Reference BioImaging to assess the phenotypic trait diversity of bryophytes within the family Scapaniaceae

Authors

Kristian Peters[1,2,3], Birgitta König-Ries[1,4,5]

Affiliations

1 – German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Puschstraße 4, 04103 Leipzig, Germany

2 – Institute of Biology/Geobotany and Botanical Garden, Martin Luther University Halle-Wittenberg, Am Kirchtor 1, 06108 Halle (Saale), Germany

3 – Bioinformatics and Scientific Data, Leibniz Institute of Plant Biochemistry, Weinberg 3, 06120 Halle (Saale), Germany

4 – Heinz-Nixdorf Chair for Distributed Information Systems, Friedrich Schiller University, Jena, Germany

5 – Michael Stifel Center Jena, Jena, Germany

corresponding author(s): Kristian Peters (kpeters@ipb-halle.de)

Abstract

Macro- and microscopic images of organisms are pivotal in biodiversity research. Despite that bioimages have manifold applications such as for assessing the diversity of form and function, FAIR bioimaging data in the context of biodiversity are still very scarce, especially for difficult taxonomic groups such as bryophytes. Here, we present a high-quality reference dataset containing macroscopic and bright-field microscopic images documenting various phenotypic attributes of the species belonging to the family of Scapaniaceae occurring in Europe. To encourage data reuse in biodiversity and adjacent research areas, we annotated the imaging data with machine-actionable meta-data using community-accepted semantics. Furthermore, raw imaging data are retained and any contextual image processing like multi-focus image fusion and stitching were documented to foster good scientific practices through source tracking and provenance. The information contained in the raw images are also of particular interest for machine learning and image segmentation used in bioinformatics and computational ecology. We expect that this richly annotated reference dataset will encourage future studies to follow our principles.

Background & Summary

In biodiversity, organisms are studied with the aim to record their diversity at the genetic, metabolic, physiological, morphological, or the ecosystem level. Despite the fact that bioimaging techniques such as macro- and microscopy are prominently used, FAIR bioimaging data in the context of biodiversity are still very scarce[1–3]. This is especially the case for taxonomically difficult and underrepresented groups such as bryophytes. Currently, there are approx. 24'000 species of bryophytes known to science[4]. Unlike vascular plants, bryophytes lack well-differentiated organs that protect them from environmental exposures and pathogens. As a result, phenotypes are often cryptic and difficult to identify visually as they have developed unique specialised metabolisms and cell structures such as oil bodies[5,6].

The highly diverse family of Scapaniaceae contains 48 taxa in Europe[7] and is an ecologically important group regarding environmental adaptations[8], the biochemistry of terpenoid natural products and other chemical structures[9–12], the metabolism of pollutants and heavy metals[13,14], and phylogenetics[15–18]. Generally, there is a considerable lack of described traits

in bryophytes[19] and especially in liverworts such as Scapaniaceae, phenotypic traits to assess the diversity of form and function are understudied[20].

Bioimaging data in the field of biodiversity is of high relevance as they allow to assess the phenotypic diversity through an analysis and assessment of images[2,21,22]. In the form of measurable phenotypic traits, biological images are the groundwork of many ecological studies[3,20,21,23,24]. Phenotypisation through recording images of anatomical and morphological attributes allows qualitative and quantitative measurements of molecular structures relating to genetics, molecular pathways and biotechnology[25–27]. Bioimages have also gained a lot of interest in citizen sciences and in the digitization of natural history collections and digital herbaria[28,29]. Furthermore, meta-synthesis methods, which synthesise disparate data sources spanning published case studies, have great potential to reveal context-dependencies within bioimaging research data[30].

Macroscopy and microscopy are characterized by physical constrains resulting in diffraction and shallow depth of field[22,31,32]. From a technical perspective, our data employs two major methods to significantly extend the depth of field and increase the resolution of the composite images: image stitching (combining several images relative to the x- and y-axes of the visible accommodation to form an image with a larger frame)[33,34] and multi-focus image fusion (merging multiple images at different focal planes of the z-axis in such a way that only regions in focus will contribute to the resulting image)[35,36]. In this regard, the raw data also allows to be reused for combining image fusion with computational super-resolution[22,37–40].

Cloud infrastructural resources are able to execute computational workflows that combine data with computational analysis tools at a large scale[41,42]. However, there is still a considerable lack of data containing machine-actionable meta-data[1,3,43–46]. To foster provenance, any raw and segmented image in this data set has been associated with a rich set of contextual and expressive meta-data[47], documenting the phenotypic attributes, and recording any digital image processing (i.e., increasing contrast, brightness, image fusion). The meta-data has been formatted with community-accepted semantics that allow for machine-actionable data-mining and to create scientific workflow modules that produce segmented composite images automatically by reusing the instructions contained in our meta-data[20,42,45,46,48].

In this Data Descriptor, we present the principles to generate reference images from raw microscopic bioimaging data and show how individual images are associated with technical and expressive meta-data. Despite that we were able to associate our images with a rich set of meta-data, we ascertain that there is still a lack of usable ontological terms and schemas in bioimaging with regard to documenting image processing and associating individual images with phenotypic attributes[3,45]. Our high-resolution images allow for large prints and zooming into images to obtain critical details, which is particularly important for species identification and for computational image analyses, computer-assisted species recognition and identification[49,50]. Despite that we have deposited the data in the two specialised imaging repositories BioStudies (containing raw and pre-processed images which enable rapid use in, i.e., machine learning approaches in computational ecology) and Imaging Data Resource (containing pre-processed and fully segmented images to be rapidly reused by ecologists), we ascertain that there is still the need for connecting macro- and microscopic bioimaging data to biodiversity platforms[3,51] such as iDigBio[52] and GBIF[53], or even the citizen scientists community-effort iNaturalist[54]. Our reference data framework facilitates the further integration of bioimaging data into other research disciplines[55] and, thus, we want to inspire future data reuse and meta-synthesis in the fields of biodiversity and computational ecology.

Methods

Representative voucher specimens were received from different herbaria. Table 1 lists all used voucher specimen and freshly collected samples. Fresh samples of *Diplophyllum taxifolium*, *Scapania cuspiduligera*, *Scapania gymnostomophila* and *Scapania subalpina* were additionally collected at various sites, put into envelopes on-site, identified and photographed afterwards. Information regarding the date, site (including geographical coordinates), habitat, substrate and other further information were collected.

Species names	Taxonomy IDs	Reference of voucher specimens	Voucher identifiers
Diplophyllum albicans (L.) Dumort.	NCBI:264775 GBIF:5286350 OTT:1055640	Urmi, E. (1999): Schweiz, Kt. Ticino, Caviono: Valle Di S. Abbondio, 704.40/106.98, ca. 850 müM	Urmi-8371
Diplophyllum obtusatum (R.M.Schust.) R.M.Schust.	NCBI:2212631 GBIF:5710146 OTT:3855725	Urmi, E. (2017): Schweiz, Kt. Ticino, Chiasso: im Bosca Penz oberhalb Bresciano, 721.27/76.88, ca. 300 müM	Urmi-10695
Diplophyllum obtusifolium (Hook.) Dumort.	NCBI:248354 GBIF:5710147 OTT:831945	Urmi, E. (1988): Schweiz, Kt. Ticino, Cugnasco: Alpe di Ruscada oberhalb Corte di mezzo, 715.00/119.00, ca. 1550 müM	Urmi-4501
Diplophyllum taxifolium (Wahlenb.) Dumort.	NCBI:248355 GBIF:5710140 OTT:831946	Urmi, E. (1975): Italien, Ot. Crigiono Mesocco, oberhalb San Bernardina, am linken Seitenback der Moesa nordwestlich von Careida Sot, Landeskarte 734.34/147.91, Felsnische an der Seitenwand der kleinen Schlucht, ca. 1680 müM.	Urmi-1060
		Peters, K. (2020): Schweiz, in Erdloch unter Felsspalte, 46.083813428, 7.923810274, 2614m +- 5m	Peters-2020-21
Douinia ovata (Dicks.) H.Buch	NCBI:248357 GBIF:2689235 OTT:831948	Urmi, E. (1980): Spanien, Prov. Oviedo (Asturias), Cangas de Narcea: Monte de Muniellos bei Tablizas oberhalb Venta Nueva, v. Madrid: 3°0'20''W / 43°2'10''N, ca. 820 müM, an trockenen Felsen	Urmi-2056
Scapania aequiloba (Schwägr.) Dumort.	NCBI:464313 GBIF:7696978 OTT:382426	Urmi, E. (1976): Schweiz, Kt. Zürich: Hütten, Chrüzbrunnen am Höhronen-Nordhang, auf anstehendem Molasse-Sandstein im Schluchtwald, ca. 910 müM, Ost-Exposition	Urmi-1111
		Urmi, E. (2018): Schweiz, Kt. Uri, Spiringen (Enklave Urnerboden): unterhalb Waldrüti, 712.09/194.26, ca. 1340 müM, Straßenböschung, Kalkfelsen	Urmi-10772
Scapania apiculata Spruce	NCBI:537844 GBIF:4277163 OTT:655684	Hürlimann, H. (1983): Japan, an morschom, trockenem Stamm in Birkenwald beim 5-gome am Mt. Fuji, Japan, ca. 2200 müM	Z-000130732
Scapania aspera M.Bernet & Bernet	NCBI:399155 GBIF:7435100 OTT:683885	Bertram, J. (2007): Schweiz, SO, Seewen, SW Fulnau, E Risenberg, 613.845/253.87 680 müM, Eingang zu einer Gasse zwischen steilen Felswänden im Fagetum, Felsabsatz, auf kantig-rauhem Kalkgestein	Bertram-3283a
Scapania brevicaulis Taylor	GBIF:4277188 OTT:3855710	Macurian, S.M. (1900)	Zürich-00000
Scapania calcicola (Arnell & J.Perss.) Ingham	NCBI:537846 GBIF:4277184 OTT:655680	Hürlimann, H. (1968): Schweiz, Schattiger Kalkfelsen im obersten Teil des Berghügels Ramstein bei Bretzol BL	Z-000129406
Scapania carinthiaca var. *massalongi* J.B.Jack ex Lindb.	GBIF:8320028	Meier, M. (2019): Schweiz, Kt. SZ, Gemeinde Sattel, Bärenfang, Zufluss zum Lauitobelbach, Koord: 692.912/213.421+-6m, Höhe: 1300 müM, Waldbach, vermodernder, ins Bachbett ragender Stamm einer Fichte oder Tanne	ZT-337051
		Høitomt, T. (2018): Norway: Oppland: Vestre Slidre, Mosåni, På dødved av gran langs elva, Host: Picea abies, UTM: 499542,6769666, Alt: 400m, 3.7.2018	TRH-B108431
Scapania compacta (Roth) Dumort.	NCBI:402640 GBIF:9148635 OTT:168110	Urmi, E. (1980): Italien, Sardegna, Prov. di Nuoro, Airtzo oberhalb des Dorfes am Weg nach Badu Arala, Ret. ital.: 1517.8 E / 4423.4 N, ca. 1020 müM, bei einem Brunnen an Felsen (kalkarm)	Urmi-1951

Scapania crassiretis Bryhn	NCBI:248358 GBIF:4277199 OTT:831941	Hürlimann, H. (1947): Italien, auf der Nordseite der Talsperre von Caraglia GR	Hürlimann-H252
Scapania curta (Mart.) Dumort.	NCBI:537847 GBIF:2689216 OTT:655678	Culmann, P. (1998)	ZT-7/93-5
Scapania cuspiduligera (Nees) Müll.Frib.	NCBI:537848 GBIF:8306353 OTT:655688	Urmi, E. (1990): Schweiz, Kt. Wallis, Evolene: Comba Louva am linken Talhang, CH1903/LV03: 601.00 / 109.00, ca. 1860 müM, Nadelwald offene Stelle auf stellenweise kalkreichem Stein	Urmi-5345
		Peters, K. (2020): Schweiz, auf Erde, Hang, 46.037564627, 7.954650076, 2240m, +-5m	Peters-2020-09
Scapania degenii Schiffn. ex Müll.Frib.	NCBI:537848 GBIF:8306353 OTT:655688	Gustafsson, L. (1989): Jämtland, Undersåker par., Ristafallen, S-side inbetween rocks near the stream, below the fall	CP-CM30374
Scapania glaucocephala (Taylor) Austin	NCBI:537849 GBIF:4277200 OTT:655686	Schuster, R.M. (1956): USA, Decaying logs, shaded swamp along small stream, Sturgeon Bay in Wildwood State Park, Emmet Co., Michigan	CP-CM30377
		Høitomt, T. (2018): Norway: Oppland: Etnedal, Fjellselve, På låg av osp i elvekant, Host: Populus tremula, UTM-WGS84: 534192,6750148, Alt: 375	TRH-B108428
Scapania gracilis Lindb.	NCBI:264776 GBIF:5837556 OTT:844266	Urmi, E. (1982): Spanien, Prov. Oviedo (Asturias), Ponga ca 2.5km oberhakb Sellaño, am Rio Ponga (rechte Seite) v. Madrid: 1°31'10''W, 43°13'55''N, ca 280 müM, sehr steiler felsiger Hang, Felsstufe, kalkarm	Urmi-2211
Scapania gymnostomophila Kaal.	NCBI:537850 GBIF:4277196 OTT:133595	Urmi, E. (2016): Schweiz, Kt. Grischun, Valsot am Strässchen nördl. v. Dorf Tschlin, 827.66/195.71, ca. 1550 müM, Wandrand an der Strasse, schieferige Kalkfelsen	Urmi-10478
		Peters, K. (2020): Schweiz, auf Erde, west-exponierter Hang, 46.039167304, 7.954680641, 2237m, +-10m	Peters-2020-16
Scapania helvetica Gottsche	NCBI:537851 GBIF:7787988 OTT:133598	Urmi, E. (1998): Schweiz, Kt. Wallis, Zwischbergen: oberhalb Bidemji im Laggintal, CH1903/LV03: 647.00/111.00, ca 2100 müM, felsiger Hang mit lockerem Grünerlen-Gebüsch, über Silikatfelsen auf Sand	Urmi-8131
		Schnyder, N. (2012): Schweiz, UR, Realp, Furkapass, S Passhöhe, See bei Pt. 2650, 675.98/157.24, 2650m, Schneeboden feuchte Erde	ZT-210488
Scapania hyperborea Jørg.	NCBI:402139 GBIF:8336103 OTT:247549	Hallingbäck, T. (2019): Province: Jämtland. Parish: Åre, Locality: Staån, myrvegetationen, Long: 12.27523, Lat: 63.59544	Hallingbäck-50160
Scapania irrigua (Nees) Nees	NCBI:537852 GBIF:2689217 OTT:133601	Urmi, E. (1992): Schweiz, Kt. Uri, Wassen: oberhalb Kartigel im Meiental, 684/174, ca 1950 müM, grasiger Hang mit felsiger Rinne, feuchtes kalkarmes Gestein	Urmi-6100
		Urmi, E. (2016): Schweiz, Kt. Graubünden, Val Müstair Buffalora, westlich vom Ofenpass 816.79/168.69, ca. 2110 müM, Flachmoor, Torfmoosbult	Urmi-10576
Scapania irrigua subsp. *irrigua* (Nees) Nees	GBIF:7420792	Söderström, L. (1991): Sweden, Västerbotten, Sävar sn, Grisselögern, Bland stenar på stranden (Ruta i rikets nät: 20K67)	UME-78146
Scapania irrigua subsp. *rufescens* (Loeske) R.M.Schust.	GBIF:8087753	Sørensen, T. (1933): Liverpool Land: Cape Hope, Lat 70.28, Long: 220.25 W	CP-CM30429
Scapania kaurinii Ryan	NCBI:537853 GBIF:4277164 OTT:133603	Laine, T. (1966): Enontekiö Lapland, Enontekiö, Porojärvat, Ridnitsohka mountain, In alpine top area, on soil, 1280m above sea level	Z-026-1966
Scapania ligulifolia R.M.Schust.	NCBI:1230444 GBIF:8288472 OTT:279012	Frisvoll, A.A. (1973): Svalbard: Dickson Land, Nordfjorden NØ. Vestskråningen av Kongressfjellet litt S for toppen, 300-400 moh	TRH-74054
Scapania lingulata H.Buch	NCBI:402641 GBIF:4277202 OTT:2063	Buch, H. (1919): Finlandia Regio aboensis, Insula Hirvensalo, St. Marie, ad rupes	Zürich Buch-1919
Scapania mucronata H.Buch	NCBI:537854 GBIF:4277191 OTT:133587	Urmi, E. (1992): Schweiz, Kt. Uri, Realp: ob den Galenbödmen, 675.00 / 159.00; ca 2530 m ü. M.	Urmi-6291
		Schnyder, N. (2003): Schweiz, NE Fleurier, Pouetta Raisse, 536.6 / 193.2, 1010 m	Schnyder-47044
		Buch, H. (1922): Finland, Regio aboensis, Angelniemi	Zürich, Buch-

			1922
Scapania nemorea (L.) Grolle	NCBI:41848 GBIF:2689220 OTT:851174	Urmi, E. (1981): Schweiz, Kt. Obwalden, Giswil: Altibach-Tobel beim Kleinteil, 654.36 / 186.25; ca. 650 m ü M.	Urmi-2110
Scapania nimbosa Taylor	NCBI:1133313 GBIF:8635654 OTT:1062948	Wallace, E. (1955)	Zürich Wallace-1955
Scapania obcordata (Berggr.) S.W.Arnell	NCBI:402140 GBIF:4277171 OTT:719916	Zemp, F. (2019): Schweiz, Kanton: Be, Gemeinde: Innertkirchen, Sustenpass, Steinsee, Koord: 675.220/174.460, Höhe: 2050m, Standort: Bachalluvion	Zemp-337400
Scapania obscura (Arnell & C.E.O.Jensen) Schiffn.	NCBI:1282292 GBIF:4277165 OTT:558046	Urmi, E. (2019): Schweiz Kt. Wallis, Oberwald beim Totensee auf dem Grimselpass, CH1903/LV03: 669.03 / 156.76, ca. 2170 müM, kleines saures Flachmoor neben dem Bach, relativ grober Silikat-Sand	Urmi-10928
		Culmann, P. (1908): Schweiz, Wallis, Grimsel, prope lac. d. Todtensee, in paludosis juxta rivulum, hygrophila, 2150m	Zürich-0000
Scapania ornithopodioides (With.) Waddell	NCBI:264777 GBIF:7512997 OTT:333110	Macricar, S.M. (1898):	ZT-1898
Scapania paludicola Loeske & Müll.Frib.	NCBI:209818 GBIF:4277131 OTT:245460	Urmi, E. (1985): Schweiz, Kt. Obwalden, Engelberg Feldmoos auf Gerschni, 673.70/184.82, ca. 1230 müM, Moor, sehr nasser Torf	Urmi-3307
Scapania paludosa (Müll.Frib.) Müll.Frib.	NCBI:537855 GBIF:4277179 OTT:133589	Urmi, E. (2019): Schweiz, Kt. Wallis, Oberwald beim Totensee auf dem Gimselpass, CG1903/LV03: 668.72/156.92, ca 2170 müM, moorige Stelle an einem Bach, teils im Wasser stehend, Silikat-Sand	Urmi-10924
Scapania parvifolia Warnst.	GBIF:4277175 OTT:3855712	Marko Lewis, 12.7.1976, Abundand amongst rocks near snow beds. Upper Ambler river, ridge above Kowalskis Creek, Mt. 3950, 67°19'N, 156°59'W. Late snow bed on N. side of alpine ridge, wet or moise, 2000 ft	CP-CM30382
Scapania praetervisa Meyl.	GBIF:4277193 OTT:7571327	Urmi, E. (1982): Schweiz, Kt. Grischun, Lumbrein hart an der Gemeindegrenze oberhalb Schlareins, CH1903/LV03: 729.14 / 173.20, ca. 1750 müM, am Anstehenden (schattig), kalkarmes schieferiges Gestein mit Ausblühungen	Urmi-2161
Scapania scandica (Arnell & H.Buch) Macvicar	NCBI:1133317 GBIF:4277166 OTT:1062956	Düll R. (1971): Schweiz, Kt. Uri, 8 km nördlich d. St. Gotthartpasses nahe d. Str. Grundgeb., ca 1570m	ZT-017155
		Urmi, E. (1999): Schweiz Kt. Ticino, Caviano Valle de S. Abbondio, CH1903/LV03: 704.40/106.98, ca. 850 müM, grasiger Felshang mit wenig Gebüsch, anstehender Silikatfels, z.T. mit etwas Feinmaterial	Urmi-8376
Scapania simmonsii Bryhn & Kaal.	NCBI:537857 GBIF:4277140 OTT:133593	Lewis, M. (1976): Upper Ambler River, 0.5 miles N. of Ulaneak Creek, base of Grey Mt. Tocky limestone slope	CP-CM30391
		Kjeld Holmen, 15.7.1961, Alaska: Old John Lake, Lat. 68.08N, Long. 145 W, Alt 0m	CM-30384
Scapania sphaerifera H.Buch & Tuom.	NCBI:537858 GBIF:5793125 OTT:133605	Bakalin, V. (2010): Russian Far East. Primorsky Territory. Shkotovsky, District, Pidan Mt., N-facing slope in upper course of Oyry Stream. Merely dry boulders along stream in Picea-Abies forest, WGS84: 43.083N, 132.7E	TRH-B9515
Scapania spitsbergensis (Lindb.) Müll.Frib.	NCBI:537859 GBIF:8990537 OTT:133607	Laine, T. (1966): Enontekiö Lapland, Enontekiö, Porojärvet, Ridnitsohka mountain, in alpine top area, between boulders, 1250m above sea level	ZT-026-1966
Scapania subalpina (Nees ex Lindenb.) Dumort.	NCBI:248359 GBIF:4277176 OTT:831943	Hürlimann (1947)	Z-000129362
		Hangartner, R. (2008): GR Medel (Lucmagn), ob Masauna, 710.60/170.65, 1720m, Felsen feuchter Glimmerschiefer	ZT-101909-159278
		Hürlimann, H. (1975): Nasser Fels zwischen Restaurant und Staumauer Zervreila (Vals GR), ca. 1850müM	Z-000129211
		Peters, K. (2020): Schweiz, auf Erde in Felsspalte, 46.08316087, 7.923246937, 2585m, +-5m	Peters-2020-23
Scapania tundrae (Arnell) H.Buch	NCBI:537860 GBIF:8547629 OTT:707922	Sørensen, T. (1933): Liverpool Land: Cape Hope, Lat 70°28'N, Long: 220°25'W	CP-CM30429
Scapania uliginosa (Lindenb.) Dumort.	NCBI:537861 GBIF:8368371 OTT:707920	Urmi, E. (1984): Schweiz, Kt. Uri, Erstfeld Oberstafel (Riedbarg), 691.18 / 183.94, ca. 1530 müM, saures Hangmoor, in Sphagnum-Polster	Urmi-2825

Scapania umbrosa (Schrad.) Dumort.	NCBI:537862 GBIF:4277183 OTT:707926	Urmi, E. (1985): Schweiz, Nidwalden, Oberdorf oberhalb Wolfboden, CH1903+/LV95: 671.00 / 199.00, ca. 1340 müM, Lichter Fichtenwald an steilem Hang, auf totem Holz	Urmi-3527
		Edwin Urmi, 26.8.1983, Schweiz, Kt. Bern, Lenk Iffigtal bei der Iffigenalp, 599.88 / 138.6, ca. 1620 müM, im Fichtenwald, auf totem Baumstumpf	Urmi-2548
Scapania undulata (L.) Dumort.	NCBI:215256 GBIF:2689218 OTT:812077	Bertram, J. (1996): VS Ried-Mörel, Reservat Aletschwald, Mittlerer Aletschwald, 645.54/138.08, 1970m, kleines Bächlein in einem Tälchen mit Alnetum viridis-Bestand auf Steinen im Bach, z.T. submers bzw. überspült	Bertram-1898e
		Hürlimann, H. (1994): Fundort, Nasser schattiger Granitfels, Bondo GR im Bondasca-Tal bei Stauwehr-Brücke, 1080müM, 76370/13288	Hürlimann-94211
Scapania verrucosa Heeg	NCBI:537863 GBIF:8119128 OTT:707924	Urmi, E. (1976): Italien, Piemonte, Prov. di Novara, Valstrona, am rechten Ufer des T. Strona westlich von Campello Monti, ret. ital.: fuso ovest 1440.8E / 5087.2 N, Bachböschung auf anstehendem Peridotit, ca 1320 müM	Urmi-1173
		P. Culmann, 22.8.1920 CM-30381	
Scapania zemliae S.W.Arnell	GBIF:7480007 OTT:7571345	Holmen, K. (1958): From brooks below a glacier-tongue, Charcots Land: Lat 71°54'N, Long 29°W, Alt 850m	CP-CM30431

Table 1: List of voucher specimens and fresh samples used in this study. All voucher specimens have been investigated. The columns list the taxonomic species identifiers (*NCBI*, *GBIF*, or *Open Tree of Life* identifiers, if available), the text on the specimen sleeves (collector, date and text on the envelopes) and the voucher specimen identifiers (the first letters either indicate the *Index Herbariorum* institution code[56] if available or the name of private collection where the specimens were stored).

For microscopy, a Zeiss Axio Scope.A1 HAL 100/HBO, 6x HD/DIC, M27, 10x/23 microscope with an achromatic-aplanatic 0.9 H D Ph DIC condenser was used with the objectives EC Plan Neofluar 2.5x/0.075 M27 (a=8.8mm), Plan-Apochromat 5x/0.16 M27 (a=12.1mm), Plan-Apochromat 10x/0.45 M27 (a=2.1mm), Plan-Apochromat 20x/0.8 M27 (a=0.55mm), and Plan-Apochromat 40x/0.95 Korr M27 (a=0.25mm) using the EC PN and the Fluar 40x/1.30 III and PA 40x/0.95 III filters for DIC. The conversion filter CB3 and the interference filter wideband green were used to improve digital reproduction of colors. For macroscopy and for preparing microscopy slides, a binocular microscope Zeiss Stemi 2000c was used (apochromatic Greenough system with a stereo angle of 11° and 100/100 switchover of camera and ocular viewing). For stand-alone macroscopic images, the objectives Canon MP-E 65mm 2.8 1-5x macro and Laowa 25mm 2.5-5.0x ultra-macro for Canon EF and the Canon EF-RF adapter.

To acquire digital images, a full-frame, high-resolution camera (Canon EOS RP, 26 megapixel) was used and adapted to the microscopes using binocular phototubes with sliding prism 30°/23 (Axio Scope.A1) and 100:0/0:100 reversed image (Stemi 2000c) using 60-T2 camera adapter for Canon EOS and Canon EF-RF adapter. The objectives Canon MP-E and Laowa 25mm were adapted directly through the Canon EF-RF adapter.

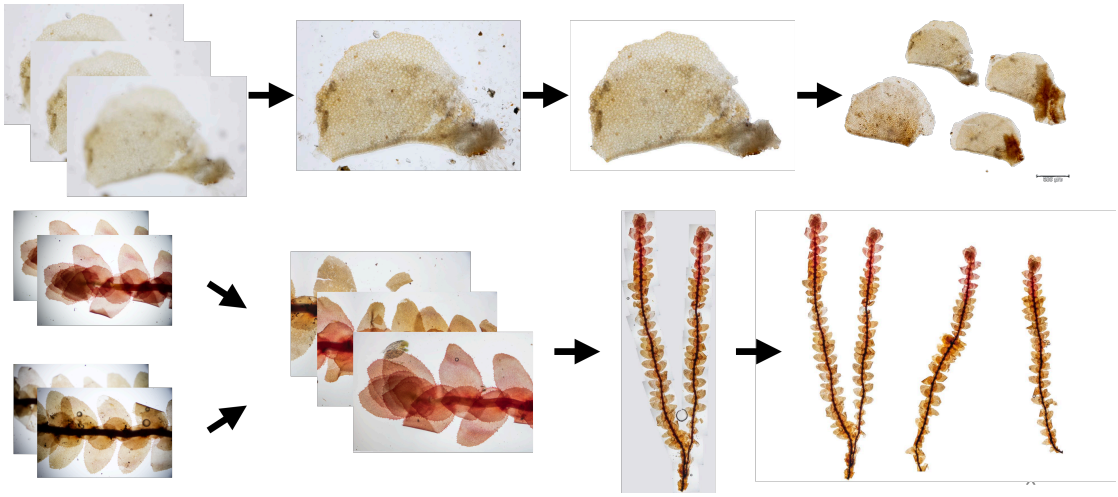

Figure 1: Two exemplary workflows used in this study to create segmented images. **(a)** Example of multi-focus image fusion where (1) several images of one object of a leaf lobe are fused into a (2) composite image. The leaf lobe as shown in the composite image is then (3) segmented and the background removed. Several leaf lobe objects are then put onto the (4) final image and a microscopic scale is applied. **(b)** Example of image stitching where (2) several fused images from (1) different segments of the same object are (3) stitched into a composite image with larger dimensions. Several of these stitched images are then put onto the (4) final image and a microscopic scale is applied.

To construct images with extended depth of field using computational methods, multiple images were recorded at different focal planes. This "focus stacking" approach was automatized for macroscopy by attaching the camera to a Cognisys StackShot macro rail fixed on a Novoflex macro stand, and for microscopy by adapting a Cognisys StackShot motor to the fine adjustment of the microscope using two cogged wheels, one small wheel (1 cm diameter) adapted on the motor and one large wheel (8.5 cm diameter) on the fine adjustment of the microscope. The two cogged wheels were coupled with a toothed belt to obtain very fine step increments of the stepping motor for high magnifications. A Cognisys StackShot controller was used to control the amount and distance of the stepping motor with the following controller settings: Dist/Rev: 3200 stp, Backlash: 0 steps, # pics: 1, Tsettle: 100.0 ms, Toff: 450.0 ms, Auto Return: yes, Speed: 3000 st/sec, Tlapse: off, Tpulse: 800.0 ms, Tramp: 100 ms, Units: steps, Torque: 6, Hi Precision: Off, LCD Backlight: 10, Mode: Auto-Step using between 25 steps (magnification 1x) and 50 steps (magnification 25x) and 100 steps (magnification 400x) (number of steps depending on aperture settings and effective magnification).

Raw images were recorded in CR3-format and pre-processed with Adobe Camera RAW. Non-destructive image processing such as corrections of the field curvature, removal of chromatic aberration, increase of contrast and brightness were performed in Adobe Camera RAW. Images were then exported to TIFF-format and any image prcoessing steps were recorded in individual Adobe XMP-files.

Multi-focus image fusion was performed on the individual images in the z-stacks using the software Helicon Focus 7.7.5 and by choosing the algorithms *depth map* and *pyramid* with different settings of radius (4, 8, 16, 24) and smoothing (2, 4). The best composite image was chosen manually and retained. When composite images contained specimen that were larger than the frame, several images were stitched together using the *panorama stitching* function in the software Affinity Photo 1.10.1.

Images were manually segmented and interfering background removed using the *flood select*, *brush selection* and *freehand selection* tools in the software Affinity Photo. In order to determine the scale, a stage micrometre was photographed separately with any of the objectives and microscope combinations. The scale was calculated per pixel for each combination (File scale_bar_distances.csv in [57]) and scale bars were put post-hoc onto the segmented images using the Python script scale_bar.py[57].

Meta-data including species name, taxonomic rank information (NCBI-Taxon, GBIF and OTT taxonomy identifiers), voucher specimen id, image acquisition date, an object description including the name of the captured phenotypic attribute(s), the used objective, microscope, and magnification were associated with any raw image based on unique respective file names. Individual file names (variable *file list*), name within an image focus stack (variable *stack name*) and name within an image stitching stack (variable *stitch name*) were recorded additionally to facilitate subsequent automized image processing in workflows. A Python script was created to put individual images as part of image stacks into directories (File create_image_stacks.py in [58]). The Python script parses the *Label* tag in the XMP-files. Any meta-data regarding image enhancement and non-destructive image processing were extracted from XMP-files using a simple Python script (File xmp_stack_to_tsv.py in [58]). The meta-data was saved in individual TSV-files and merged using a helper Python script (File tsv_merge.py in [58]). Table 2 lists all fields which were extracted from the XMP.

Semantic identifier	**Tag name**
{http://www.w3.org/1999/02/22-rdf-syntax-ns#}	about
{http://ns.adobe.com/xap/1.0/}	ModifyDate
{http://ns.adobe.com/xap/1.0/}	CreateDate
{http://ns.adobe.com/xap/1.0/}	MetadataDate
{http://ns.adobe.com/xap/1.0/}	Rating
{http://ns.adobe.com/tiff/1.0/}	Make
{http://ns.adobe.com/tiff/1.0/}	Model
{http://ns.adobe.com/tiff/1.0/}	Orientation
{http://ns.adobe.com/tiff/1.0/}	ImageWidth
{http://ns.adobe.com/tiff/1.0/}	ImageLength
{http://ns.adobe.com/exif/1.0/}	ExifVersion
{http://ns.adobe.com/exif/1.0/}	ExposureTime
{http://ns.adobe.com/exif/1.0/}	ShutterSpeedValue
{http://ns.adobe.com/exif/1.0/}	FNumber
{http://ns.adobe.com/exif/1.0/}	ApertureValue
{http://ns.adobe.com/exif/1.0/}	ExposureProgram
{http://ns.adobe.com/exif/1.0/}	SensitivityType
{http://ns.adobe.com/exif/1.0/}	RecommendedExposureIndex
{http://ns.adobe.com/exif/1.0/}	ExposureBiasValue
{http://ns.adobe.com/exif/1.0/}	MaxApertureValue
{http://ns.adobe.com/exif/1.0/}	MeteringMode
{http://ns.adobe.com/exif/1.0/}	FocalLength
{http://ns.adobe.com/exif/1.0/}	CustomRendered
{http://ns.adobe.com/exif/1.0/}	ExposureMode
{http://ns.adobe.com/exif/1.0/}	WhiteBalance
{http://ns.adobe.com/exif/1.0/}	SceneCaptureType
{http://ns.adobe.com/exif/1.0/}	FocalPlaneXResolution
{http://ns.adobe.com/exif/1.0/}	FocalPlaneYResolution
{http://ns.adobe.com/exif/1.0/}	FocalPlaneResolutionUnit
{http://ns.adobe.com/exif/1.0/}	DateTimeOriginal
{http://ns.adobe.com/exif/1.0/}	PixelXDimension
{http://ns.adobe.com/exif/1.0/}	PixelYDimension
{http://purl.org/dc/elements/1.1/}	format
{http://ns.adobe.com/exif/1.0/aux/}	SerialNumber
{http://ns.adobe.com/exif/1.0/aux/}	LensInfo
{http://ns.adobe.com/exif/1.0/aux/}	Lens
{http://ns.adobe.com/exif/1.0/aux/}	LensID

{http://ns.adobe.com/exif/1.0/aux/}	LensSerialNumber
{http://ns.adobe.com/exif/1.0/aux/}	ImageNumber
{http://ns.adobe.com/exif/1.0/aux/}	ApproximateFocusDistance
{http://ns.adobe.com/exif/1.0/aux/}	FlashCompensation
{http://ns.adobe.com/exif/1.0/aux/}	Firmware
{http://cipa.jp/exif/1.0/}	LensModel
{http://ns.adobe.com/photoshop/1.0/}	DateCreated
{http://ns.adobe.com/photoshop/1.0/}	SidecarForExtension
{http://ns.adobe.com/photoshop/1.0/}	EmbeddedXMPDigest
{http://ns.adobe.com/xap/1.0/mm/}	DocumentID
{http://ns.adobe.com/xap/1.0/mm/}	PreservedFileName
{http://ns.adobe.com/xap/1.0/mm/}	OriginalDocumentID
{http://ns.adobe.com/xap/1.0/mm/}	InstanceID
{http://ns.adobe.com/camera-raw-settings/1.0/}	Version
{http://ns.adobe.com/camera-raw-settings/1.0/}	ProcessVersion
{http://ns.adobe.com/camera-raw-settings/1.0/}	WhiteBalance
{http://ns.adobe.com/camera-raw-settings/1.0/}	Temperature
{http://ns.adobe.com/camera-raw-settings/1.0/}	Tint
{http://ns.adobe.com/camera-raw-settings/1.0/}	Sharpness
{http://ns.adobe.com/camera-raw-settings/1.0/}	LuminanceSmoothing
{http://ns.adobe.com/camera-raw-settings/1.0/}	ColorNoiseReduction
{http://ns.adobe.com/camera-raw-settings/1.0/}	VignetteAmount
{http://ns.adobe.com/camera-raw-settings/1.0/}	ShadowTint
{http://ns.adobe.com/camera-raw-settings/1.0/}	RedHue
{http://ns.adobe.com/camera-raw-settings/1.0/}	RedSaturation
{http://ns.adobe.com/camera-raw-settings/1.0/}	GreenHue
{http://ns.adobe.com/camera-raw-settings/1.0/}	GreenSaturation
{http://ns.adobe.com/camera-raw-settings/1.0/}	BlueHue
{http://ns.adobe.com/camera-raw-settings/1.0/}	BlueSaturation
{http://ns.adobe.com/camera-raw-settings/1.0/}	GrayMixerRed
{http://ns.adobe.com/camera-raw-settings/1.0/}	GrayMixerOrange
{http://ns.adobe.com/camera-raw-settings/1.0/}	GrayMixerYellow
{http://ns.adobe.com/camera-raw-settings/1.0/}	GrayMixerGreen
{http://ns.adobe.com/camera-raw-settings/1.0/}	GrayMixerAqua
{http://ns.adobe.com/camera-raw-settings/1.0/}	GrayMixerBlue
{http://ns.adobe.com/camera-raw-settings/1.0/}	GrayMixerPurple
{http://ns.adobe.com/camera-raw-settings/1.0/}	GrayMixerMagenta
{http://ns.adobe.com/camera-raw-settings/1.0/}	SplitToningShadowHue
{http://ns.adobe.com/camera-raw-settings/1.0/}	SplitToningShadowSaturation
{http://ns.adobe.com/camera-raw-settings/1.0/}	SplitToningHighlightHue
{http://ns.adobe.com/camera-raw-settings/1.0/}	SplitToningHighlightSaturation
{http://ns.adobe.com/camera-raw-settings/1.0/}	SplitToningBalance
{http://ns.adobe.com/camera-raw-settings/1.0/}	ParametricShadows
{http://ns.adobe.com/camera-raw-settings/1.0/}	ParametricDarks
{http://ns.adobe.com/camera-raw-settings/1.0/}	ParametricLights
{http://ns.adobe.com/camera-raw-settings/1.0/}	ParametricHighlights
{http://ns.adobe.com/camera-raw-settings/1.0/}	ParametricShadowSplit
{http://ns.adobe.com/camera-raw-settings/1.0/}	ParametricMidtoneSplit
{http://ns.adobe.com/camera-raw-settings/1.0/}	ParametricHighlightSplit
{http://ns.adobe.com/camera-raw-settings/1.0/}	SharpenRadius
{http://ns.adobe.com/camera-raw-settings/1.0/}	SharpenDetail
{http://ns.adobe.com/camera-raw-settings/1.0/}	SharpenEdgeMasking
{http://ns.adobe.com/camera-raw-settings/1.0/}	PostCropVignetteAmount
{http://ns.adobe.com/camera-raw-settings/1.0/}	GrainAmount
{http://ns.adobe.com/camera-raw-settings/1.0/}	ColorNoiseReductionDetail
{http://ns.adobe.com/camera-raw-settings/1.0/}	ColorNoiseReductionSmoothness
{http://ns.adobe.com/camera-raw-settings/1.0/}	LensProfileEnable
{http://ns.adobe.com/camera-raw-settings/1.0/}	LensManualDistortionAmount
{http://ns.adobe.com/camera-raw-settings/1.0/}	PerspectiveVertical
{http://ns.adobe.com/camera-raw-settings/1.0/}	PerspectiveHorizontal
{http://ns.adobe.com/camera-raw-settings/1.0/}	PerspectiveRotate
{http://ns.adobe.com/camera-raw-settings/1.0/}	PerspectiveScale
{http://ns.adobe.com/camera-raw-settings/1.0/}	PerspectiveAspect
{http://ns.adobe.com/camera-raw-settings/1.0/}	PerspectiveUpright

{http://ns.adobe.com/camera-raw-settings/1.0/}	PerspectiveX
{http://ns.adobe.com/camera-raw-settings/1.0/}	PerspectiveY
{http://ns.adobe.com/camera-raw-settings/1.0/}	AutoLateralCA
{http://ns.adobe.com/camera-raw-settings/1.0/}	Exposure2012
{http://ns.adobe.com/camera-raw-settings/1.0/}	Contrast2012
{http://ns.adobe.com/camera-raw-settings/1.0/}	Highlights2012
{http://ns.adobe.com/camera-raw-settings/1.0/}	Shadows2012
{http://ns.adobe.com/camera-raw-settings/1.0/}	Whites2012
{http://ns.adobe.com/camera-raw-settings/1.0/}	Blacks2012
{http://ns.adobe.com/camera-raw-settings/1.0/}	Clarity2012
{http://ns.adobe.com/camera-raw-settings/1.0/}	Dehaze
{http://ns.adobe.com/camera-raw-settings/1.0/}	Texture
{http://ns.adobe.com/camera-raw-settings/1.0/}	ToneMapStrength
{http://ns.adobe.com/camera-raw-settings/1.0/}	ConvertToGrayscale
{http://ns.adobe.com/camera-raw-settings/1.0/}	OverrideLookVignette
{http://ns.adobe.com/camera-raw-settings/1.0/}	ToneCurveName
{http://ns.adobe.com/camera-raw-settings/1.0/}	ToneCurveName2012
{http://ns.adobe.com/camera-raw-settings/1.0/}	CameraProfile
{http://ns.adobe.com/camera-raw-settings/1.0/}	CameraProfileDigest
{http://ns.adobe.com/camera-raw-settings/1.0/}	LensProfileSetup
{http://ns.adobe.com/camera-raw-settings/1.0/}	LensProfileName
{http://ns.adobe.com/camera-raw-settings/1.0/}	LensProfileFilename
{http://ns.adobe.com/camera-raw-settings/1.0/}	LensProfileDigest
{http://ns.adobe.com/camera-raw-settings/1.0/}	LensProfileDistortionScale
{http://ns.adobe.com/camera-raw-settings/1.0/}	LensProfileChromaticAberrationScale
{http://ns.adobe.com/camera-raw-settings/1.0/}	LensProfileVignettingScale
{http://ns.adobe.com/camera-raw-settings/1.0/}	UprightVersion
{http://ns.adobe.com/camera-raw-settings/1.0/}	UprightCenterMode
{http://ns.adobe.com/camera-raw-settings/1.0/}	UprightCenterNormX
{http://ns.adobe.com/camera-raw-settings/1.0/}	UprightCenterNormY
{http://ns.adobe.com/camera-raw-settings/1.0/}	UprightFocalMode
{http://ns.adobe.com/camera-raw-settings/1.0/}	UprightFocalLength35mm
{http://ns.adobe.com/camera-raw-settings/1.0/}	UprightPreview
{http://ns.adobe.com/camera-raw-settings/1.0/}	UprightDependentDigest
{http://ns.adobe.com/camera-raw-settings/1.0/}	UprightTransformCount
{http://ns.adobe.com/camera-raw-settings/1.0/}	UprightTransform_0
{http://ns.adobe.com/camera-raw-settings/1.0/}	UprightTransform_1
{http://ns.adobe.com/camera-raw-settings/1.0/}	UprightTransform_2
{http://ns.adobe.com/camera-raw-settings/1.0/}	UprightTransform_3
{http://ns.adobe.com/camera-raw-settings/1.0/}	UprightTransform_4
{http://ns.adobe.com/camera-raw-settings/1.0/}	UprightTransform_5
{http://ns.adobe.com/camera-raw-settings/1.0/}	UprightFourSegmentsCount
{http://ns.adobe.com/camera-raw-settings/1.0/}	ToggleStyleDigest
{http://ns.adobe.com/camera-raw-settings/1.0/}	ToggleStyleAmount
{http://ns.adobe.com/camera-raw-settings/1.0/}	HasSettings
{http://ns.adobe.com/camera-raw-settings/1.0/}	CropTop
{http://ns.adobe.com/camera-raw-settings/1.0/}	CropLeft
{http://ns.adobe.com/camera-raw-settings/1.0/}	CropBottom
{http://ns.adobe.com/camera-raw-settings/1.0/}	CropRight
{http://ns.adobe.com/camera-raw-settings/1.0/}	CropAngle
{http://ns.adobe.com/camera-raw-settings/1.0/}	CropConstrainToWarp
{http://ns.adobe.com/camera-raw-settings/1.0/}	HasCrop
{http://ns.adobe.com/camera-raw-settings/1.0/}	AlreadyApplied
{http://ns.adobe.com/camera-raw-settings/1.0/}	RawFileName
{http://ns.adobe.com/camera-raw-settings/1.0/}	Saturation
{http://ns.adobe.com/camera-raw-settings/1.0/}	Vibrance
{http://ns.adobe.com/camera-raw-settings/1.0/}	HueAdjustmentRed
{http://ns.adobe.com/camera-raw-settings/1.0/}	HueAdjustmentOrange
{http://ns.adobe.com/camera-raw-settings/1.0/}	HueAdjustmentYellow
{http://ns.adobe.com/camera-raw-settings/1.0/}	HueAdjustmentGreen
{http://ns.adobe.com/camera-raw-settings/1.0/}	HueAdjustmentAqua
{http://ns.adobe.com/camera-raw-settings/1.0/}	HueAdjustmentBlue
{http://ns.adobe.com/camera-raw-settings/1.0/}	HueAdjustmentPurple
{http://ns.adobe.com/camera-raw-settings/1.0/}	HueAdjustmentMagenta

{http://ns.adobe.com/camera-raw-settings/1.0/}	SaturationAdjustmentRed
{http://ns.adobe.com/camera-raw-settings/1.0/}	SaturationAdjustmentOrange
{http://ns.adobe.com/camera-raw-settings/1.0/}	SaturationAdjustmentYellow
{http://ns.adobe.com/camera-raw-settings/1.0/}	SaturationAdjustmentGreen
{http://ns.adobe.com/camera-raw-settings/1.0/}	SaturationAdjustmentAqua
{http://ns.adobe.com/camera-raw-settings/1.0/}	SaturationAdjustmentBlue
{http://ns.adobe.com/camera-raw-settings/1.0/}	SaturationAdjustmentPurple
{http://ns.adobe.com/camera-raw-settings/1.0/}	SaturationAdjustmentMagenta
{http://ns.adobe.com/camera-raw-settings/1.0/}	LuminanceAdjustmentRed
{http://ns.adobe.com/camera-raw-settings/1.0/}	LuminanceAdjustmentOrange
{http://ns.adobe.com/camera-raw-settings/1.0/}	LuminanceAdjustmentYellow
{http://ns.adobe.com/camera-raw-settings/1.0/}	LuminanceAdjustmentGreen
{http://ns.adobe.com/camera-raw-settings/1.0/}	LuminanceAdjustmentAqua
{http://ns.adobe.com/camera-raw-settings/1.0/}	LuminanceAdjustmentBlue
{http://ns.adobe.com/camera-raw-settings/1.0/}	LuminanceAdjustmentPurple
{http://ns.adobe.com/camera-raw-settings/1.0/}	LuminanceAdjustmentMagenta
{http://ns.adobe.com/camera-raw-settings/1.0/}	DefringePurpleAmount
{http://ns.adobe.com/camera-raw-settings/1.0/}	DefringePurpleHueLo
{http://ns.adobe.com/camera-raw-settings/1.0/}	DefringePurpleHueHi
{http://ns.adobe.com/camera-raw-settings/1.0/}	DefringeGreenAmount
{http://ns.adobe.com/camera-raw-settings/1.0/}	DefringeGreenHueLo
{http://ns.adobe.com/camera-raw-settings/1.0/}	DefringeGreenHueHi
{http://ns.adobe.com/xap/1.0/}	Label
{http://ns.adobe.com/camera-raw-settings/1.0/}	VignetteMidpoint

Table 2: List of semantic identifiers used to annotate the image-enhancement performed in Adobe Camera-RAW.

Raw camera and pre-processed imaging data in CR3 and TIFF format, respectively, were uploaded to BioStudies using the command line IBM Aspera software tool *ascp* version 3.8.1.161274 to ensure that data has been transmitted without errors. Sparse file check summing was enabled to ensure integrity of files during transfer (parameter *-k 2*). The raw bioimaging data is available under the BioStudies identifier S-BIAD188.

Semantic identifier	**Name**
{http://edamontology.org/data_1060}	File base name
{http://rs.tdwg.org/dwc/terms/acceptedScientificName}	Species name
{http://rs.tdwg.org/dwc/terms/TaxonID}	TaxonID
{http://rs.tdwg.org/dwc/terms/measurementMethod}	Measurement Method
{http://www.openmicroscopy.org/rdf/2016-06/ome_core/Instrument}	Microscope
{http://www.openmicroscopy.org/rdf/2016-06/ome_core/nominalMagnification}	Magnification
{http://www.openmicroscopy.org/rdf/2016-06/ome_core/ContrastMethod}	Contrast
{http://www.openmicroscopy.org/rdf/2016-06/ome_core/Objective}	Microscope Objective
{http://rs.tdwg.org/dwc/terms/basisOfRecord}	Basis of Record
{http://rs.tdwg.org/dwc/terms/PreservedSpecimen}	Voucher specimen
{http://rs.tdwg.org/dwc/terms/EarliestDateCollected}	Collection Date
{http://rs.tdwg.org/dwc/terms/recordedBy}	Collector
{http://rs.tdwg.org/dwc/terms/identifiedBy}	Determined
{http://rs.tdwg.org/dwc/terms/geodeticDatum}	Geodetic datum
{http://www.ebi.ac.uk/efo/EFO_0005020}	Latitude
{http://www.ebi.ac.uk/efo/EFO_0005021}	Longitude
{http://rs.tdwg.org/dwc/terms/verbatimElevation}	Elevation
{http://rs.tdwg.org/dwc/terms/coordinatePrecision}	Precision
{http://purl.obolibrary.org/obo/OBI_0001048}	Camera

Table 3: List of semantic terms used to annotate the segmented images.

Pre-processed images were converted to the Bio-Formats OME-TIFF format[59] by creating intermediate ZARR-pyramid tiles using the bioformats2raw converter version 0.4.0 and then using the raw2ometiff version 0.3.0 software tool to create the final pyramid images. In order to improve data reuse and to enable linking bioimaging data to ecological data repositories, individual fully segmented and processed images were associated with

standardised geolocation information. Swiss Topo CH1903/LV03 coordinates were converted to WGS84 using Swisstopo-WGS84-LV03[60]. The processed images were further associated with the meta-data information listed in Table 3. A helper script was implemented in R to facilitate the generation of TSV tables for data upload to the Image Data Resource (IDR) repository (_tsv_res_2_idr.r in [61]). Processed images and the meta-data aggregated in a TSV table were uploaded to IDR using the software Globus Connect Personal 3.1.6. The dataset is available under the identifier idr0134.

Data Records

Two separate data records were created in order to enable rapid use of the data in machine learning and biodiversity approaches.

(1) The camera raw images (Canon CR3-format), the pre-processed images (16-bit TIFF-format), and the contextual meta-data were deposited in BioStudies under the identifier S-BIAD188 (https://www.ebi.ac.uk/biostudies/studies/S-BIAD188). The data record consists of a total of 223'989 individual raw image files partitioned into 48 species. The entire data record has a total size of approx. 12 TB.

(2) The pre-processed and fully segmented and processed images along with meta-data were deposited in OME-TIFF and JPEG-format, respectively, in the Image Data Resource (IDR) repository under the identifier idr0134 (https://idr.openmicroscopy.org/search/?query=Name:134). The data record consists of a total of 4233 pre-processed and 905 fully processed imaged files. The data record has a total size of approx. 14 TB.

Technical Validation

Biological validation of species identity and visible phenotypic attributes in the pre-processed images were performed by consulting the external experts Edwin Urmi, Heike Hofmann, Vadim Bakalin and Kristian Hassel. Photos of the herbarium specimen CM-30377 originating from North America (Table 1) show quite different properties when compared to the voucher specimen B-108428 originating from Northern Europe. Hence, the taxonomic status of the species *Scapania glaucocephala* is not yet fully clear[15]. Photos of CM-30377 may, thus, relate to the species *Scapania scapanioides* (C.Massal.) Grolle, which is listed in [7] as separate species occurring in Europe. Further, *S. brevicaulis* and *S. degenii* may comprise taxonomically identical species and additional research is needed to resolve their phylogenetic status. Images from this study can help the to clarify relationships of phenotypic attributes and the phylogenetic and taxonomic status of cryptic species.

To validate the technical quality of fused composite images, multi-focused image fusion methods were applied with different settings to the individual images in stacks. Composite images were manually inspected and the best image retained. Generally, classic Laplacian pyramid transform-based methods such as *Pyramid Maximum Contrast* implemented in the software Helicon Focus produce good results in complex cases with regard to intersecting objects, edges and many images, but these algorithms increase contrast and glare and they are prone to noise and artefacts and generally are considered less accurate regarding reproduction of microscopic objects[62–65]. The deterministic depth map-based method implemented in Helicon Focus first calculates depth maps from intermediary images based on the absolute difference in brightness of corresponding pixels in the source images and smoothed intermediary images and then generates the composite image from the source image pixels with indices differing from the indices in the smoothed depth map[66]. Whereas

larger values for the parameter *radius* increase blur along edges, lower values can introduce artefacts, while the amount of blur along the transition between fused areas of individual images can be controlled with the parameter *smoothing*. While the depth map-based method generally produces accurate reproductions of microscopic objects, in some circumstances and especially with high magnifications, it can generate visible and large artefacts and blur around the edges (boundary regions) (Fig. 2). Recently, machine learning-based methods have been applied to focus-based image fusion tasks that may be superior to deterministic approaches[36]. Although there have been proposed some algorithms specifically for microscopic imaging, there is a considerable lack of usable implementations and a lack of microscopic training data for machine learning-based algorithms[36].

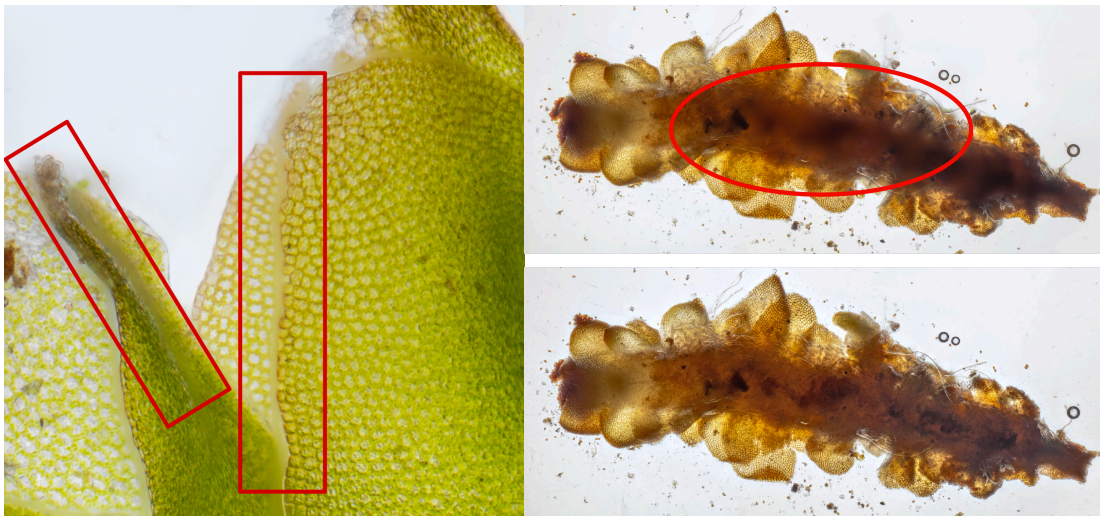

Figure 2: Deficiencies of multi-focus image fusion methods. Red circles and bars were drawn *post-hoc* with Affinity Photo to indicate the deficient regions in the images, thus, the regions where multi-focus image fusion methods can produce blur and artefacts in composite images of microscopic objects. **(a)** Crop of *IMG_1532-1621 Scapania cuspiduligera stature ventral side*. Visible blur along edges (boundary regions) of overlapping leaves (Parameter settings: *Method: Depth Map, Radius: 8, Smoothing: 4*). **(b)** *IMG_0107-0226 Scapania ligulifolia stature dorsal side* (Parameter settings: *Method: Depth Map, Radius: 32, Smoothing: 20*).

In order to facilitate the automatic processing of images, Python scripts have been written which are available as Open Source software in github[57,58]. These scripts use meta-data information to put individual images into image stacks to perform focus-based image fusion and image stitching tasks. Most of the work has been implemented manually. Scientific workflows allow to fully automate the entire process combining images with software tools utilising the machine-actionable information contained in the meta-data[42,48,67]. Meta-data has been validated using procedures described in [20,45,46] and standardised vocabularies were following the FAIR guiding principles[1]. When improved algorithms have been developed, the entire pipeline can be re-run resulting in improved segmented images without any further intervention. This data reuse and the rich documentation in meta-data will result in foster good scientific practices through source tracking and provenance.

Usage Notes

Raw camera and pre-processed imaging data, the fully segmented and processed images, and the meta-data are available under the terms of the license Creative Commons CC BY 4.0.

Open-Source software scripts and code[57,58] are available under the terms of the BSD 3-Clause license.

Code Availability

Software code and scripts used in this study are available as Open Source in github: https://github.com/korseby/create_image_stacks [58], https://github.com/korseby/scale_bar [57], https://github.com/korseby/bioimage_submission [61]. Python scripts were tested under Python 3.7 and require the additional modules PIL, pandas, xml, csv, errno, sys, os, argparse, glob, pathlib and re. R scripts were tested under R 4.1.3 and require the additional packages parallel, foreach, and doMC. Shell scripts were tested using Bourne Again Shell (bash) 5.1.16.

Acknowledgements

KP acknowledges the support of iDiv (funded by the German Research Foundation, DFG-FZT 118, 202548816) and Swissbryophytes. Fully segmented and processed images of 27 *Scapania* species occurring in Switzerland were originally produced for the website http://www.swissbryophytes.ch and are also available at the Zurich Open Repository Archive [68–71,71–93]. We would also like to thank the external experts Edwin Urmi, Heike Hofmann, Vadim Bakalin and Kristian Hassel for providing voucher specimens and for performing validation of species identities on images of visible traits and especially Edwin Urmi for help with identification of fresh samples.

Author contributions

KP performed the entire study. BKR supervised the study.

Competing interests

No competing interests.

References

1. Wilkinson, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* **3**, 160018 (2016).
2. Ellenberg, J. *et al.* A call for public archives for biological image data. *Nat Methods* **15**, 849–854 (2018).
3. Löffler, F., Wesp, V., König-Ries, B. & Klan, F. Dataset search in biodiversity research: Do metadata in data repositories reflect scholarly information needs? *PLoS ONE* **16**, e0246099 (2021).
4. Asakawa, Y., Ludwiczuk, A. & Nagashima, F. Phytochemical and biological studies of bryophytes. *Phytochemistry* **91**, 52–80 (2013).
5. He, X., Sun, Y. & Zhu, R.-L. The Oil Bodies of Liverworts: Unique and Important Organelles in Land Plants. *Critical Reviews in Plant Sciences* **32**, 293–302 (2013).
6. Kanazawa, T. *et al.* The liverwort oil body is formed by redirection of the secretory pathway. *Nat Commun* **11**, 6152 (2020).

7. Hodgetts, N. G. *et al.* An annotated checklist of bryophytes of Europe, Macaronesia and Cyprus. *Journal of Bryology* **42**, 1–116 (2020).
8. Spitale, D. Switch between competition and facilitation within a seasonal scale at colony level in bryophytes. *Oecologia* **160**, 471–482 (2009).
9. Andersen, N. H. *et al.* Sesquiterpenes of nine european liverworts from the genera, Anastrepta, bazzania, jungermannia, lepidozia and Scapania. *Phytochemistry* **16**, 1731–1751 (1977).
10. Guo, L. *et al.* Chemical Composition, Antifungal and Antitumor Properties of Ether Extracts of Scapania verrucosa Heeg. and its Endophytic Fungus Chaetomium fusiforme. *Molecules* **13**, 2114–2125 (2008).
11. Bukvicki, D. R. *et al.* Assessment of the Chemical Composition and *In Vitro* Antimicrobial Potential of Extracts of the Liverwort *Scapania Aspera*. *Natural Product Communications* **8**, 1934578X1300800 (2013).
12. Han, J. *et al.* Terpenoids from Chinese Liverworts *Scapania* spp. *J. Nat. Prod.* **84**, 1210–1215 (2021).
13. Vázquez, M. D., López, J. & Carballeira, A. Uptake of Heavy Metals to the Extracellular and Intracellular Compartments in Three Species of Aquatic Bryophyte. *Ecotoxicology and Environmental Safety* **44**, 12–24 (1999).
14. Samecka-Cymerman, A., Kolon, K. & Kempers, A. J. Heavy Metals in Aquatic Bryophytes from the Ore Mountains (Germany). *Ecotoxicology and Environmental Safety* **52**, 203–210 (2002).
15. Heinrichs, J. *et al.* A phylogeny of the northern temperate leafy liverwort genus Scapania (Scapaniaceae, Jungermanniales). *Molecular Phylogenetics and Evolution* **62**, 973–985 (2012).
16. Vana, J., Hentschel, J., Müller, J. & Heinrichs, J. Taxonomic novelties in Scapania. *PHYTOKEYS* **10**, 13 (2012).
17. Choi, S. S., Min, J., Kwon, W. & Park, J. The complete mitochondrial genome of *Scapania ampliata* Steph., 1897 (Scapaniaceae, Jungermanniales). *Mitochondrial DNA Part B* **6**, 686–688 (2021).
18. Choi, S. S., Bakalin, V. A., Kwon, W. & Park, J. The complete mitochondrial genome of *Douinia plicata* (Lindb.) Konstant. et. Vilnet (Scapaniaceae, Jungermanniales). *Mitochondrial DNA Part B* **6**, 789–791 (2021).
19. Bernhardt-Römermann, M., Poschlod, P. & Hentschel, J. BryForTrait - A life-history trait database of forest bryophytes. *J Veg Sci* **29**, 798–800 (2018).
20. Schneider, F. D. *et al.* Towards an ecological trait-data standard. *Methods Ecol Evol* **10**, 2006–2019 (2019).
21. Kommineni, V. K. *et al.* Comprehensive leaf size traits dataset for seven plant species from digitised herbarium specimen images covering more than two centuries. *BDJ* **9**, e69806 (2021).
22. Meijering, E., Carpenter, A. E., Peng, H., Hamprecht, F. A. & Olivo-Marin, J.-C. Imagining the future of bioimage analysis. *Nat Biotechnol* **34**, 1250–1255 (2016).
23. Cornelissen, J. H. C., Lang, S. I., Soudzilovskaia, N. A. & During, H. J. Comparative Cryptogam Ecology: A Review of Bryophyte and Lichen Traits that Drive Biogeochemistry. *Annals of Botany* **99**, 987–1001 (2007).
24. Díaz, S. *et al.* The global spectrum of plant form and function. *Nature* **529**, 167–171 (2016).
25. Brodribb, T. J., Carriquí, M., Delzon, S., McAdam, S. A. M. & Holbrook, N. M. Advanced vascular function discovered in a widespread moss. *Nat. Plants* **6**, 273–279 (2020).
26. Duckett, J. G. & Pressel, S. Of mosses and vascular plants. *Nat. Plants* **6**, 184–185 (2020).
27. Horn, A. *et al.* Natural Products from Bryophytes: From Basic Biology to Biotechnological Applications. 28.

28. Schindel, D. E. & Cook, J. A. The next generation of natural history collections. *PLoS Biol* **16**, e2006125 (2018).
29. Hedrick, B. P. *et al.* Digitization and the Future of Natural History Collections. *BioScience* **70**, 243–251 (2020).
30. Gurevitch, J., Koricheva, J., Nakagawa, S. & Stewart, G. Meta-analysis and the science of research synthesis. *Nature* **555**, 175–182 (2018).
31. Valdecasas, A. G., Marshall, D., Becerra, J. M. & Terrero, J. J. On the extended depth of focus algorithms for bright ®eld microscopy. 11 (2001).
32. Goldsmith, N. T. DEEP FOCUS; A DIGITAL IMAGE PROCESSING TECHNIQUE TO PRODUCE IMPROVED FOCAL DEPTH IN LIGHT MICROSCOPY. *Image Anal Stereol* **19**, 163 (2011).
33. Nasibov, A., Nasibov, H. & Hacizade, F. Seamless image stitching algorithm using radiometric lens calibration for high resolution optical microscopy. in *2009 Fifth International Conference on Soft Computing, Computing with Words and Perceptions in System Analysis, Decision and Control* 1–4 (IEEE, 2009). doi:10.1109/ICSCCW.2009.5379500.
34. Wang, Z. & Yang, Z. Review on image-stitching techniques. *Multimedia Systems* **26**, 413–430 (2020).
35. Piper, J. Software-Based Stacking Techniques to Enhance Depth of Field and Dynamic Range in Digital Photomicrography. in *Histology Protocols* (eds. Hewitson, T. D. & Darby, I. A.) vol. 611 193–210 (Humana Press, 2010).
36. Liu, Y., Wang, L., Cheng, J., Li, C. & Chen, X. Multi-focus image fusion: A Survey of the state of the art. *Information Fusion* **64**, 71–91 (2020).
37. Jianchao Yang, Wright, J., Huang, T. S., & Yi Ma. Image Super-Resolution Via Sparse Representation. *IEEE Trans. on Image Process.* **19**, 2861–2873 (2010).
38. Yin, H., Li, S. & Fang, L. Simultaneous image fusion and super-resolution using sparse representation. *Information Fusion* **14**, 229–240 (2013).
39. Yu, Z., Liu, S., Zhu, D., Kuang, C. & Liu, X. Parallel detecting super-resolution microscopy using correlation based image restoration. *Optics Communications* **404**, 139–146 (2017).
40. Yang, B., Zhong, J., Li, Y. & Chen, Z. Multi-focus image fusion and super-resolution with convolutional neural network. *Int. J. Wavelets Multiresolut Inf. Process.* **15**, 1750037 (2017).
41. Peters, K. *et al.* PhenoMeNal: processing and analysis of metabolomics data in the cloud. *GigaScience* **8**, (2019).
42. Goble, C. *et al.* FAIR Computational Workflows. *Data Intellegence* **2**, 108–121 (2020).
43. Atkinson, M., Gesing, S., Montagnat, J. & Taylor, I. Scientific workflows: Past, present and future. *Future Generation Computer Systems* **75**, 216–227 (2017).
44. Miksa, T., Simms, S., Mietchen, D. & Jones, S. Ten principles for machine-actionable data management plans. *PLoS Comput Biol* **15**, e1006750 (2019).
45. Samuel, S., Taubert, F., Walther, D., König-Ries, B. & Bücker, H. M. Towards Reproducibility of Microscopy Experiments. *D-Lib Magazine* **23**, (2017).
46. Kunis, S. *et al.* MDEmic: a metadata annotation tool to facilitate management of FAIR image data in the bioimaging community. *Nat Methods* (2021) doi:10.1038/s41592-021-01288-z.
47. Samuel, S. & König-Ries, B. End-to-End provenance representation for the understandability and reproducibility of scientific experiments using a semantic approach. *J Biomed Semant* **13**, 1 (2022).
48. Wratten, L., Wilm, A. & Göke, J. Reproducible, scalable, and shareable analysis pipelines with bioinformatics workflow managers. *Nat Methods* **18**, 1161–1168 (2021).
49. Hansen, O. L. P. *et al.* Species-level image classification with convolutional neural network enables insect identification from habitus images. *Ecol Evol* **10**, 737–747 (2020).
50. Høye, T. T. *et al.* Deep learning and computer vision will transform entomology. *Proc Natl Acad Sci USA* **118**, e2002545117 (2021).

51. König, C. *et al.* Biodiversity data integration—the significance of data resolution and domain. *PLoS Biol* **17**, e3000183 (2019).
52. Nelson, G. & Paul, D. L. DiSSCo, iDigBio and the Future of Global Collaboration. *BISS* **3**, e37896 (2019).
53. Culina, A. *et al.* Navigating the unfolding open data landscape in ecology and evolution. *Nat Ecol Evol* **2**, 420–426 (2018).
54. Seltzer, C. Making Biodiversity Data Social, Shareable, and Scalable: Reflections on iNaturalist & citizen science. *BISS* **3**, e46670 (2019).
55. Borgman, C. L. & Bourne, P. E. Why it takes a village to manage and share data. 20.
56. Holmgren, P. K. & Holmgren, N. H. INDEX HERBARIORUM. *TAXON* **40**, 687–692 (1991).
57. Peters, K. *korseby/scale_bar: v0.2*. (Zenodo, 2021). doi:10.5281/ZENODO.5592446.
58. Peters, K. *korseby/create_image_stacks: v0.3*. (Zenodo, 2021). doi:10.5281/ZENODO.5592436.
59. Besson, S. *et al.* Bringing Open Data to Whole Slide Imaging. in *Digital Pathology* (eds. Reyes-Aldasoro, C. C., Janowczyk, A., Veta, M., Bankhead, P. & Sirinukunwattana, K.) vol. 11435 3–10 (Springer International Publishing, 2019).
60. Marti, U. & Dupraz, H. *Swisstopo Scripts GPS WGS84 <-> LV03 (CH1903)*. (2021).
61. Peters. *korseby/bioimage_submission: v0.1.0*. (Zenodo, 2022). doi:10.5281/ZENODO.6447017.
62. Adelson, E. H., Anderson, C. H., Bergen, J. R., Burt, P. J. & Ogden, J. M. Pyramid methods in image processing. 9 (1984).
63. Ogden, J. M., Adelson, E. H., Bergen, J. R. & Burt, P. J. Pyramid-based computer graphics. 12 (1985).
64. Toet, A. Image fusion by a ratio of low-pass pyramid. *Pattern Recognition Letters* **9**, 245–253 (1989).
65. Liu, Z., Tsukada, K., Hanasaki, K., Ho, Y. K. & Dai, Y. P. Image fusion by using steerable pyramid. *Pattern Recognition Letters* **22**, 929–939 (2001).
66. Kozub, D. Focus stacking of captured images. 11.
67. Perkel, J. M. Computational pipelines turn raw data into reproducible scientific knowledge. 2.
68. Urmi, E., Hofmann, H. & Schubiger, C. Scapania aspera Bernet & M.Bernet. (2020) doi:10.5167/UZH-197490.
69. Urmi, E., Hofmann, H. & Schubiger, C. Scapania subalpina (Lindenb.) Dumort. (2020) doi:10.5167/UZH-197517.
70. Urmi, E., Hofmann, H. & Schubiger, C. Scapania undulata (L.) Dumort. (2020) doi:10.5167/UZH-197522.
71. Urmi, E., Peters, K. & Schubiger, C. Scapania nemorea subsp. nemorea (L.) Grolle. (2020) doi:10.5167/UZH-205668.
72. Urmi, E., Peters, K. & Schubiger, C. Scapania aequiloba (Schwägr.) Dumort. (2020) doi:10.5167/UZH-197488.
73. Urmi, E., Peters, K. & Schubiger, C. Scapania apiculata Spruce. (2020) doi:10.5167/UZH-197489.
74. Urmi, E., Peters, K. & Schubiger, C. Scapania calcicola (Arnell & J.Perss.) Ingham. (2020) doi:10.5167/UZH-197492.
75. Urmi, E., Peters, K. & Schubiger, C. Scapania carinthiaca Lindb. (2020) doi:10.5167/UZH-197494.
76. Urmi, E., Peters, K. & Schubiger, C. Scapania compacta (Roth) Dumort. (2020) doi:10.5167/UZH-197496.
77. Urmi, E., Peters, K. & Schubiger, C. Scapania curta (Mart.) Dumort. (2020) doi:10.5167/UZH-197497.
78. Urmi, E., Peters, K. & Schubiger, C. Scapania cuspiduligera (Nees) Müll.Frib. (2020) doi:10.5167/UZH-197499.

79. Urmi, E., Peters, K. & Schubiger, C. Scapania degenii Müll.Frib. (2020) doi:10.5167/UZH-197500.
80. Urmi, E., Peters, K. & Schubiger, C. Scapania gracilis Lindb. (2020) doi:10.5167/UZH-197503.
81. Urmi, E., Peters, K. & Schubiger, C. Scapania gymnostomophila Kaal. (2020) doi:10.5167/UZH-197504.
82. Urmi, E., Peters, K. & Schubiger, C. Scapania helvetica Gottsche. (2020) doi:10.5167/UZH-197505.
83. Urmi, E., Peters, K. & Schubiger, C. Scapania irrigua subsp. irrigua (Nees) Nees. (2020) doi:10.5167/UZH-197506.
84. Urmi, E., Peters, K. & Schubiger, C. Scapania irrigua subsp. rufescens (Loeske) R.M.Schust. (2020) doi:10.5167/UZH-197507.
85. Urmi, E., Peters, K. & Schubiger, C. Scapania mucronata subsp. mucronata H.Buch. (2020) doi:10.5167/UZH-197508.
86. Urmi, E., Peters, K. & Schubiger, C. Scapania mucronata subsp. praetervisa (Meyl.) R.M.Schust. (2020) doi:10.5167/UZH-197509.
87. Urmi, E., Peters, K. & Schubiger, C. Scapania obscura (Arnell & C.E.O.Jensen) Schiffn. (2020) doi:10.5167/UZH-197511.
88. Urmi, E., Peters, K. & Schubiger, C. Scapania paludicola Loeske & Müll.Frib. (2020) doi:10.5167/UZH-197513.
89. Urmi, E., Peters, K. & Schubiger, C. Scapania paludosa (Müll.Frib.) Müll.Frib. (2020) doi:10.5167/UZH-197514.
90. Urmi, E., Peters, K. & Schubiger, C. Scapania scandica (Arnell & H.Buch) Macvicar. (2020) doi:10.5167/UZH-197515.
91. Urmi, E., Peters, K. & Schubiger, C. Scapania uliginosa (Lindenb.) Dumort. (2020) doi:10.5167/UZH-197520.
92. Urmi, E., Peters, K. & Schubiger, C. Scapania umbrosa (Schrad.) Dumort. (2020) doi:10.5167/UZH-197521.
93. Urmi, E., Peters, K. & Schubiger, C. Scapania verrucosa Heeg. (2020) doi:10.5167/UZH-197523.