* **13**, 1678–1679 (2007).
68. Jähne, B. *Spatio-Temporal Image Processing: Theory and Scientific Applications* (Springer-Verlag, 1993).
69. Advani, S. G. & Tucker, C. L. The use of tensors to describe and predict fiber orientation in short fiber composites. *Journal of Rheology* **31**, 751–784 (1987).

## Acknowledgements

This work was supported by the ESRF proposal MD-957 and shift IN-996 (ID19 beamline), the LabEx Tec 21 (Investissements d'Avenir - grant agreement n° ANR-11-LABX-0030), the AGIR-PEPS 2015 *FIBROSCILLE* (UGA-CNRS) and the INSIS PEPS 2016 *MICROPLI* (CNRS). We would like to thank Philippe Masson, Clare Peine, Stéphane Priem, Véronique Curri, Guillaume Bravais, Maxime Terrien, Carole Neuville, Emilie Mathieu, Florian Martoia, Vincent Fernandez and Edward Andò for their helpful assistance.

## Author Contributions

L.B., L.O., N.H.B. and P.C. designed the project. L.B., T.C., L.O., N.H.B., S.R.D.R., X.L., B.F., Y.R., A.M.L. and P.C. were involved in the preparation, conception and realisation of the microtomographic experiments. Y.R. and P.C. ensured the biological tissues' excision and conservation. X.L. designed and processed the 3D-printed containment set-up for laryngeal pieces. S.R.D.R., E.B. and B.F. performed the optimisation of the ESRF optical set-ups and the X-ray scan acquisition. A.M.L. conducted the histological experiments. L.B., T.C., L.O. and N.H.B. processed the collected data and analysed the results. T.L. and T.C. were involved in the 3D orientation map calculations. L.B., T.C., L.O., N.H.B., S.R.D.R. and A.M.L. wrote and reviewed the main manuscript text.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-31849-w>.

**Competing Interests:** The authors declare 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) 2018