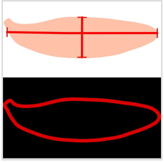


Mar 11, 2021

Extracting shape and size information from fungal spores

Alexander Ordynets[1], Sarah Keßler[1], Christina Willemsens[1]

[1]Universität Kassel

1 Works for me dx.doi.org/10.17504/protocols.io.bdeii3ce

Alexander Ordynets
Universität Kassel

SUBMIT TO PLOS ONE

ABSTRACT

The protocol explains extracting and quantifying the shape and size information from the fungal spores. The complex information about the outline will be transformed into few values that can be used as shape descriptors in comparative analyses.
The photos of spores from the light microscopy serve as input data. The shape information is processed with the elliptic Fourier analysis and afterward with the principal component analysis (PCA). The size information is captured as linear measurements of length and width. The average per specimen values for shape and size traits are then calculated in R computing environment.
The protocol can be also applied to the images from other kinds of microscopy or to images of macroscopic objects (plants, animals, or their organs). The protocol contains example files and screenshots that should assist in the processing of own data. See "Materials" for the list of used software and "Guidelines and Warnings" for the flowchart illustrating the protocol.

THIS PROTOCOL ACCOMPANIES THE FOLLOWING PUBLICATION

Ordynets A, Keßler S, Langer E. Quantitative analysis of spore shapes improves identification of fungi (in preparation)

DOI

dx.doi.org/10.17504/protocols.io.bdeii3ce

PROTOCOL CITATION

Alexander Ordynets, Sarah Keßler, Christina Willemsens 2021. Extracting shape and size information from fungal spores. **protocols.io**
https://dx.doi.org/10.17504/protocols.io.bdeii3ce

MANUSCRIPT CITATION please remember to cite the following publication along with this protocol

Ordynets A, Keßler S, Langer E. Quantitative analysis of spore shapes improves identification of fungi (in preparation)

KEYWORDS

image processing, geometric morphometrics, morphology, trait, outline-based morphometrics, linear measurement, image analysis, shape analysis, outline analysis, elliptic Fourier analysis

LICENSE

IMAGE ATTRIBUTION

Attribution 4.0 International (CC BY 4.0) - https://creativecommons.org/licenses/by/4.0/

CREATED

Mar 09, 2020

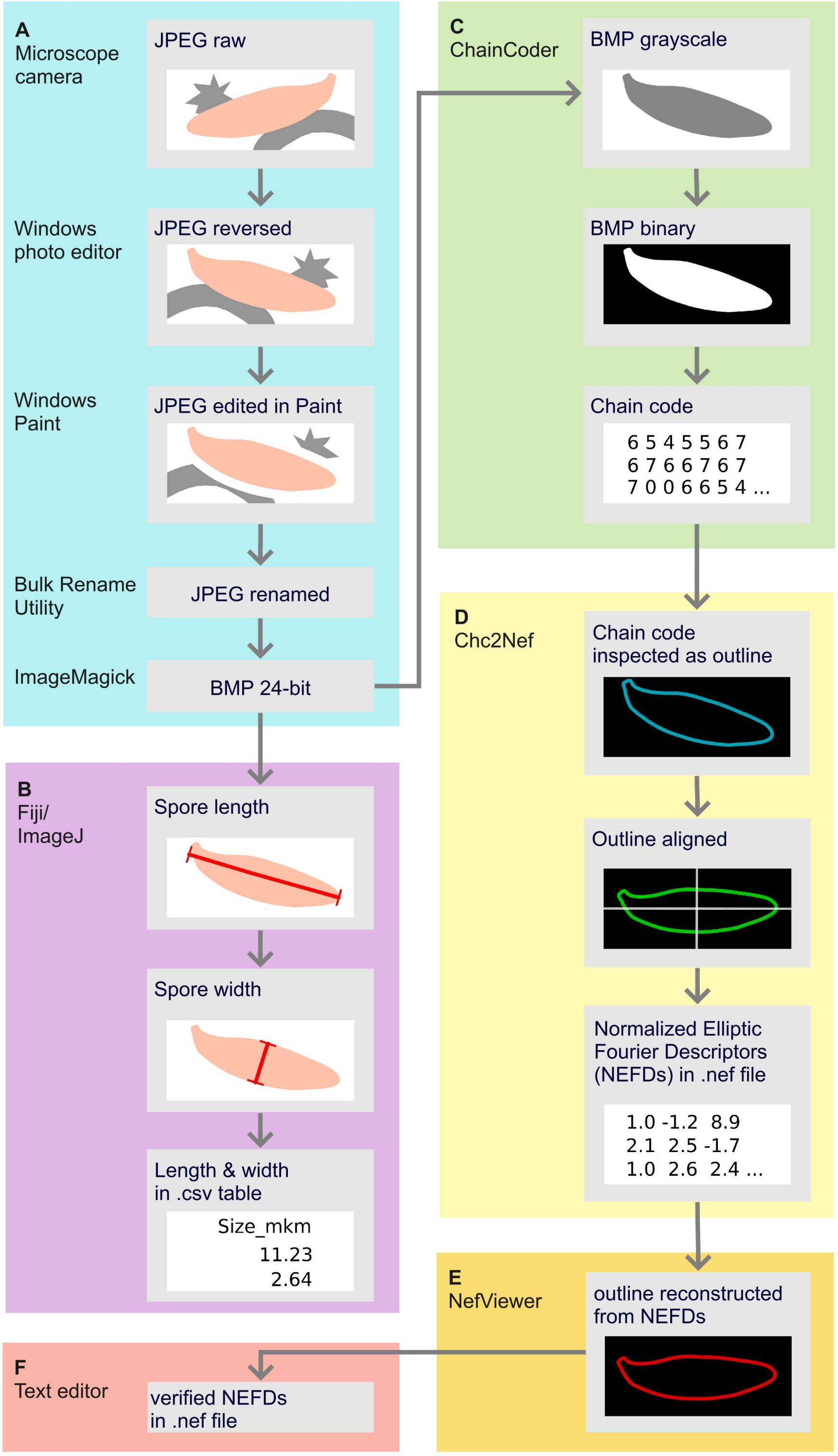
LAST MODIFIED

Mar 11, 2021

PROTOCOL INTEGER ID

33962

Citation: Alexander Ordynets, Sarah KeÃÂler, Christina Willemsens (03/11/2021). Extracting shape and size information from fungal spores. https://dx.doi.org/10.17504/protocols.io.bdeii3ce



Flowchart illustrating the current protocol: extraction of shape and size information from a single fungal spore. Created with Inkscape (https://inkscape.org/) and Miro (https://miro.com/).

Citation: Alexander Ordynets, Sarah KeÃÂler, Christina Willemsens (03/11/2021). Extracting shape and size information from fungal spores. https://dx.doi.org/10.17504/protocols.io.bdeii3ce

MATERIALS TEXT

The current protocol was developed on a computer with Windows 10 operating system but should be also applicable for other operating systems.

For imaging, our lab uses Leica DM500 light microscope and images are captured (as JPEG files) with a built-in ICC 50 HD Camera using Leica Application Suite EZ software versions 3.2.1 and 3.4.2 (Leica Microsystems GmbH, Wetzlar, Germany). The size of our captured images was 1024 × 768 or 2048 ×1536 pixels while the resolution was always 96 dpi.

The protocol requires the following software installed (the version we used and citations provided)

Bulk Rename Utility v. 3.3.1.0
TGRMN Software Company (2019) Bulk Rename Utility. TGRMN Software Company. Available from: https://www.bulkrenameutility.co.uk/.

Fiji 2.1.0/1.53c that uses Java 1.8.0_172 [64-bit]
Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez J-Y, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A (2012) Fiji: an open-source platform for biological-image analysis. Nature Methods 9: 676–682. https://doi.org/10.1038/nmeth.2019

ImageMagick v. 7.0.10
The ImageMagick Development Team (2020) ImageMagick. Available from: https://imagemagick.org.

R v 4.0.3; v. 3.6.3
R Core Team (2020) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Available from: https://www.r-project.org/.

RStudio v. 1.3.1093
RStudio Team (2020) RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA. Available from: http://www.rstudio.com/.

SHAPE package v. 1.3: including programs ChainCoder, Ch2Nef, NefViewer, PrinComp and PrinPrint.
Iwata H, Ukai Y (2002) SHAPE: A Computer Program Package for Quantitative Evaluation of Biological Shapes Based on Elliptic Fourier Descriptors. Journal of Heredity 93: 384–385. https://doi.org/10.1093/jhered/93.5.384
http://lbm.ab.a.u-tokyo.ac.jp/~iwata/shape/ [the detailed manual from SHAPE authors is also here!]

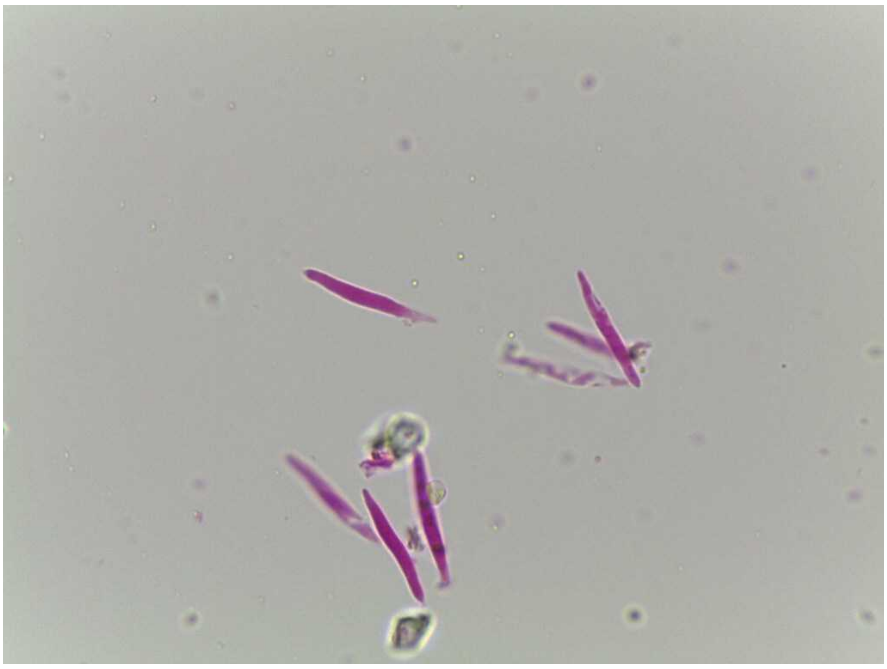
Text editor
e.g. Notepad++ https://notepad-plus-plus.org/

Windows PowerShell or command prompt for Windows machines. Other operating systems would need their shell tools like e.g. Bash.

Image capture

1 Prepare a microscope slide. Place small pieces of fungal dried herbarium specimen in a drop of 3% aqueous solution of potassium hydroxide (KOH) mixed with 1% aqueous solution of Phloxine. The volume proportions of KOH vs. Phloxine in a mixed drop should be approximately 3:1. Examine squash preparations at 1000× magnification and find the places, if possible, with numerous spores in a view field. Capture the image, e.g. as JPEG file.

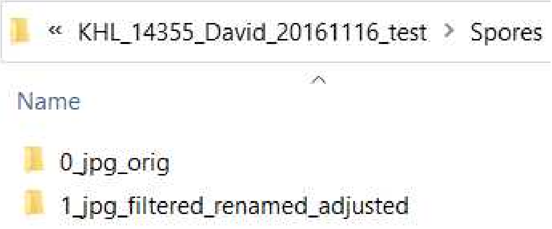
Citation: Alexander Ordynets, Sarah KeÃÂler, Christina Willemsens (03/11/2021). Extracting shape and size information from fungal spores. https://dx.doi.org/10.17504/protocols.io.bdeii3ce



Example of raw JPEG image as captured by microscope camera Leica ICC 50 HD

Image pre-processing

2 Place raw JPEG files with images from a single specimen into one folder.
Create a copy of this folder and further work only with the copy. In this way, the original JPEGs will stay unchanged and serve as a backup.



« KHL_14355_David_20161116_test › Spores

Name

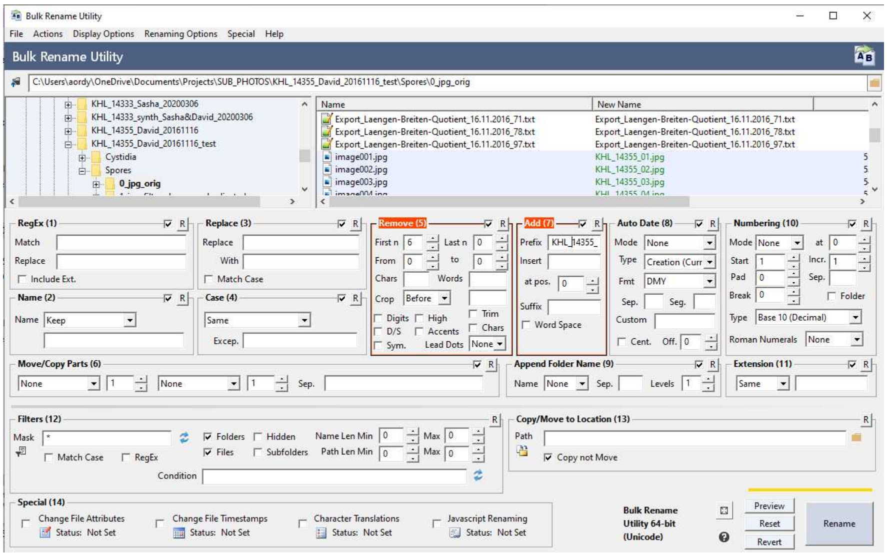
0_jpg_orig

1_jpg_filtered_renamed_adjusted

Folder for one specimen containing two folders: one for raw (0_) and one for processed JPEG images (1_)

3 For each of the captured images, the default file name should be modified to include the specimen name. This is necessary for further correct summarizing of information at the specimen level.
Make bulk photo file renaming with Bulk Rename Utility

Citation: Alexander Ordynets, Sarah KeÃÂler, Christina Willemsens (03/11/2021). Extracting shape and size information from fungal spores. https://dx.doi.org/10.17504/protocols.io.bdeii3ce



Bulk Rename Utility

C:\Users\aordy\OneDrive\Documents\Projects\SUB_PHOTOS\KHL_14355_David_20161116_test\Spores\0_jpg_orig

- KHL_14333_Sasha_20200306
- KHL_14333_synth_Sasha&David_20200306
- KHL_14355_David_20161116
- KHL_14355_David_20161116_test
  - Cystidia
  - Spores
    - 0_jpg_orig

Name	New Name
Export_Laengen-Breiten-Quotient_16.11.2016_71.txt	Export_Laengen-Breiten-Quotient_16.11.2016_71.txt
Export_Laengen-Breiten-Quotient_16.11.2016_78.txt	Export_Laengen-Breiten-Quotient_16.11.2016_78.txt
Export_Laengen-Breiten-Quotient_16.11.2016_97.txt	Export_Laengen-Breiten-Quotient_16.11.2016_97.txt
image001.jpg	KHL_14355_01.jpg
image002.jpg	KHL_14355_02.jpg
image003.jpg	KHL_14355_03.jpg

Remove (5) ☑: First n 6, Last n 0, From 0 to 0, Crop Before, Lead Dots None

Add (7) ☑: Prefix KHL_14355_, at pos. 0

Name (2) ☑: Name Keep; Case (4) ☑: Same; Auto Date (8): Mode None, Type Creation (Curr, Fmt DMY; Numbering (10): Mode None, Start 1, Incr. 1, Pad 0, Break 0, Type Base 10 (Decimal), Roman Numerals None; Extension (11): Same; Filters (12): Mask *, ☑ Folders, ☑ Files; Copy/Move to Location (13): ☑ Copy not Move

Screenshot showing how to make bulk name changes for the images with Bulk Rename Utility. In this case, the pattern "image0" is removed and the pattern "KHL_14355_" is added to all image names.

A	B
Raw file name	Edited file name
image001.jpg	KHL_14355_01.jpg

Example of a raw (left) and edited (right) file name at this processing step

4 For further processing, select images that contain one or several healthy (not broken, with intact cell wall) mature (not attached to basidia) and well-focused spores. Focus only on spores that lay clearly in lateral projection, i.e. so that hilar appendix (approximate locus of attachment to basidia) is directed to the side as far as possible from the longest axis of the spore. Furthermore, for meaningful shape analysis, spores have to be alignable, i.e. with hilar appendix stretched to the same direction if the spores are aligned. We agreed that the same direction means the upper left corner if the spore is placed strictly horizontally and with the base at the left and apex to the right. Some images have to be mirrored (reversed along the horizontal axis) to meet this criterion. This can be done with a built-in photo editor of Windows 10. It can be that some spores are usable from the image before reversing, and some - after. In such cases, both image versions are kept for analysis but have to be marked correspondingly (see the next protocol step for details of naming the images).

Additional manual adjustments can be applied to some images to enhance the clarity of spore outlines and prevent the overlap with the structures that are not spores (crystals or hyphae). These adjustments are made with the Paint 3D or classical Paint tool of Windows 10. These adjustments do not affect the geometric properties of the spore outlines but make spore contour extraction more efficient.

Citation: Alexander Ordynets, Sarah KeÃÂler, Christina Willemsens (03/11/2021). Extracting shape and size information from fungal spores. https://dx.doi.org/10.17504/protocols.io.bdeii3ce

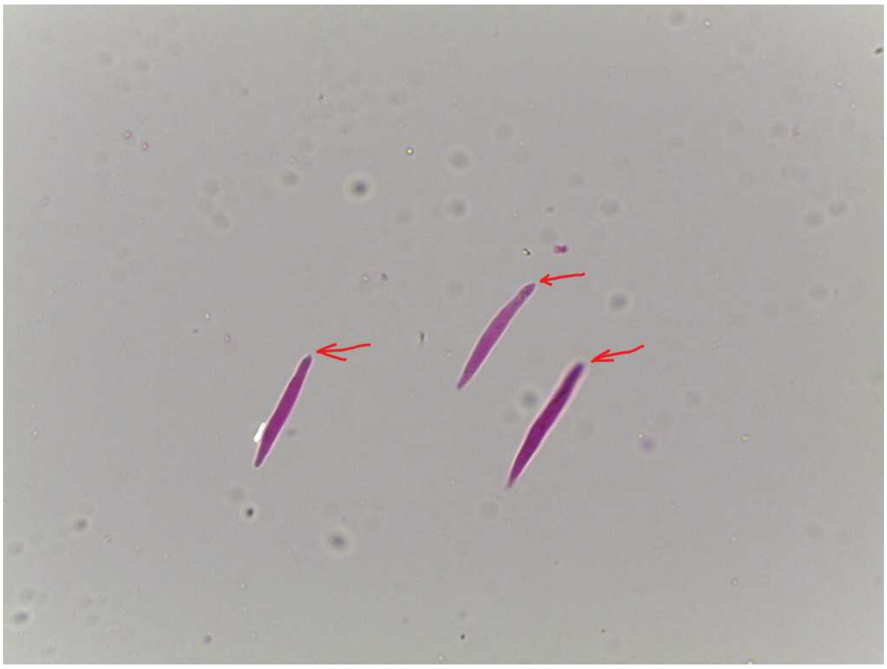
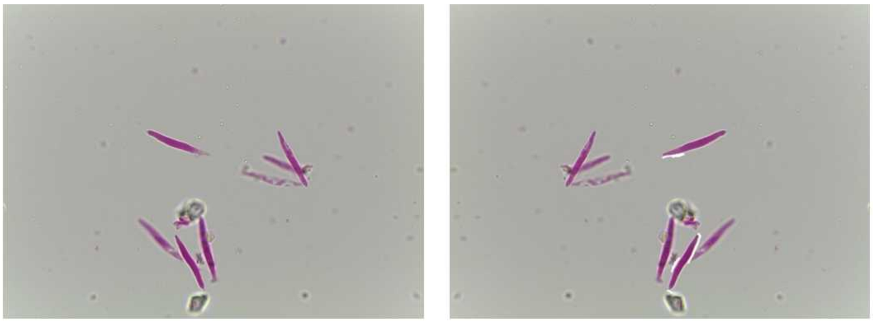


Image from which three spore outlines will be extracted. No image mirroring is required because if to imaginably rotate the spores so that their longest dimension is strictly horizontal, the hilar appendix (see arrow) will be directed to the upper left corner of the image. However, little manual adjustment of the background near the leftmost spore was required. The spore overlapped with a crystal and the latter was masked with a white line in Windows Paint 3D.



Original (left) and mirrored (right) versions of one image. Mirroring allows extracting the outlines of two spores. However, again, graphical artifacts have to be removed by drawing manually the white lines around parts of these spores in Windows Paint 3D.

5 Careful naming of the images allows keeping track of the target spores and distinguishing them from the rest of the information in the image. Below is the general pattern for naming the edited JPEGs and an example of a whole set of images for one specimen.

A
VoucherID_ImageNr_ rev[if made]_ NrOfSporesxLocation_paint[if made]

An overall pattern of the name for the processed image, with different elements separated by underscores.

Citation: Alexander Ordynets, Sarah KeÃÂler, Christina Willemsens (03/11/2021). Extracting shape and size information from fungal spores. https://dx.doi.org/10.17504/protocols.io.bdeii3ce

A	B
KHL_14355_01_1xT_paint.jpg	image #1 with one "good" spore at the top, Paint 3D tool was applied
KHL_14355_02_2xC_paint.jpg	image #2 with two spores at the center, Paint 3D was applied
KHL_14355_06_1xC_paint.jpg	image #6 with one spore at the centre, Paint 3D was applied
KHL_14355_06_rev_1xL.jpg	image #6, reversed version, with one spore in the left part
KHL_14355_07_3x_paint.jpg	image #7 with three spores, no reversing or adjusting with Paint 3D made
KHL_14355_08_1xTR_paint.jpg	image #8 at the top-right, Paint 3D was applied
KHL_14355_13_rev_2x_R_B.jpg	image #13, usable after reversing, with two spores one of which at the right and one at the bottom part of the image
KHL_14355_14_4x.jpg	image #14 with four spores, usable "as is"
KHL_14355_21_1xCL_paint.jpg	image #21 with one spore at the center-left, Paint tool was applied
KHL_14355_21_rev_1xBR_paint.jpg	image #21 after reversing, with one spore at the bottom-right, Paint 3D was applied

Explanations of the image names to understand the naming system

Ten edited images mentioned just above can be accessed from the attached archive.

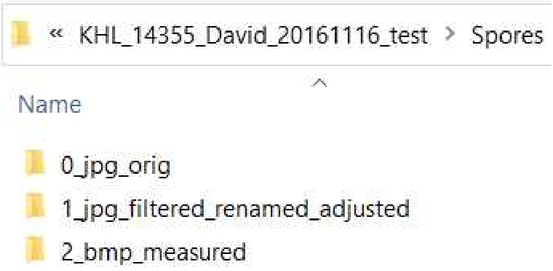
1_jpg_filtered_renamed_adjusted.zip The pattern

6 Create a third folder and move there the ready JPEGs (i.e. that are edited and re-named). Bulk-convert the JPEGs into BMP files with ImageMagick.
For this, open Windows PowerShell in this folder and type:

```
magick mogrify -format bmp *.jpg
```

After BMP files were created, remove the JPEGs from this folder.

7 For each specimen, keep the uniform folder structure. Keep the original JPEGs and processed JPEGs and BMPs in separate folders.


« KHL_14355_David_20161116_test › Spores

Name

0_jpg_orig

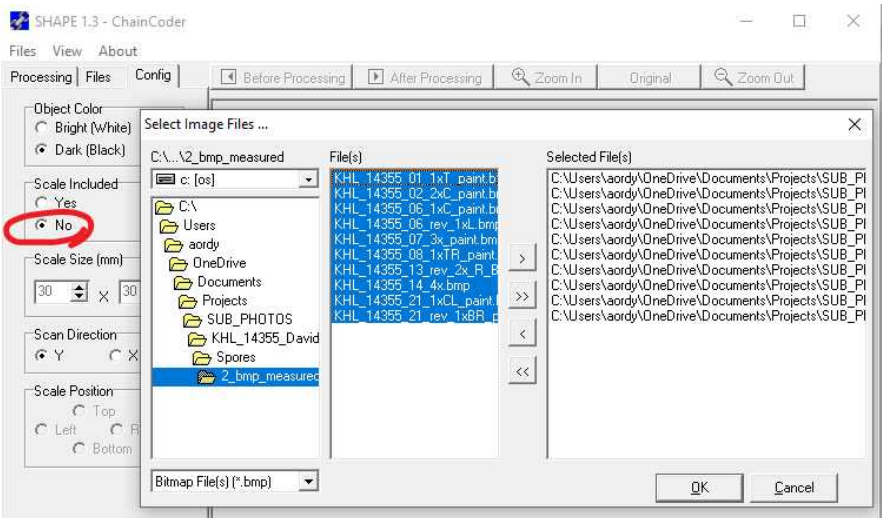
1_jpg_filtered_renamed_adjusted

2_bmp_measured

SHAPE: ChainCoder

8 Open ChainCoder program of the SHAPE package.
When configuring ChainCoder, for proper capturing of objects from the very first one, specify that the scale is missing on the pictures. Otherwise, the first spore will be considered as a scale.

Citation: Alexander Ordynets, Sarah KeÃÂler, Christina Willemsens (03/11/2021). Extracting shape and size information from fungal spores. https://dx.doi.org/10.17504/protocols.io.bdeii3ce



Object Color: ○ Bright (White) ● Dark (Black)

Scale Included: ○ Yes ● No

Scale Size (mm): 30 x 30

Scan Direction: ● Y ○ X

Scale Position: ○ Top ○ Left ○ R ○ Bottom

Select Image Files ...

C:\...\2_bmp_measured

c: [os]

```
C:\
  Users
    aordy
      OneDrive
        Documents
          Projects
            SUB_PHOTOS
              KHL_14355_David
                Spores
                  2_bmp_measured
```

Bitmap File(s) (*.bmp)

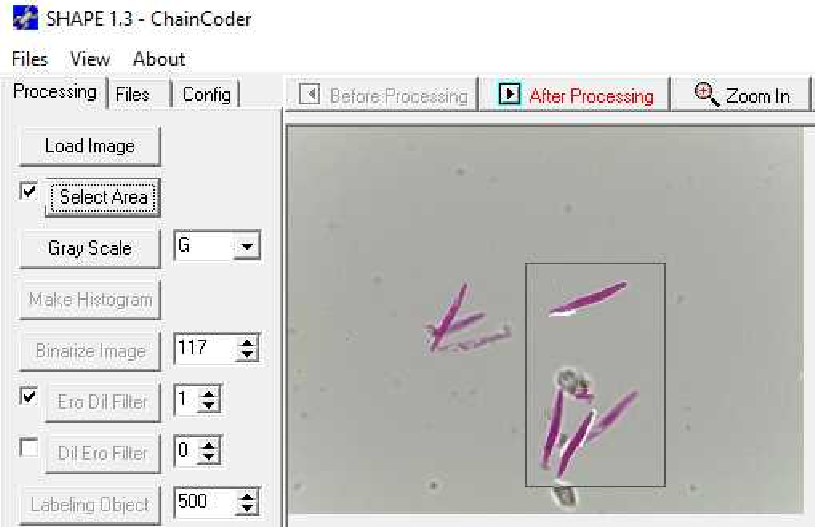
File(s)

```
KHL_14355_01_1xT_paint.b
KHL_14355_02_2xC_paint.b
KHL_14355_06_1xC_paint.b
KHL_14355_06_rev_1xL.bmp
KHL_14355_07_3x_paint.bm
KHL_14355_08_1xTR_paint.
KHL_14355_13_rev_2x_R_B
KHL_14355_14_4x.bmp
KHL_14355_21_1xCL_paint.
KHL_14355_21_rev_1xBR_p
```

Selected File(s)

```
C:\Users\aordy\OneDrive\Documents\Projects\SUB_PI
C:\Users\aordy\OneDrive\Documents\Projects\SUB_PI
C:\Users\aordy\OneDrive\Documents\Projects\SUB_PI
C:\Users\aordy\OneDrive\Documents\Projects\SUB_PI
C:\Users\aordy\OneDrive\Documents\Projects\SUB_PI
C:\Users\aordy\OneDrive\Documents\Projects\SUB_PI
C:\Users\aordy\OneDrive\Documents\Projects\SUB_PI
C:\Users\aordy\OneDrive\Documents\Projects\SUB_PI
C:\Users\aordy\OneDrive\Documents\Projects\SUB_PI
C:\Users\aordy\OneDrive\Documents\Projects\SUB_PI
```

9 Select files containing all pictures for one specimen. They are BMPs located in their own folder, according to the instructions above.
Then proceed the furtherpictures one-by-one.
Click "Load Image".
Click 'Select Area' to crop the image and leave just the area of interest.



☑ Select Area

Gray Scale: G

Binarize Image: 117

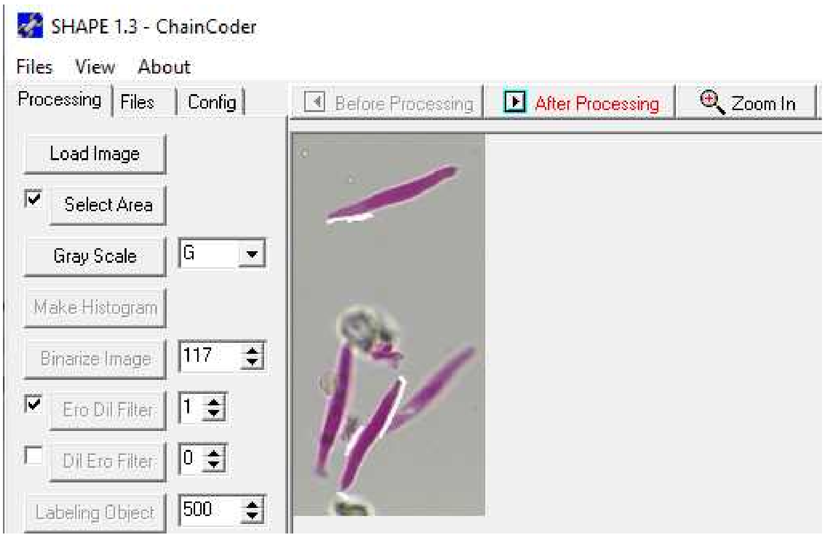
☑ Ero Dil Filter: 1

☐ Dil Ero Filter: 0

Labeling Object: 500

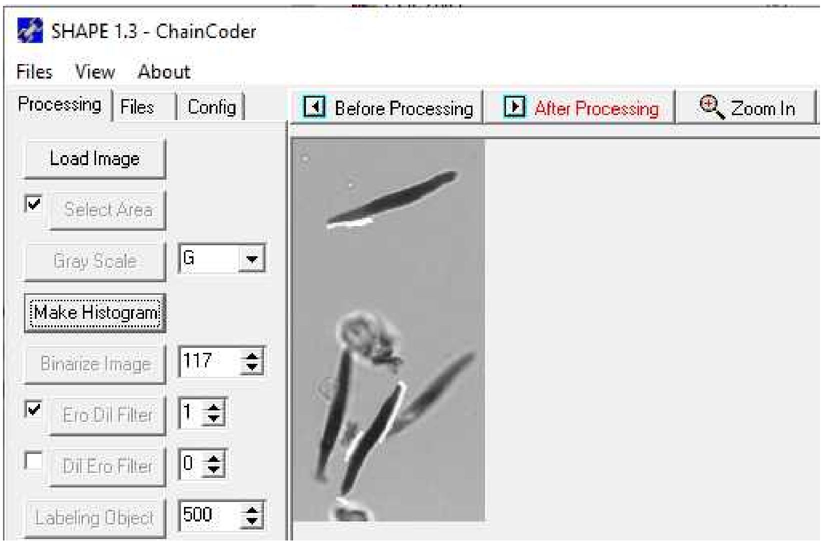
The full-size image where the fragment of interest was selected

Citation: Alexander Ordynets, Sarah KeÃÂler, Christina Willemsens (03/11/2021). Extracting shape and size information from fungal spores. https://dx.doi.org/10.17504/protocols.io.bdeii3ce

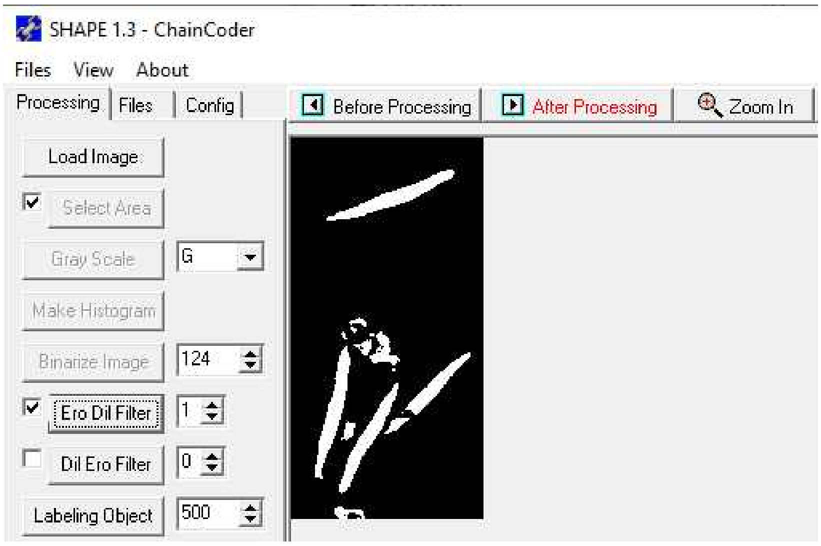


Cropped (just internally in ChainCoder) fragment of interest that will further be processed

10 Turn the image to grayscale by using the green channel. It is a "G" option right to the "Gray Scale" button. Green works better than other channels for the spores that are colored in pink hues with Phloxine.



11 Click 'Make Histogram' which will propose the threshold for identifying the pixels as belonging to the object vs. belonging to the background. The threshold that ChainCoder proposes itself usually works well. For the pictures with lower contrast, selecting the value slightly smaller than proposed by ChainCoder can help. A binary (black and white) image will be created.



☑ Select Area

Gray Scale: G

Binarize Image: 124

☑ Ero Dil Filter: 1

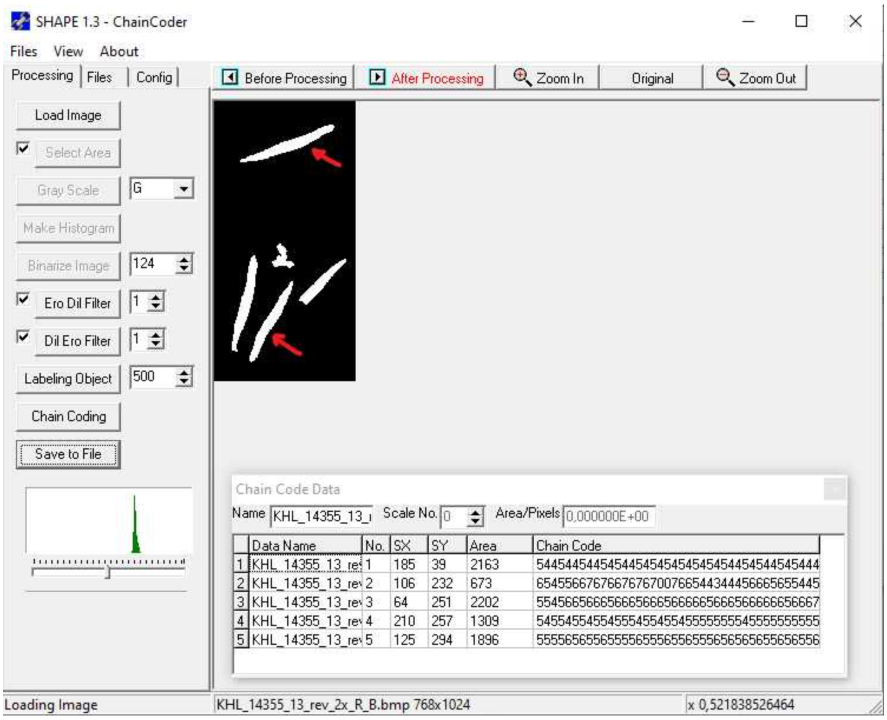
□ Dil Ero Filter: 0

Labeling Object: 500

12 To remove the shape artefacts like protrusions, apply erosion-dilation filter, the option "Ero Dil Filter", with number iterations set to one. To remove shape artefacts like cavities, if such are present, use the dilation-erosion filter, "Dil Ero Filter", with iteration specified arbitrarily as one or two, very rarely three to four per image. In most cases, "Dil Ero Filter" will not be necessary.

13 Click the button "Labeling Object" to select the objects that are larger than the specified number of pixels.
The default 500 works for most cases but for specimens with small spores (<7 mkm long) set the value of 200 or 100.
Click "Chain Coding" to get the chain codes for the spores. Chain code is a sequence of x and y coordinates describing the outline of an object.
Write the chain codes from the given image to a file ("Save to File") and proceed with the next image of this specimen.
When ready with all images for one specimen, close ChainCoder.

Citation: Alexander Ordynets, Sarah KeÃÂler, Christina Willemsens (03/11/2021). Extracting shape and size information from fungal spores. https://dx.doi.org/10.17504/protocols.io.bdeii3ce



☑ Select Area

Gray Scale: G

Binarize Image: 124

☑ Ero Dil Filter: 1

☑ Dil Ero Filter: 1

Labeling Object: 500

Name: KHL_14355_13_ Scale No. 0 Area/Pixels 0,000000E+00

	Data Name	No.	SX	SY	Area	Chain Code
1	KHL_14355_13_re	1	185	39	2163	5445445445454454545454545454454544545444
2	KHL_14355_13_re	2	106	232	673	6545566767667676700766544344456665655445
3	KHL_14355_13_re	3	64	251	2202	5545665666566656665666665666566666656667
4	KHL_14355_13_re	4	210	257	1309	5455455455455545545545555555545555555555
5	KHL_14355_13_re	5	125	294	1896	5555656556555565556556555656565655656556

Loading Image KHL_14355_13_rev_2x_R_B.bmp 768x1024 x 0,521838526464

Image after all filtering steps and taking the chain codes. One can see that five objects were left, out of which only two are objects of interest while three others are either spore with an artifact or not spore. It is not possible here to remove the not target objects but will be possible in Chc2Nef software (see steps below).

The ready .chc file with the chain codes for one specimen is attached below.

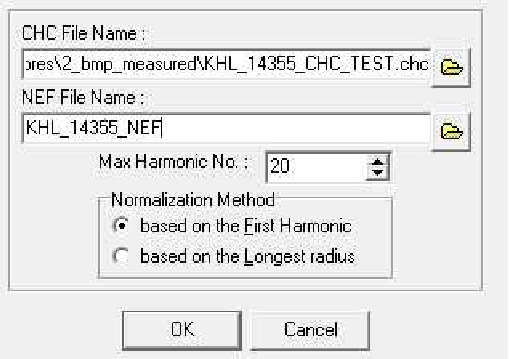
KHL_14355_CHC_TEST.chc

14 If the ChainCoder crashes at the step "Labeling Object", it is probably due to the difficulty to select the spores because of their small size. If this is the first image in the given specimen, restart ChainCoder and use a smaller value next to the "Labeling Object" button, e.g. 200 or 100.
If ChainCoder crashes not on the first image that is analyzed for the given specimen but later, be aware of the following. The chain codes from previous images (processed before the crash) are successfully written to .chc file and remain unaffected. The chain code(-s) for the image on which the crash happened will not be saved to .chc file. To solve the problem, close and then re-open ChainCoder. Import the image on which the crash happened and all subsequent images. Save the chain codes with a smaller value for "Labeling Object" button, e.g. 200 or 100. Write the chain codes to a new .chc file. After all images for the given specimen are processed, combine manually the contents of two .chc files. In this way, the chain codes obtained before the program crash and after the crash will be in a single file.

SHAPE: Chc2Nef

15 Open Chc2Nef program of the SHAPE package.
Select the chain-code file (.chc) to work on and create a NEF file (.nef) where the Normalized Elliptic Fourier Descriptors (NEFD) will be saved. Please note that .nef format used by Ch2Nef is not related to the Nikon Electronic Format that is abbreviated in the same way.
In the field for "Max Harmonic No" stay with the default value of 20. Each harmonic is a level of detalization of the contour information. Twenty harmonics represent the shape of the object very well, as authors of SHAPE package show in their tutorial (see the Reference to Iwata and Ukai 2002 in "Materials" panel). As a normalization method, use the approach based on the first harmonic.

Citation: Alexander Ordynets, Sarah KeÃÂler, Christina Willemsens (03/11/2021). Extracting shape and size information from fungal spores. https://dx.doi.org/10.17504/protocols.io.bdeii3ce



CHC File Name : ɔres\2_bmp_measured\KHL_14355_CHC_TEST.chc

NEF File Name : KHL_14355_NEF

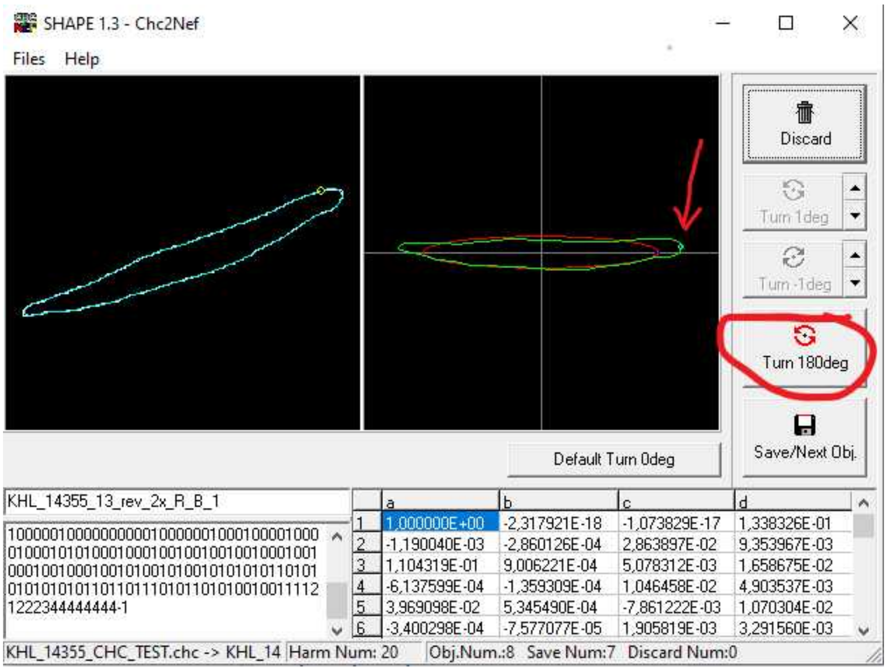
Max Harmonic No. : 20

Normalization Method

● based on the First Harmonic

○ based on the Longest radius

16 Press the "Start" button to check one by one all objects stored in the .chc file. When confident that the object is an intact spore (and not some cell fragment of the fungal fruiting body or not a crystal) and object orientation is correct (spore's hilar appendix is in the upper left quarter - already now or after 180-degree rotation (button "Turn 180deg"), and spore outline does not contain artifacts, send the NEFD data on this spore to the .nef file with the "Save/Next Obj." button. Proceed until all objects in the .chc file will be examined and good spores will be sent to a .nef file.



KHL_14355_13_rev_2x_R_B_1

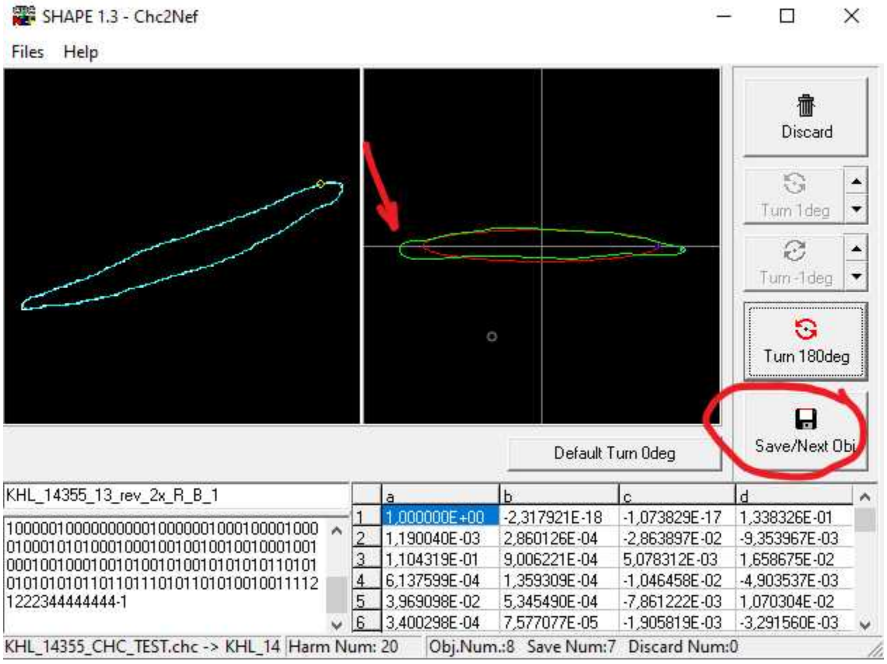
```
10000010000000000010000001000100001000
0100010101000100010010010010010001001
0001001000100101001010010101010110101
0101010101101101110101101010010011112
1222344444444-1
```

	a	b	c	d
1	1,000000E+00	-2,317921E-18	-1,073829E-17	1,338326E-01
2	-1,190040E-03	-2,860126E-04	2,863897E-02	9,353967E-03
3	1,104319E-01	9,006221E-04	5,078312E-03	1,658675E-02
4	-6,137599E-04	-1,359309E-04	1,046458E-02	4,903537E-03
5	3,969098E-02	5,345490E-04	-7,861222E-03	1,070304E-02
6	-3,400298E-04	-7,577077E-05	1,905819E-03	3,291560E-03

KHL_14355_CHC_TEST.chc -> KHL_14 Harm Num: 20 Obj.Num.:8 Save Num:7 Discard Num:0

Spore in Chc2Nef: original orientation on the image (left) and placement horizontally (right). However, as the hilar appendix does not stretch to the upper left corner of the image, the 180-degree rotation has to be applied.

Citation: Alexander Ordynets, Sarah KeÃÂler, Christina Willemsens (03/11/2021). Extracting shape and size information from fungal spores. https://dx.doi.org/10.17504/protocols.io.bdeii3ce



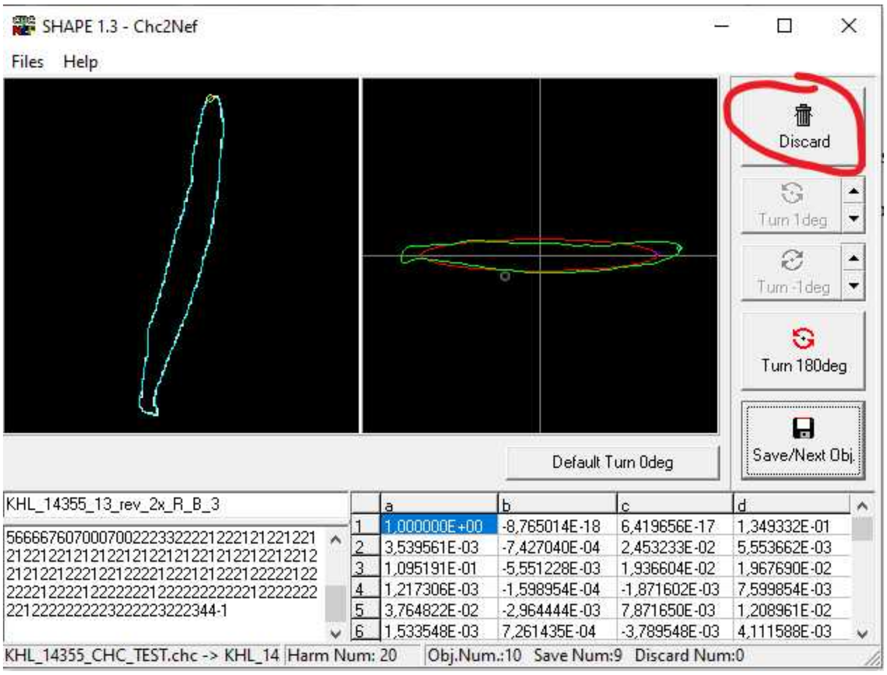
KHL_14355_13_rev_2x_R_B_1

```
1000001000000000001000000100010000100001000
0100010101000100010010010010010001001
00010010001001010010100101010110101
01010101011011011101011010100100111112
1222344444444-1
```

	a	b	c	d
1	1,000000E+00	-2,317921E-18	-1,073829E-17	1,338326E-01
2	1,190040E-03	2,860126E-04	-2,863897E-02	-9,353967E-03
3	1,104319E-01	9,006221E-04	5,078312E-03	1,658675E-02
4	6,137599E-04	1,359309E-04	-1,046458E-02	-4,903537E-03
5	3,969098E-02	5,345490E-04	-7,861222E-03	1,070304E-02
6	3,400298E-04	7,577077E-05	-1,905819E-03	-3,291560E-03

KHL_14355_CHC_TEST.chc -> KHL_14 Harm Num: 20 Obj.Num.:8 Save Num:7 Discard Num:0

It is the same spore as shown in the screenshot above but with 180-degree rotation applied. This is what we aimed for - the spore´s hilar appendix stretches to the upper left corner. We can save this outline to the .nef file and go on with the next object.



KHL_14355_13_rev_2x_R_B_3

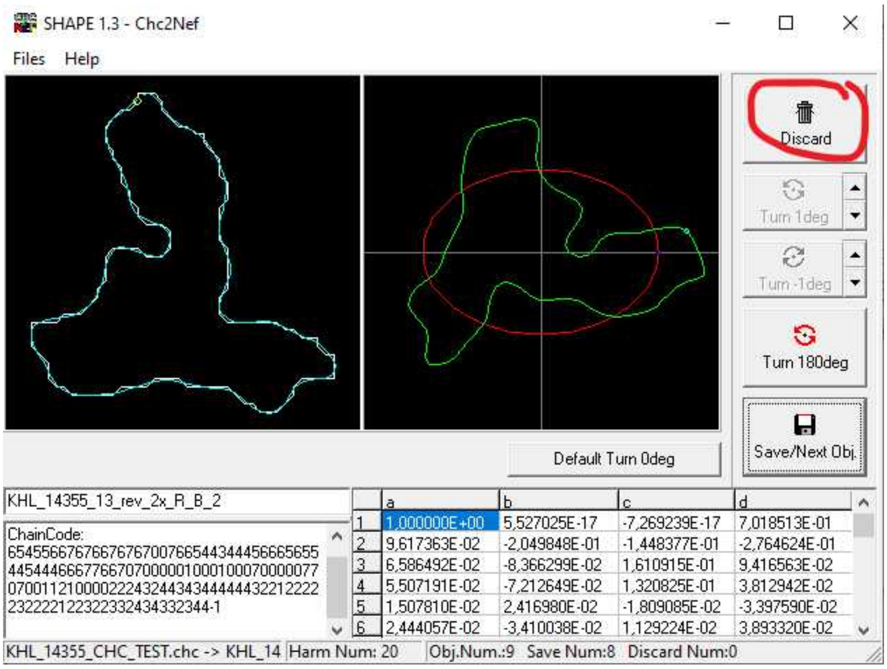
```
566667607000700222332222122212122122 1
212212212121221212212122121221221221 2
21212212221221222212221212221222221 22
222212222122222221222222222221222222 2
2212222222223222223222344-1
```

	a	b	c	d
1	1,000000E+00	-8,765014E-18	6,419656E-17	1,349332E-01
2	3,539561E-03	-7,427040E-04	2,453233E-02	5,553662E-03
3	1,095191E-01	-5,551228E-03	1,936604E-02	1,967690E-02
4	1,217306E-03	-1,598954E-04	-1,871602E-03	7,599854E-03
5	3,764822E-02	-2,964444E-03	7,871650E-03	1,208961E-02
6	1,533548E-03	7,261435E-04	-3,789548E-03	4,111588E-03

KHL_14355_CHC_TEST.chc -> KHL_14 Harm Num: 20 Obj.Num.:10 Save Num:9 Discard Num:0

This spore has some outline artifact - proximal part looks as a hook - and has to be discarded. i.e. not sent to the .nef file.

Citation: Alexander Ordynets, Sarah KeÃÂler, Christina Willemsens (03/11/2021). Extracting shape and size information from fungal spores. https://dx.doi.org/10.17504/protocols.io.bdeii3ce



KHL_14355_13_rev_2x_R_B_2

```
ChainCode:
65455667676676767007665443444566665655
4454446667766707000001000100070000077
0700112100002224324434344444432212222
232222122322332434332344-1
```

	a	b	c	d
1	1,000000E+00	5,527025E-17	-7,269239E-17	7,018513E-01
2	9,617363E-02	-2,049848E-01	-1,448377E-01	-2,764624E-01
3	6,586492E-02	-8,366299E-02	1,610915E-01	9,416563E-02
4	5,507191E-02	-7,212649E-02	1,320825E-01	3,812942E-02
5	1,507810E-02	2,416980E-02	-1,809085E-02	-3,397590E-02
6	2,444057E-02	-3,410038E-02	1,129224E-02	3,893320E-02

KHL_14355_CHC_TEST.chc -> KHL_14 Harm Num: 20 Obj.Num.:9 Save Num:8 Discard Num:0

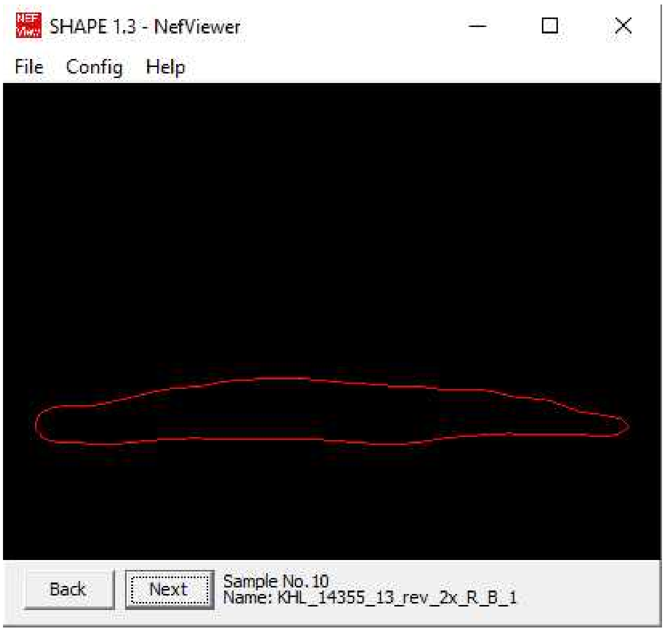
This object is not a spore and has to be discarded.

Ready .nef file with all spores of one specimen is attached below.

KHL_14355_NEF_TEST.zip

NefViewer

17 Open NefViewer to check that .nef file contains only "good" spores and only properly aligned spores.
If some not-spore or "bad" spore was mistakenly saved to the .nef file during the work in Ch2Nef, it is possible to open the .nef file in the text editor, to find the block of text that corresponds to NEFD of not target object and remove it.



Sample No. 10
Name: KHL_14355_13_rev_2x_R_B_1

Screenshot of NefViewer where the spores are checked one-by-one for being target objects and properly oriented. This screenshots shows the spore from image #13 that was demonstrated in Chc2Nef and was accepted for outline extraction after 180-degree rotation.

Citation: Alexander Ordynets, Sarah KeÃÂler, Christina Willemsens (03/11/2021). Extracting shape and size information from fungal spores. https://dx.doi.org/10.17504/protocols.io.bdeii3ce

This is how the NEFD code of the spore shown just above looks. The name allows identifying the piece of .nef file that corresponds to a spore as shown in the NefViewer.

```
KHL_14355_13_rev_2x_R_B_1
  1.0000000e+00  2.1160887e-17 -2.4515702e-17  1.3371701e-01
  1.1878045e-03  2.8494892e-04 -2.8696282e-02 -9.3624850e-03
  1.1043572e-01  8.9990081e-04  4.9820320e-03  1.6446673e-02
  6.1344530e-04  1.3195405e-04 -1.0501531e-02 -4.7999060e-03
  3.9693850e-02  5.3018451e-04 -7.9845890e-03  1.0748659e-02
  3.3567316e-04  7.2965502e-05 -1.7759919e-03 -3.2183380e-03
  1.9871497e-02  3.8722821e-04 -8.8202003e-04  8.0072021e-03
  2.0927678e-04  4.0102762e-05  2.0974518e-03 -2.4564797e-03
  1.1833938e-02  3.9541219e-04 -1.6406683e-03  3.8397591e-03
  2.5905899e-04  1.9414425e-05  3.3191499e-04 -2.0674836e-03
  7.7229891e-03  3.0094319e-04 -5.4529650e-04  4.4606975e-03
  2.7419533e-04 -3.9919521e-06 -3.6893713e-04 -1.5747850e-03
  5.3438975e-03  2.8856453e-04  6.8803535e-04  3.2827272e-03
  2.7557006e-04 -1.6574449e-05 -6.6664217e-04 -1.4059339e-03
  3.8668808e-03  2.8037904e-04  1.1016597e-03  2.4471208e-03
  2.3571746e-04 -7.9720335e-05  3.1456216e-04  3.9309737e-04
  2.9183263e-03  2.6148980e-04  2.8826305e-04  2.2490390e-03
  2.3599510e-04 -9.3077036e-05 -8.5402535e-06  6.3461805e-04
  2.2447106e-03  2.5143066e-04  1.2153733e-04  1.9634976e-03
  2.3956809e-04 -6.8605422e-05 -4.6178449e-04 -2.9917648e-04
```

Command prompt and text editor

18 When all NEF files for separate specimens are obtained, place their copies in one folder and combine them into a single file in Windows PowerShell. Open PowerShell in the folder where .nef files for separate specimens are placed and type:

```
Get-Content *.nef | Set-Content 30specimens_raw.nef
```

"30specimens.nef" is the arbitrary name of the resulting file, which informs of the amount of data in it.
Access the file here:

30specimens_raw.zip

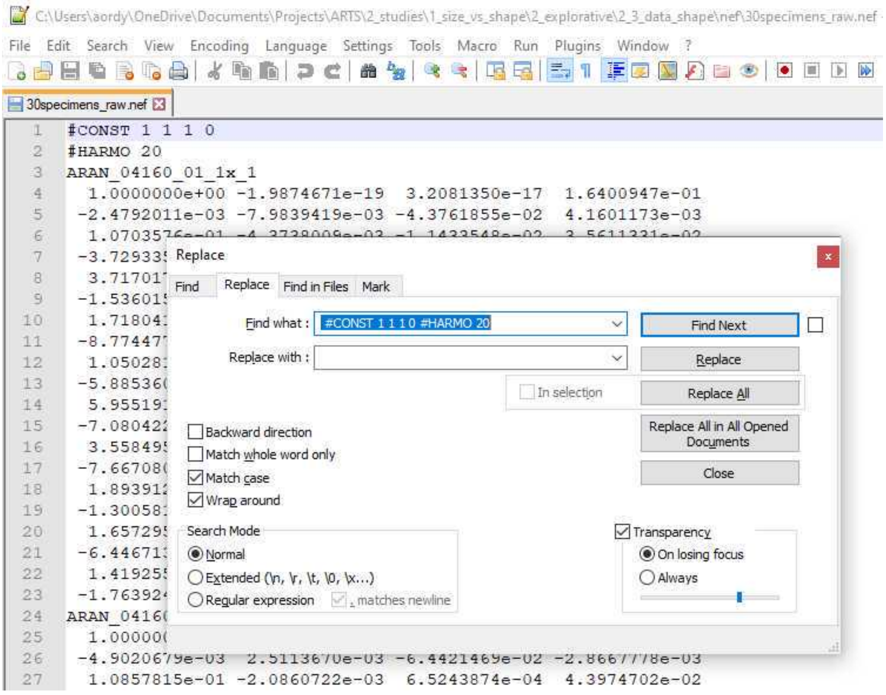
19 After the combined .nef is ready, remove the repeated text (three lines, the first line is empty) which was present at the top of each individual file from the previous step:

```

#CONST 1 1 1 0
#HARMO 20
```

I use Notepad to remove the specified text (literally replace with nothing). Thus, the resulting .nef file will have this text string only at the top of the file, i.e. two first lines, but nowhere else.

Citation: Alexander Ordynets, Sarah KeÃÂler, Christina Willemsens (03/11/2021). Extracting shape and size information from fungal spores. https://dx.doi.org/10.17504/protocols.io.bdeii3ce


C:\Users\aordy\OneDrive\Documents\Projects\ARTS\2_studies\1_size_vs_shape\2_explorative\2_3_data_shape\nef\30specimens_raw.nef

File Edit Search View Encoding Language Settings Tools Macro Run Plugins Window ?
30specimens_raw.nef

```
#CONST 1 1 1 0
#HARMO 20
ARAN_04160_01_1x_1
  1.0000000e+00 -1.9874671e-19  3.2081350e-17  1.6400947e-01
 -2.4792011e-03 -7.9839419e-03 -4.3761855e-02  4.1601173e-03
  1.0703576e-01 -4.3738009e-03 -1.1433548e-02  3.5611331e-02
 -3.72933
  3.71701
 -1.53601
  1.71804
 -8.77447
  1.05028
 -5.88536
  5.95519
 -7.08042
  3.55849
 -7.66708
  1.89391
 -1.30058
  1.65729
 -6.44671
  1.41925
 -1.76392
ARAN_0416
  1.00000
 -4.9020679e-03  2.5113670e-03 -6.4421469e-02 -2.8667778e-03
  1.0857815e-01 -2.0860722e-03  6.5243874e-04  4.3974702e-02
```

Replace

Find what : #CONST 1 1 1 0 #HARMO 20

Replace with :

□ In selection

□ Backward direction
□ Match whole word only
☑ Match case
☑ Wrap around

Search Mode
● Normal
○ Extended (\n, \r, \t, \0, \x...)
○ Regular expression ☑ . matches newline

☑ Transparency
● On losing focus
○ Always

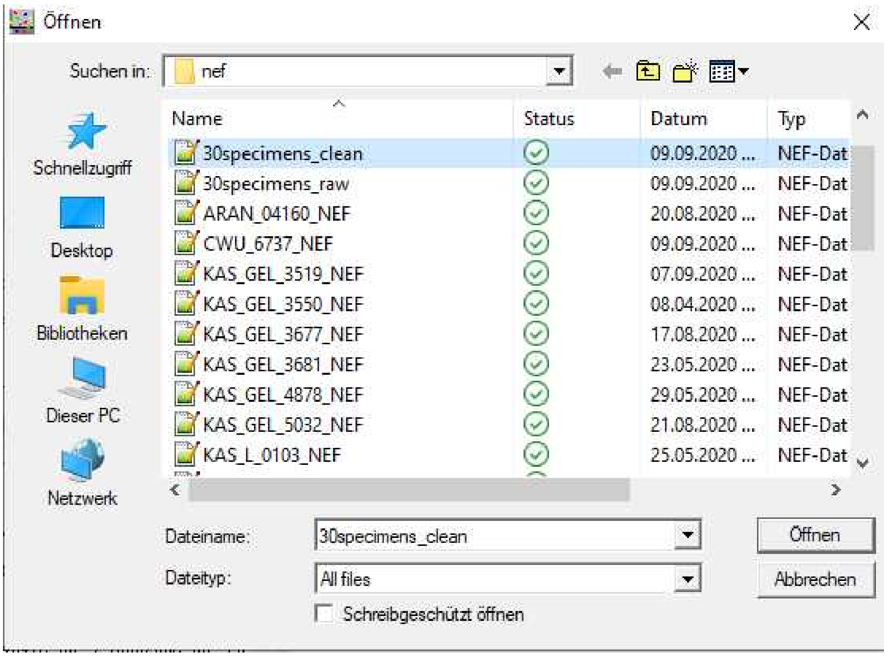
The adjusted .nef file should get another name, e.g. "30specimens_cleaned.nef".
If the replacement does not work in a single step, make it in two steps: first for the row starting as "#CONST" and then for the row starting with "#HARMO". NB! Remember that at the top of the file, these two lines have to be kept.

Access the file here:

30specimens_clean.zip

SHAPE: PrinComp+PrinPrint

20 Open PrinComp program to perform a Principal Component Analysis. Import the combined.nef file from the previous step.


Öffnen

Suchen in: nef

Name	Status	Datum	Typ
30specimens_clean	✓	09.09.2020 ...	NEF-Dat
30specimens_raw	✓	09.09.2020 ...	NEF-Dat
ARAN_04160_NEF	✓	20.08.2020 ...	NEF-Dat
CWU_6737_NEF	✓	09.09.2020 ...	NEF-Dat
KAS_GEL_3519_NEF	✓	07.09.2020 ...	NEF-Dat
KAS_GEL_3550_NEF	✓	08.04.2020 ...	NEF-Dat
KAS_GEL_3677_NEF	✓	17.08.2020 ...	NEF-Dat
KAS_GEL_3681_NEF	✓	23.05.2020 ...	NEF-Dat
KAS_GEL_4878_NEF	✓	29.05.2020 ...	NEF-Dat
KAS_GEL_5032_NEF	✓	21.08.2020 ...	NEF-Dat
KAS_L_0103_NEF	✓	25.05.2020 ...	NEF-Dat

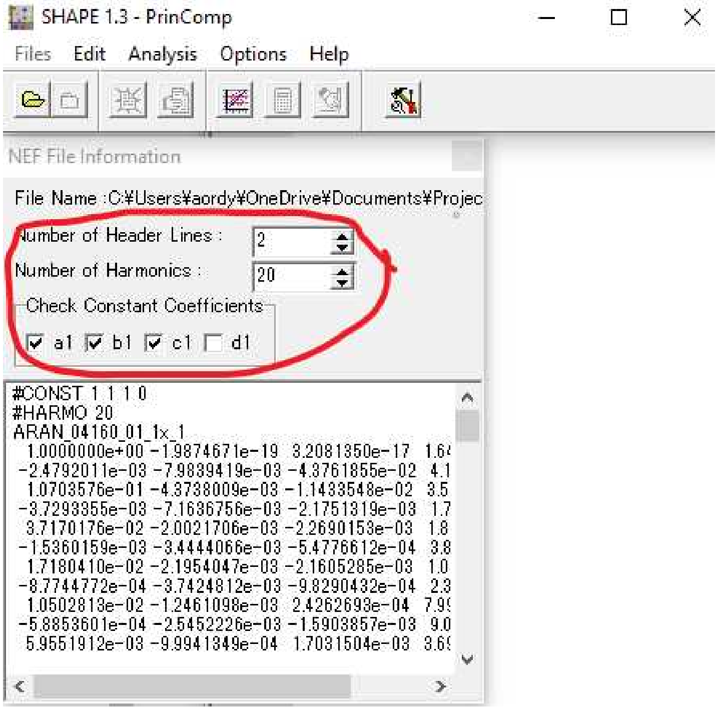
Dateiname: 30specimens_clean

Dateityp: All files

□ Schreibgeschützt öffnen

Citation: Alexander Ordynets, Sarah KeÃÂler, Christina Willemsens (03/11/2021). Extracting shape and size information from fungal spores. https://dx.doi.org/10.17504/protocols.io.bdeii3ce

21 If PrinComp successfully reads the file in, the windows as shown below will appear. Note the three first coefficients that PrinComp selects itself as constant: a1, b1, and c1. Here, nothing has to be changed.



File Name :C:¥Users¥aordy¥OneDrive¥Documents¥Projec

Number of Header Lines : 2

Number of Harmonics : 20

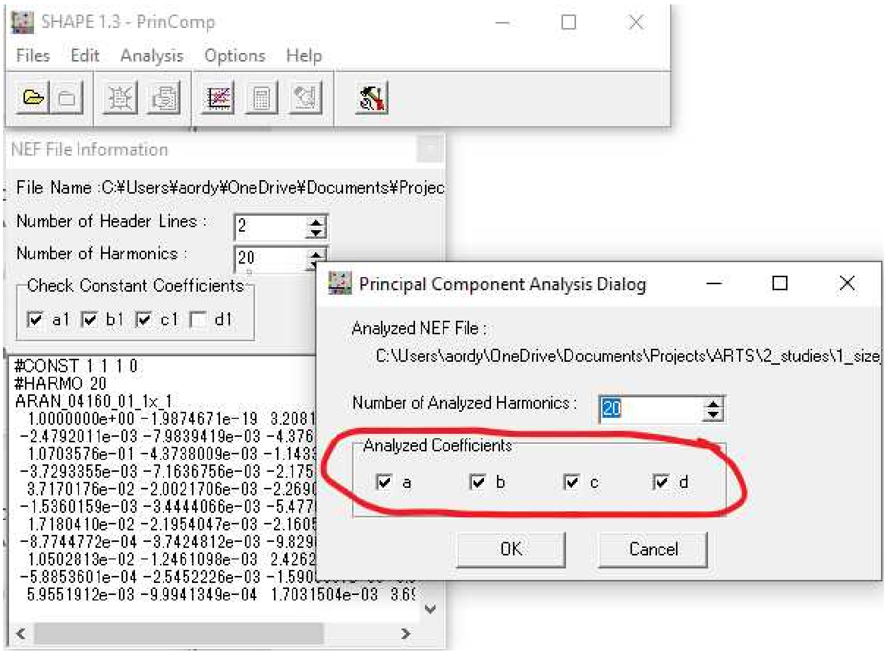
Check Constant Coefficients

☑ a1 ☑ b1 ☑ c1 □ d1

```
#CONST 1 1 1 0
#HARMO 20
ARAN_04160_01_1x_1
  1.0000000e+00 -1.9874671e-19  3.2081350e-17  1.64
 -2.4792011e-03 -7.9839419e-03 -4.3761855e-02  4.1
  1.0703576e-01 -4.3738009e-03 -1.1433548e-02  3.5
 -3.7293355e-03 -7.1636756e-03 -2.1751319e-03  1.7
  3.7170176e-02 -2.0021706e-03 -2.2690153e-03  1.8
 -1.5360159e-03 -3.4444066e-03 -5.4776612e-04  3.8
  1.7180410e-02 -2.1954047e-03 -2.1605285e-03  1.0
 -8.7744772e-04 -3.7424812e-03 -9.8290432e-04  2.3
  1.0502813e-02 -1.2461098e-03  2.4262693e-04  7.9
 -5.8853601e-04 -2.5452226e-03 -1.5903857e-03  9.0
  5.9551912e-03 -9.9941349e-04  1.7031504e-03  3.6
```

22 Start Principal Component Analysis by clicking the respective icon (with a plot).
The program can perform three versions of the analysis:
i) considering only the symmetric shape variation (leave selected only coefficients **a** and **d**);
ii) considering only the asymmetric shape variation (leave selected only coefficients **b** and **c**);
iii) considering total shape variation not differentiated into symmetric and asymmetric (keep selected **a, b, c,** and **d**). This will be called hereinafter a global variation.

This protocol demonstrates the analysis of global shape variation, i.e. the case when all four coefficients a, b, c and d will be selected for analyses.


SHAPE 1.3 - PrinComp
Files Edit Analysis Options Help
NEF File Information
Principal Component Analysis Dialog
OK
Cancel

File Name :C:¥Users¥aordy¥OneDrive¥Documents¥Projec

Number of Header Lines : 2

Number of Harmonics : 20

Check Constant Coefficients: ☑ a1 ☑ b1 ☑ c1 ☐ d1

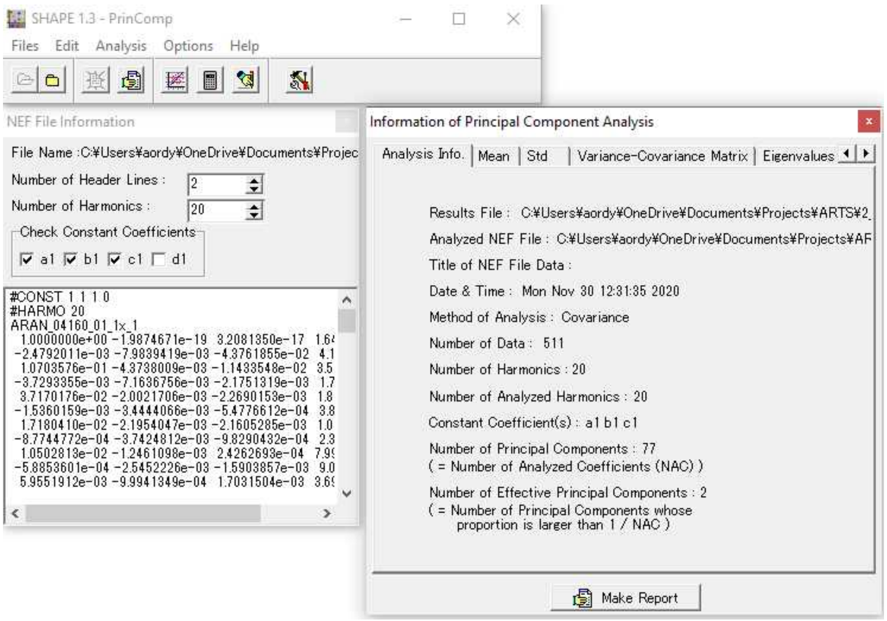
```
#CONST 1 1 1 0
#HARMO 20
ARAN_04160_01_1x_1
  1.0000000e+00 -1.9874671e-19  3.2081
 -2.4792011e-03 -7.9839419e-03 -4.376
  1.0703576e-01 -4.3738009e-03 -1.1433
 -3.7293355e-03 -7.1636756e-03 -2.175
  3.7170176e-02 -2.0021706e-03 -2.2690
 -1.5360159e-03 -3.4444066e-03 -5.477
  1.7180410e-02 -2.1954047e-03 -2.1605
 -8.7744772e-04 -3.7424812e-03 -9.829
  1.0502813e-02 -1.2461098e-03  2.4262
 -5.8853601e-04 -2.5452226e-03 -1.590
  5.9551912e-03 -9.9941349e-04  1.7031504e-03  3.6
```

Analyzed NEF File :
C:\Users\aordy\OneDrive\Documents\Projects\ARTS\2_studies\1_size

Number of Analyzed Harmonics : 20

Analyzed Coefficients: ☑ a ☑ b ☑ c ☑ d

23 The window with general information on the analysis will appear. At this moment, the file of the format .txt will be saved that contains general information about the performed analysis. It is good to keep the results from PCA in a separate folder.


SHAPE 1.3 - PrinComp
Files Edit Analysis Options Help
NEF File Information
Information of Principal Component Analysis
Analysis Info.
Mean
Std
Variance-Covariance Matrix
Eigenvalues
Make Report

File Name :C:¥Users¥aordy¥OneDrive¥Documents¥Projec

Number of Header Lines : 2

Number of Harmonics : 20

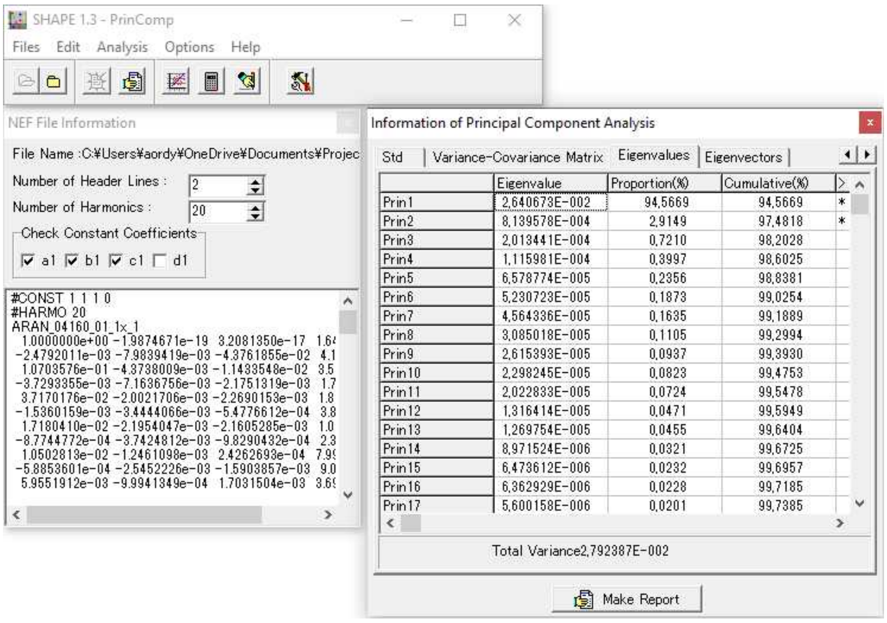
Check Constant Coefficients: ☑ a1 ☑ b1 ☑ c1 ☐ d1

```
#CONST 1 1 1 0
#HARMO 20
ARAN_04160_01_1x_1
  1.0000000e+00 -1.9874671e-19  3.2081350e-17  1.6
 -2.4792011e-03 -7.9839419e-03 -4.3761855e-02  4.1
  1.0703576e-01 -4.3738009e-03 -1.1433548e-02  3.5
 -3.7293355e-03 -7.1636756e-03 -2.1751319e-03  1.7
  3.7170176e-02 -2.0021706e-03 -2.2690153e-03  1.8
 -1.5360159e-03 -3.4444066e-03 -5.4776612e-04  3.8
  1.7180410e-02 -2.1954047e-03 -2.1605285e-03  1.0
 -8.7744772e-04 -3.7424812e-03 -9.8290432e-04  2.3
  1.0502813e-02 -1.2461098e-03  2.4262693e-04  7.9
 -5.8853601e-04 -2.5452226e-03 -1.5903857e-03  9.0
  5.9551912e-03 -9.9941349e-04  1.7031504e-03  3.6
```

```
Results File :  C:¥Users¥aordy¥OneDrive¥Documents¥Projects¥ARTS¥2_
Analyzed NEF File : C:¥Users¥aordy¥OneDrive¥Documents¥Projects¥AF
Title of NEF File Data :
Date & Time :  Mon Nov 30 12:31:35 2020
Method of Analysis :  Covariance
Number of Data :  511
Number of Harmonics : 20
Number of Analyzed Harmonics : 20
Constant Coefficient(s) : a1 b1 c1
Number of Principal Components : 77
( = Number of Analyzed Coefficients (NAC) )
Number of Effective Principal Components : 2
( = Number of Principal Components whose
     proportion is larger than 1 / NAC )
```

In the different panels, various elements of the output can be checked. For example, in the panel "Eigenvalues" one can see how much variation in % each principal component can explain:

Citation: Alexander Ordynets, Sarah KeÃÂler, Christina Willemsens (03/11/2021). Extracting shape and size information from fungal spores. https://dx.doi.org/10.17504/protocols.io.bdeii3ce


SHAPE 1.3 - PrinComp
Files Edit Analysis Options Help
NEF File Information
Information of Principal Component Analysis
Std
Variance-Covariance Matrix
Eigenvalues
Eigenvectors
Make Report

File Name :C:¥Users¥aordy¥OneDrive¥Documents¥Projec

Number of Header Lines : 2

Number of Harmonics : 20

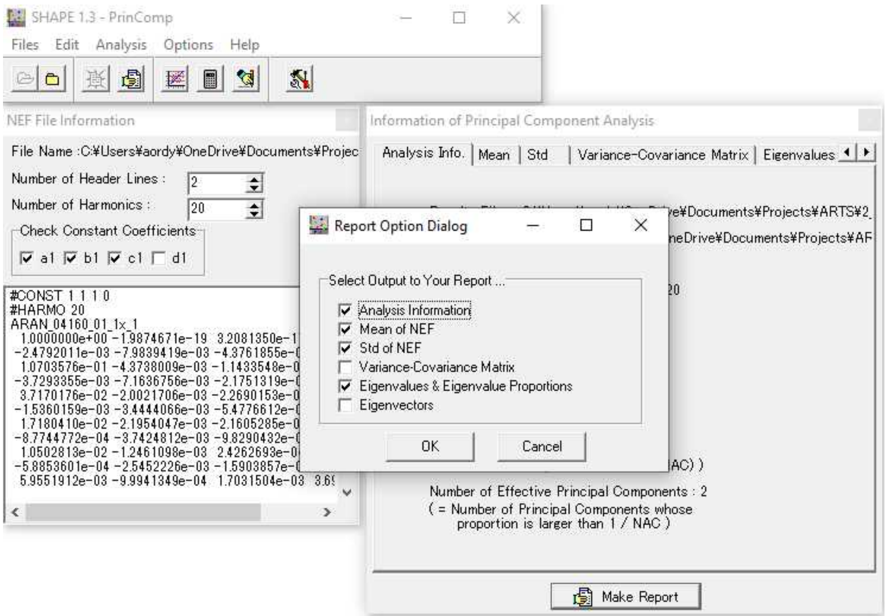
Check Constant Coefficients: ☑ a1 ☑ b1 ☑ c1 □ d1

```
#CONST 1 1 1 0
#HARMO 20
ARAN_04160_01_1x_1
  1.0000000e+00 -1.9874671e-19  3.2081350e-17  1.6
 -2.4792011e-03 -7.9839419e-03 -4.3761855e-02  4.1
  1.0703576e-01 -4.3738009e-03 -1.1433548e-02  3.5
 -3.7293355e-03 -7.1636756e-03 -2.1751319e-03  1.7
  3.7170176e-02 -2.0021706e-03 -2.2690153e-03  1.8
 -1.5360159e-03 -3.4444066e-03 -5.4776612e-04  3.8
  1.7180410e-02 -2.1954047e-03 -2.1605285e-03  1.0
 -8.7744772e-04 -3.7424812e-03 -9.8290432e-04  2.3
  1.0502813e-02 -1.2461098e-03  2.4262693e-04  7.9
 -5.8853601e-04 -2.5452226e-03 -1.5903857e-03  9.0
  5.9551912e-03 -9.9941349e-04  1.7031504e-03  3.6
```

	Eigenvalue	Proportion(%)	Cumulative(%)	>
Prin1	2,640673E-002	94,5669	94,5669	*
Prin2	8,139578E-004	2,9149	97,4818	*
Prin3	2,013441E-004	0,7210	98,2028	
Prin4	1,115981E-004	0,3997	98,6025	
Prin5	6,578774E-005	0,2356	98,8381	
Prin6	5,230723E-005	0,1873	99,0254	
Prin7	4,564336E-005	0,1635	99,1889	
Prin8	3,085018E-005	0,1105	99,2994	
Prin9	2,615393E-005	0,0937	99,3930	
Prin10	2,298245E-005	0,0823	99,4753	
Prin11	2,022833E-005	0,0724	99,5478	
Prin12	1,316414E-005	0,0471	99,5949	
Prin13	1,269754E-005	0,0455	99,6404	
Prin14	8,971524E-006	0,0321	99,6725	
Prin15	6,473612E-006	0,0232	99,6957	
Prin16	6,362929E-006	0,0228	99,7185	
Prin17	5,600158E-006	0,0201	99,7385	

Total Variance2,792387E-002

24 Click "Make Report", agree with the output components that the program offers to select.



File Name :C:¥Users¥aordy¥OneDrive¥Documents¥Projec

Number of Header Lines : 2

Number of Harmonics : 20

Check Constant Coefficients: ☑ a1 ☑ b1 ☑ c1 □ d1

ve¥Documents¥Projects¥ARTS¥2_

neDrive¥Documents¥Projects¥AF

Select Output to Your Report ...

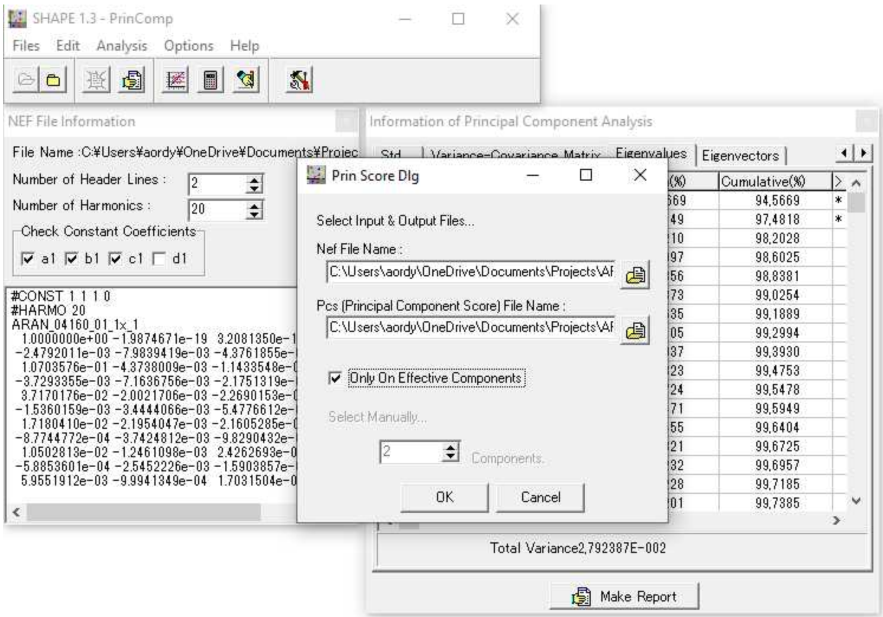
☑ Analysis Information
☑ Mean of NEF
☑ Std of NEF
□ Variance-Covariance Matrix
☑ Eigenvalues & Eigenvalue Proportions
□ Eigenvectors

Number of Effective Principal Components : 2
(= Number of Principal Components whose proportion is larger than 1 / NAC)

A file of the format .pcr will be exported with the detailed results. It can be opened in the text editor.

25 On the main panel of PrinComp, click "Calculate Principal Component Scores" (calculator icon). To export ONLY the scores of the PC axes that have an eigenvalue >1, select the box "Only On Effective Components". In this example, there are two effective components.
A file of the format .pcs will be exported with the principal component scores for each spore.

Citation: Alexander Ordynets, Sarah KeÃÂler, Christina Willemsens (03/11/2021). Extracting shape and size information from fungal spores. https://dx.doi.org/10.17504/protocols.io.bdeii3ce



File Name :C:¥Users¥aordy¥OneDrive¥Documents¥Projec

Number of Header Lines : 2

Number of Harmonics : 20

Check Constant Coefficients: ☑ a1 ☑ b1 ☑ c1 ☐ d1

```
#CONST 1 1 1 0
#HARMO 20
ARAN_04160_01_1x_1
  1.0000000e+00 -1.9874671e-19  3.2081350e-1
 -2.4792011e-03 -7.9839419e-03 -4.3761855e-
  1.0703576e-01 -4.3738009e-03 -1.1433548e-0
 -3.7293355e-03 -7.1636756e-03 -2.1751319e-
  3.7170176e-02 -2.0021706e-03 -2.2690153e-0
 -1.5360159e-03 -3.4444066e-03 -5.4776612e-
  1.7180410e-02 -2.1954047e-03 -2.1605285e-0
 -8.7744772e-04 -3.7424812e-03 -9.8290432e-
  1.0502813e-02 -1.2461098e-03  2.4262693e-0
 -5.8853601e-04 -2.5452226e-03 -1.5903857e-
  5.9551912e-03 -9.9941349e-04  1.7031504e-0
```

Select Input & Output Files...

Nef File Name : C:\Users\aordy\OneDrive\Documents\Projects\AF

Pcs (Principal Component Score) File Name : C:\Users\aordy\OneDrive\Documents\Projects\AF

☑ Only On Effective Components

Select Manually... 2 Components.

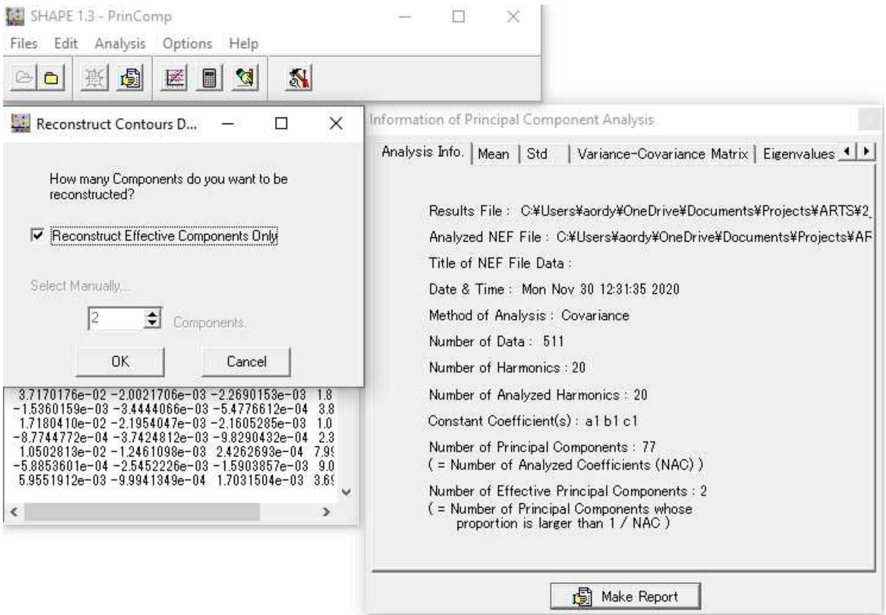
(%)	Cumulative(%)	>
669	94,5669	*
49	97,4818	*
10	98,2028	
97	98,6025	
56	98,8381	
73	99,0254	
35	99,1889	
05	99,2994	
37	99,3930	
23	99,4753	
24	99,5478	
71	99,5949	
55	99,6404	
21	99,6725	
32	99,6957	
28	99,7185	
01	99,7385	

Total Variance2,792387E-002

Access the file with score here:

30specimens_clean_glob.pcs

26 Click "Reconstruct Principal Component Contours" (icon with a hand and pencil) and propose a file name for the resulting .pcc file.



How many Components do you want to be reconstructed?

☑ Reconstruct Effective Components Only

Select Manually... 2 Components.

```
 3.7170176e-02 -2.0021706e-03 -2.2690153e-03  1.8
-1.5360159e-03 -3.4444066e-03 -5.4776612e-04  3.8
 1.7180410e-02 -2.1954047e-03 -2.1605285e-03  1.0
-8.7744772e-04 -3.7424812e-03 -9.8290432e-04  2.3
 1.0502813e-02 -1.2461098e-03  2.4262693e-04  7.9
-5.8853601e-04 -2.5452226e-03 -1.5903857e-03  9.0
 5.9551912e-03 -9.9941349e-04  1.7031504e-03  3.6
```

Results File : C:¥Users¥aordy¥OneDrive¥Documents¥Projects¥ARTS¥2_

Analyzed NEF File : C:¥Users¥aordy¥OneDrive¥Documents¥Projects¥AF

Title of NEF File Data :

Date & Time : Mon Nov 30 12:31:35 2020

Method of Analysis : Covariance

Number of Data : 511

Number of Harmonics : 20

Number of Analyzed Harmonics : 20

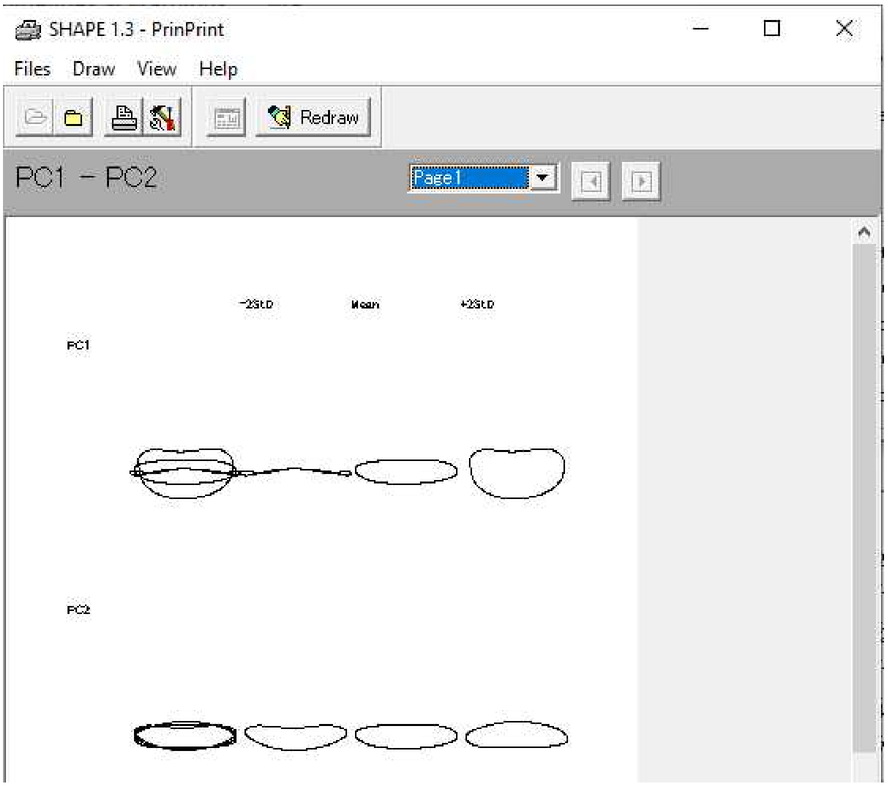
Constant Coefficient(s) : a1 b1 c1

Number of Principal Components : 77
(= Number of Analyzed Coefficients (NAC))

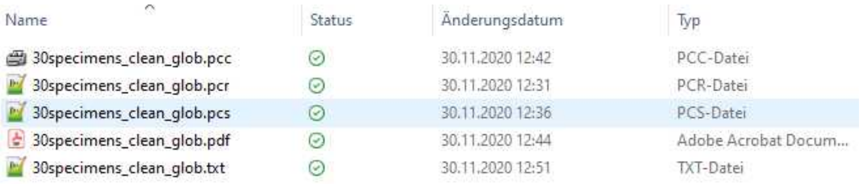
Number of Effective Principal Components : 2
(= Number of Principal Components whose proportion is larger than 1 / NAC)

27 The program "PrinPrint" will be launched automatically and will display the information stored in the .pcc file. Mean and plus and minus two standard deviations of the shape for each effective principal component will be displayed. Print this graphical output to the .pdf file (via printer icon).
Note that you can play with the position, line thickness, and colors of the displayed outlines (tools icon) before printing to a file.

Citation: Alexander Ordynets, Sarah KeÃÂler, Christina Willemsens (03/11/2021). Extracting shape and size information from fungal spores. https://dx.doi.org/10.17504/protocols.io.bdeii3ce



28 As an outcome, you should have a folder with five files: .txt, .pcr, .pcs, .pcc, and .pdf. Keep the file names identical for simplicity. All files are anyway of different formats but all (except pdf) can be viewed in the text editor. The .pcs is the most important for further analyses.



Name	Status	Änderungsdatum	Typ
30specimens_clean_glob.pcc	✓	30.11.2020 12:42	PCC-Datei
30specimens_clean_glob.pcr	✓	30.11.2020 12:31	PCR-Datei
30specimens_clean_glob.pcs	✓	30.11.2020 12:36	PCS-Datei
30specimens_clean_glob.pdf	✓	30.11.2020 12:44	Adobe Acrobat Docum...
30specimens_clean_glob.txt	✓	30.11.2020 12:51	TXT-Datei

R: shape data summary

29 Average the PCA scores of individual spores at the specimen level in R. Do this step if you need a single trait value for a specimen and not interested in the shape variation between the spores of a single specimen.
If more than one PCA analysis was performed (e.g. separately for symmetric and asymmetric variation), apply this operation to each PCA outcome separately.
R code for the case of global shape variation is provided below.

```
# load pca scores
library(here)
scores_glob_raw <- read.csv(here("2_3_datashape", "pca_202009", "3_glob",
"30specimens_clean_glob.pcs"),  sep="\t")

# add column to be filled in later with specimen IDs
scores_glob_raw<-cbind(Specimen_ID=NA, scores_glob_raw)

# create the "strings" object:
library(stringr)
```

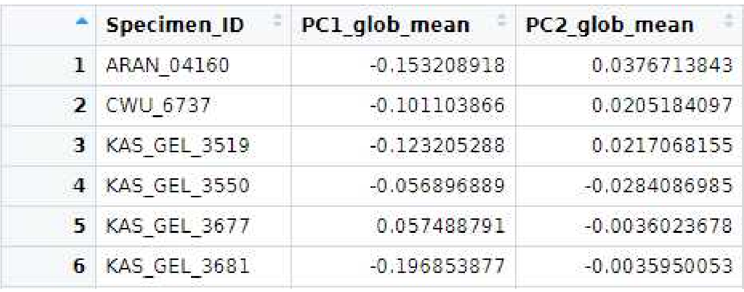
Citation: Alexander Ordynets, Sarah KeÃÂler, Christina Willemsens (03/11/2021). Extracting shape and size information from fungal spores. https://dx.doi.org/10.17504/protocols.io.bdeii3ce

```
strings <- c("^ARAN", "^CWU", "^KHL", "^LR", "^LY", "^Ordynets")

# fill in the column "Specimen_ID" using the content of DATA_NAME and regular expressions
scores_glob_raw$Specimen_ID <- ifelse(str_detect(scores_glob_raw$DATA_NAME, paste(strings,
collapse = "|")), gsub("^([^_]*_[^_]*)_.*$", "\\1", scores_glob_raw$DATA_NAME),
gsub("^([^_]*_[^_]*_[^_]*)_.*$", "\\1", scores_glob_raw$DATA_NAME))

# summarize global variation at the specimen level
library(dplyr)
scoresglobsumm <- scoresglobraw %>%
group_by(Specimen_ID) %>%
summarise(PC1_glob_mean=mean(PC1),
          PC2_glob_mean=mean(PC2),
          .groups = 'drop')
```

30 The resulting data frame should look as follows:



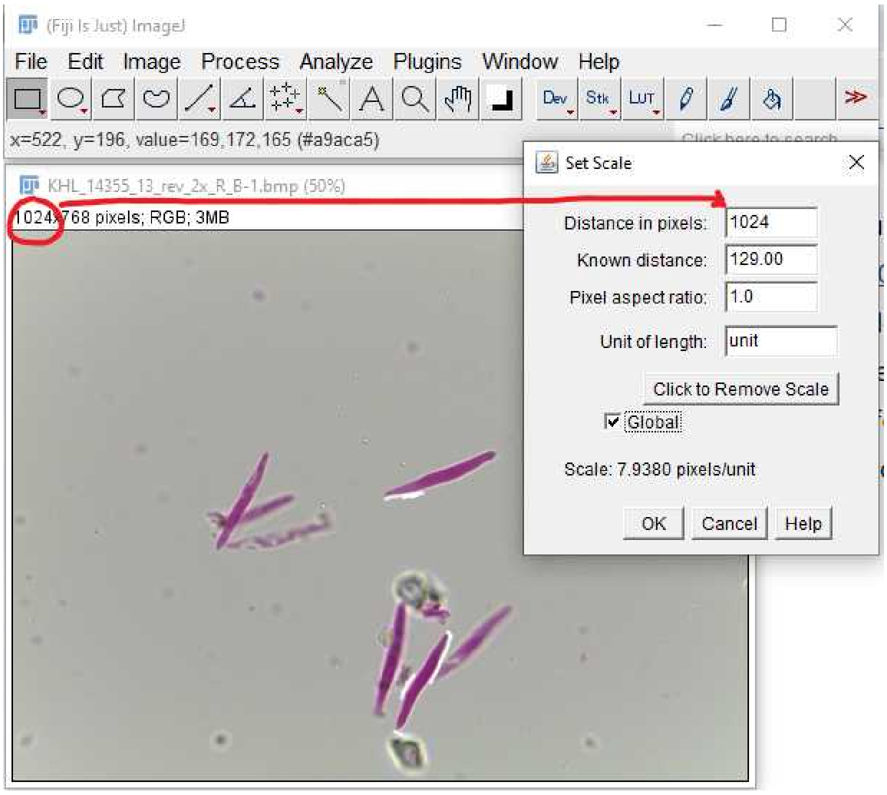
	Specimen_ID	PC1_glob_mean	PC2_glob_mean
1	ARAN_04160	-0.153208918	0.0376713843
2	CWU_6737	-0.101103866	0.0205184097
3	KAS_GEL_3519	-0.123205288	0.0217068155
4	KAS_GEL_3550	-0.056896889	-0.0284086985
5	KAS_GEL_3677	0.057488791	-0.0036023678
6	KAS_GEL_3681	-0.196853877	-0.0035950053

Screenshot from RStudio of the "score_glob_summ" data frame

Fiji: linear measurements

31 For exactly those spores for which NEFD were successfully obtained, take also the linear measurements.
Measure spore length and width with Fiji distribution of ImageJ .
Open Fiji. Open the picture and set the scale (Analyze > Set Scale). In our lab, we know that full image width equals 129 mkm. However, an image can be of different sizes in pixels. In our dataset, it is either 1024 or 2048. This has to be accounted for in the "Set Scale" window. Tick the box "Global" to make the scaling valid for all images from this specimen if you are sure they are all of the same sizes.

Citation: Alexander Ordynets, Sarah KeÃÂler, Christina Willemsens (03/11/2021). Extracting shape and size information from fungal spores. https://dx.doi.org/10.17504/protocols.io.bdeii3ce



x=522, y=196, value=169,172,165 (#a9aca5)

KHL_14355_13_rev_2x_R_B-1.bmp (50%)

1024x768 pixels; RGB; 3MB

Distance in pixels: 1024

Known distance: 129.00

Pixel aspect ratio: 1.0

Unit of length: unit

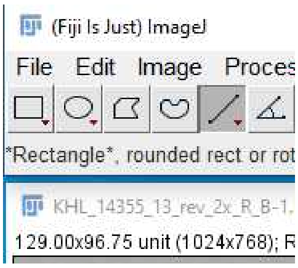
Click to Remove Scale

☑ Global

Scale: 7.9380 pixels/unit

OK Cancel Help

32 Select measuring mode "Straight".



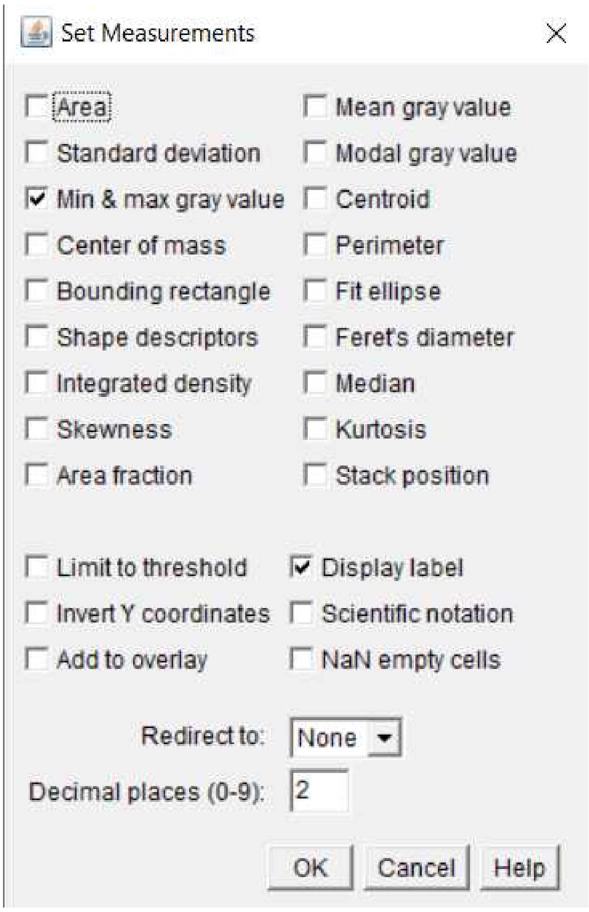
Rectangle, rounded rect or rot

KHL_14355_13_rev_2x_R_B-1.

129.00x96.75 unit (1024x768); R

33 In Analyze > Set Measurements, specify which measurements exactly have to be saved. Select "Min and max grey value" and "Display label". In this way, the program will save starting and ending points of the line to calculate the length, and labels will be kept for each measurement, to be able to understand from which image and for which specimen the measurement was taken

Citation: Alexander Ordynets, Sarah KeÃÂler, Christina Willemsens (03/11/2021). Extracting shape and size information from fungal spores. https://dx.doi.org/10.17504/protocols.io.bdeii3ce



Set Measurements

☐ Area ☐ Mean gray value
☐ Standard deviation ☐ Modal gray value
☑ Min & max gray value ☐ Centroid
☐ Center of mass ☐ Perimeter
☐ Bounding rectangle ☐ Fit ellipse
☐ Shape descriptors ☐ Feret's diameter
☐ Integrated density ☐ Median
☐ Skewness ☐ Kurtosis
☐ Area fraction ☐ Stack position

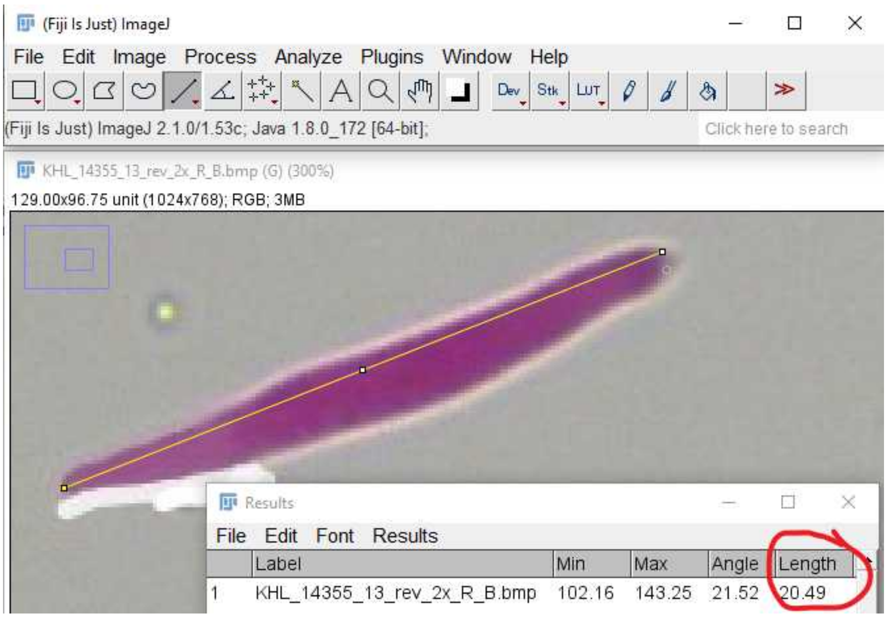
☐ Limit to threshold ☑ Display label
☐ Invert Y coordinates ☐ Scientific notation
☐ Add to overlay ☐ NaN empty cells

Redirect to: None

Decimal places (0-9): 2

34 Fiji does not differentiate whether the measurement is of a category “length” or “width”. It will actually call all the measurements "Length" and will put them in a single column. To cope with it, measure constantly firstly the length and then the width of each spore. Later, it will be possible to split the measurements like we need because the length will always be in an odd place in a column and width - in an even place.
Start with the length. Connect the farthermost points along the long spore axis by holding a left mouse button pressed. If the line was drawn not as you expected, make it again - measurement has not been saved yet. If the line is fine, click Ctrl+M, and a table "Results" will pop up. The rightmost column will contain a value in micrometers (because we set the scale earlier).

Citation: Alexander Ordynets, Sarah KeÃŸler, Christina Willemsens (03/11/2021). Extracting shape and size information from fungal spores. https://dx.doi.org/10.17504/protocols.io.bdeii3ce

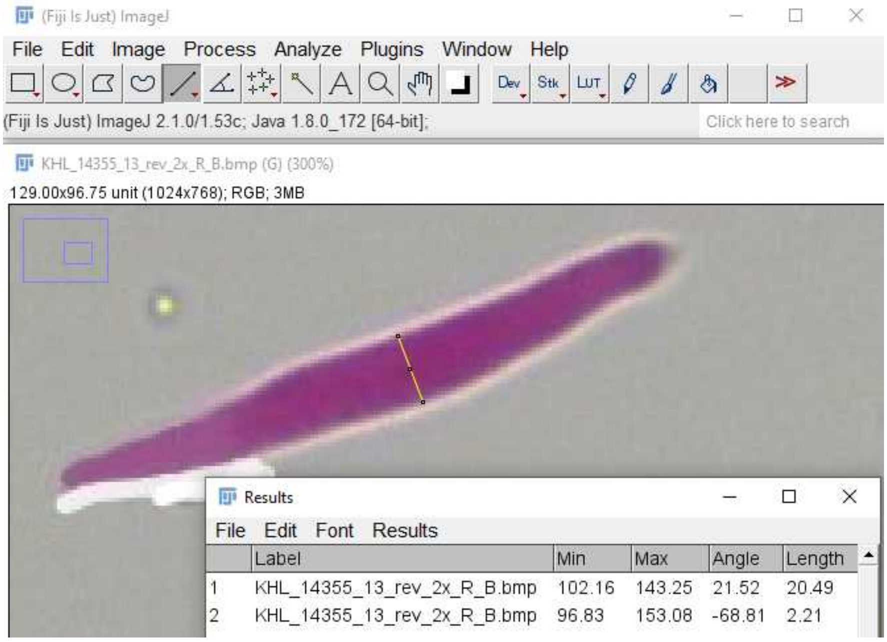

(Fiji Is Just) ImageJ
File Edit Image Process Analyze Plugins Window Help
Dev Stk LUT
Click here to search
KHL_14355_13_rev_2x_R_B.bmp (G) (300%)
Results
File Edit Font Results

(Fiji Is Just) ImageJ 2.1.0/1.53c; Java 1.8.0_172 [64-bit];

129.00x96.75 unit (1024x768); RGB; 3MB

	Label	Min	Max	Angle	Length
1	KHL_14355_13_rev_2x_R_B.bmp	102.16	143.25	21.52	20.49

35 Now measure the width - it will be a second measurement for this spore, i.e. recorded in the second row of the "Results" table.


(Fiji Is Just) ImageJ
File Edit Image Process Analyze Plugins Window Help
Dev Stk LUT
Click here to search
KHL_14355_13_rev_2x_R_B.bmp (G) (300%)
Results
File Edit Font Results

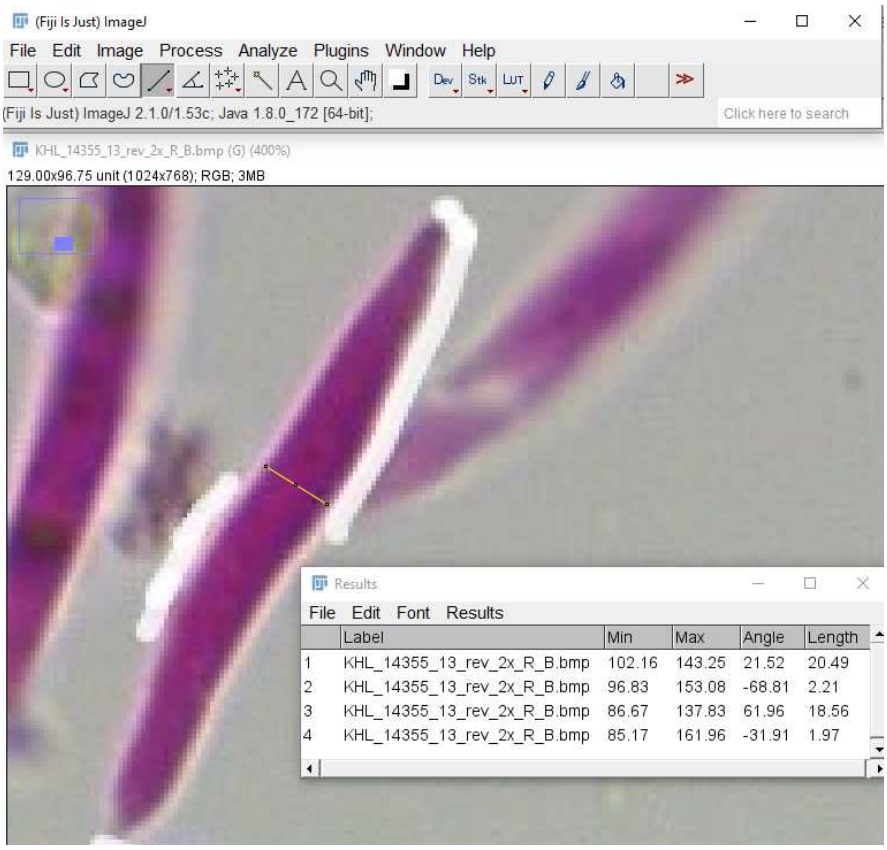
(Fiji Is Just) ImageJ 2.1.0/1.53c; Java 1.8.0_172 [64-bit];

129.00x96.75 unit (1024x768); RGB; 3MB

	Label	Min	Max	Angle	Length
1	KHL_14355_13_rev_2x_R_B.bmp	102.16	143.25	21.52	20.49
2	KHL_14355_13_rev_2x_R_B.bmp	96.83	153.08	-68.81	2.21

36 After another spore is measured (again: first length, then width), the table will contain four values in the "Length" column.

Citation: Alexander Ordynets, Sarah KeÃÂler, Christina Willemsens (03/11/2021). Extracting shape and size information from fungal spores. https://dx.doi.org/10.17504/protocols.io.bdeii3ce

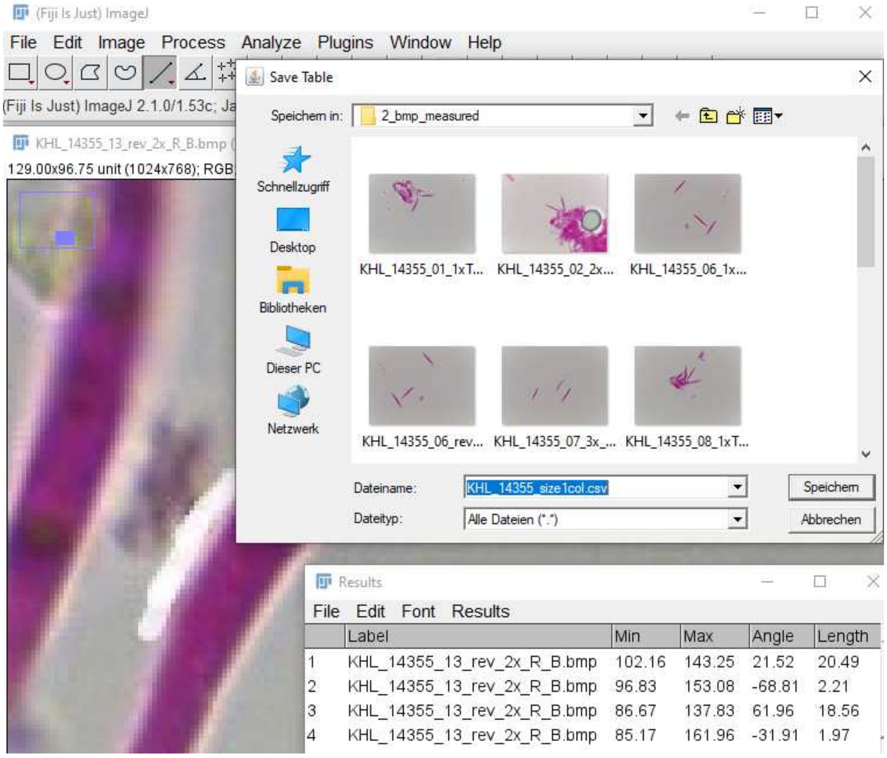

(Fiji Is Just) ImageJ
File Edit Image Process Analyze Plugins Window Help
Click here to search
KHL_14355_13_rev_2x_R_B.bmp (G) (400%)
Results
File Edit Font Results

(Fiji Is Just) ImageJ 2.1.0/1.53c; Java 1.8.0_172 [64-bit];

129.00x96.75 unit (1024x768); RGB; 3MB

	Label	Min	Max	Angle	Length
1	KHL_14355_13_rev_2x_R_B.bmp	102.16	143.25	21.52	20.49
2	KHL_14355_13_rev_2x_R_B.bmp	96.83	153.08	-68.81	2.21
3	KHL_14355_13_rev_2x_R_B.bmp	86.67	137.83	61.96	18.56
4	KHL_14355_13_rev_2x_R_B.bmp	85.17	161.96	-31.91	1.97

37 Repeat the procedure for all target spores of the given specimen. This means closing the processed image and opening one by one the others. The measurements will be put to a single table that has to be saved as a .csv file.



(Fiji Is Just) ImageJ 2.1.0/1.53c; Ja

KHL_14355_13_rev_2x_R_B.bmp (

129.00x96.75 unit (1024x768); RGB

Speichern in: 2_bmp_measured

Dateiname: KHL_14355_size1col.csv

Dateityp: Alle Dateien (*.*)

	Label	Min	Max	Angle	Length
1	KHL_14355_13_rev_2x_R_B.bmp	102.16	143.25	21.52	20.49
2	KHL_14355_13_rev_2x_R_B.bmp	96.83	153.08	-68.81	2.21
3	KHL_14355_13_rev_2x_R_B.bmp	86.67	137.83	61.96	18.56
4	KHL_14355_13_rev_2x_R_B.bmp	85.17	161.96	-31.91	1.97

Citation: Alexander Ordynets, Sarah KeÃÂler, Christina Willemsens (03/11/2021). Extracting shape and size information from fungal spores. https://dx.doi.org/10.17504/protocols.io.bdeii3ce

38 NB! If you keep Fiji open and go on with the next specimen, check whether the image size in pixels differs there. If yes, re-do the calibration (Analyze > Set Scale).

R: size data summary

39 Pool data for separate specimens in R. The pre-requisite is that .csv files are placed in a single folder.
R code:

```
library(here)
library(readr)
import.size <- dir(here("2_2_data_size"), pattern = "*.csv", full.names = T)
data_size_raw <- suppressWarnings(plyr::ldply(import.size, read_csv))
```

40 Move the resulting measurements into separate length and width columns in R.
R code:

```
library(dplyr)
# get odd values
data_size_raw_odd <- data_size_raw %>% dplyr::slice(which(row_number() %% 2 == 1))
colnames(data_size_raw_odd)[which(names(data_size_raw_odd) == "Label")] <- "Photo_ID"

# get even values
data_size_raw_even <- data_size_raw %>% dplyr::slice(which(row_number() %% 2 == 0))
colnames(data_size_raw_even)[which(names(data_size_raw_even) == "Length")] <- "Width"

# bind the columns
data_size_tiny <- bind_cols(data_size_raw_odd, data_size_raw_even)

# rename the columns
data_size_tiny_simple <-data_size_tiny[, c("Photo_ID", "Length", "Width")]
```

41 Based on length and width values, calculate their ratio to use it as an additional variable in comparative analyses.
R code:

```
data_size_tiny_simple <- transform(data_size_tiny_simple, Length_to_width_ratio = Length /
Width)
data_size_tiny_simple$Length_to_width_ratio <-
round(data_size_tiny_simple$Length_to_width_ratio, 2)
```

42 Create the new column "Specimen_ID" and fill it using the information from the column "Photo_ID" and regular expressions.
The challenge in this case is that some Specimen IDs will inherit from Photo IDs a part of the string before the second underscore while some - before the third underscore.
R code:

```
data_size_tiny_simple <- cbind(Specimen_ID=NA, datasizetinysimple)

library(stringr)
strings <- c("^ARAN", "^CWU", "^KHL", "^LR", "^LY", "^Ordynets")
data_size_tiny_simple$Specimen_ID <- ifelse(str_detect(datsizetiny_simple$Photo_ID,
paste(strings, collapse = "|")), gsub("^([^_]*_[^_]*)_.*$", "\\1",
data_size_tiny_simple$Photo_ID), gsub("^([^_]*_[^_]*_[^_]*)_.*$", "\\1",
data_size_tiny_simple$Photo_ID))
```

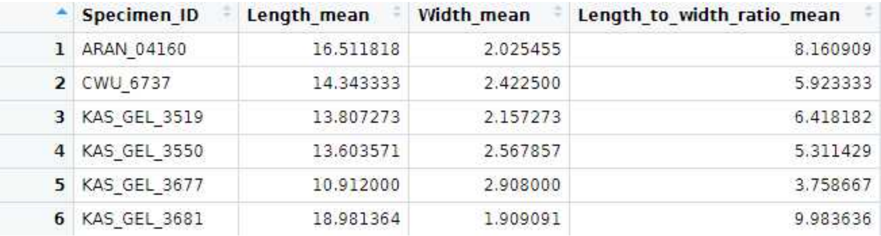
Citation: Alexander Ordynets, Sarah KeÃÂler, Christina Willemsens (03/11/2021). Extracting shape and size information from fungal spores. https://dx.doi.org/10.17504/protocols.io.bdeii3ce

For those wishing more details, the R code above using regex is based on the solution from:
## https://stackoverflow.com/questions/39366759/regex-to-extract-values-between-2-underscores-including-a-value-that-is-an-unde
## https://stackoverflow.com/questions/7449564/regex-return-all-before-the-second-occurrence

43 Summarize the data on a specimen level, by averaging, if you need a single trait value per specimen.
R code:

```
library(dplyr)
size_summ <- data_size_tiny_simple %>%
  group_by(Specimen_ID) %>%
  summarise(Length_mean=mean(Length),
            Width_mean=mean(Width),
            Length_to_width_ratio_mean=mean(Length_to_width_ratio),
            .groups = 'drop')
```

44 The resulting data frame should look as follows:



	Specimen_ID	Length_mean	Width_mean	Length_to_width_ratio_mean
1	ARAN_04160	16.511818	2.025455	8.160909
2	CWU_6737	14.343333	2.422500	5.923333
3	KAS_GEL_3519	13.807273	2.157273	6.418182
4	KAS_GEL_3550	13.603571	2.567857	5.311429
5	KAS_GEL_3677	10.912000	2.908000	3.758667
6	KAS_GEL_3681	18.981364	1.909091	9.983636

Screenshot from RStudio of the "size_summ" data frame

The averaged per specimen values of traits (in our example spore shape PC scores and size linear measurements) can be used in further analyses in R or beyond.
End of the protocol.