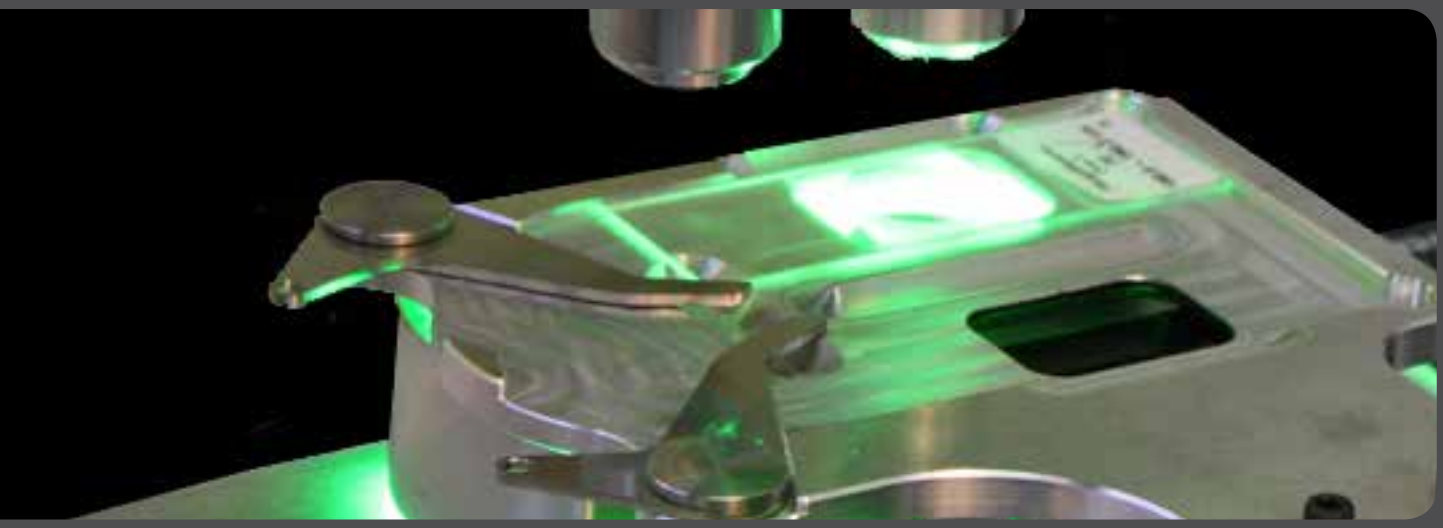
Classifynder™

newsletter issue #2| **05.13**



Classifynder; an intelligent assistant to the Palynologist. Explore new oceans.

ABOUT THE CLASSIFY*NDER*

- No other affordable desk top system exists today that is able to achieve an acceptable level of accuracy, repeatability and consistency in the location and classification of fresh pollen and spores mounted on glass slides. It offers researchers the ability to explore new horizons in the science of palynology and provides significant time and labour savings.
- The Classifynder software runs on a conventional PC or laptop, is easy to use and requires no specialist knowledge of image processing techniques or user interventions.
- Additional benefits include the automated creation of libraries of pollen and spores that are in focus across the full depth of field and metadata very suitable for the use in pollen database or key creation. A researcher can subsequently return to explore all or any of the slide objects without remounting the slide. The system also visualises the spatial distribution of all objects of interest on the slide so that a check can be made of any object segregation or related slide quality issues.
- Continued improvements over the last two years have resulted in many gains for users. Read on to learn about what has been done, who is doing it and where we are going.

This Newsletter has been prepared for our current users and those considering the purchase of our systems.

In this edition:

Editorial:

«New Oceans»

by Emeritus Professor John Flenley, Massey University, New Zealand

We are close to the centenary of Palynology, which started in Sweden in 1916 when Lennart von Post published the first pollen diagram. So it is timely to look back and note the ways in which the subject has developed, how Palynology has contributed to other branches of science and its current limitations.

Lennart von Post concentrated on tree pollen and related their changes in frequency to changes in climate during the post-glacial era. These data contributed greatly to geology at the time, when there was still confusion in the minds of some over the difference between ice-ages and deluges.

The next phase of development began in 1941 when Iversen became the first to recognise Neolithic forest clearances in Scandinavia. The relationship with archaeology has continued to grow ever since. Pollen assemblages give a generalised view of happenings over quite a sizeable area. Archaeology gives detailed evidence of what happened at one point. Both contribute to an understanding of human pre-history.

Many other applications of palynology have arisen over the years. For example pre-Quaternary palynology has given us much information about the date of evolutionary origin of many taxa. It is also invaluable in the stratigraphic studies which have been essential in the search for oil and gas.

On another front, the careful analysis of pollen which is collected by bees in their search for nectar has given rise to Melissopalynology. This can tell us much about where honey came from and thus assist in the detection of unscrupulous dilution of valuable honey varieties with cheaper types. The valuable types, such as Manuka honey, may yet have medical significance.

It isn't just bees that collect pollen: humans do also, mostly unwittingly. This is of value in forensic pollen analysis and the tracking of persons or objects of interest. Of course pollen can be a cause of serious allergies to some people and pollen traps to survey the pollen content of the atmosphere are in place in many countries in the world. Atmospheric pollen data is now regularly analysed and is invaluable in predicting and mitigating the effects for serious allergy suffers. The same data may also be used by Palaeoclimatologists to make more detailed surveys of past climates.

As a result of all these developments – and many others – there are thousands of trained Palynologists throughout the world, who spend long hours looking down microscopes.

But Palynology as we know it today could make bigger strides if we could have larger pollen counts per sample. Most Palaeoecologists working in temperate regions are happy with 150 tree pollen grains per sample, as this gives a pollen diagram in which it is possible to draw zone boundaries based on the frequencies of the major trees. But larger counts per sample would enable analysis of rare pollen types to reach statistically meaningful levels. And if one moves to the tropical rain forest, counts of several hundred tree pollen grains are indeed needed, because the forest is so much more bio-diverse.

Another improvement would be more frequent sampling in the analysis of cores to give finer temporal resolution. At present samples every 10cm apart are commonly used, which typically represents one sample about every 100 years. This can limit the chance of discovery of individual significant events such as storms, fires, early

human disturbance, or volcanic events.

Further, present day palynology can be insufferably slow. If you are drilling for oil, delays in obtaining pollen results can cost many thousands of dollars as you waste time drilling beyond the prospective oil-bearing horizon. If you are an academic, results come in too slowly to satisfy your managerial bosses!

This is where the Classifynder comes in. Classifynder is a digital microscope which can scan a slide and image all pollen-like objects with a resolution equivalent to x1000 magnification, and a depth of field similar to SEM. It then compares the images with reference images and classifies the images using mathematical techniques. The resulting groups are shown to the operator, who can correct any mis-classification if necessary. The overall accuracy of this process compares rather favourably with traditional manual methods. The result is more pollen counts and classifications in less time (and with less eye strain).

So as the centenary of Palynology is nearly upon us, perhaps it is time for the Palynological community to give these digital imaging and classification approaches more serious consideration. As noted by André Gide, winner of the 1947 Nobel Prize in literature, "Man cannot discover new oceans unless he has the courage to lose sight of the shore."

News:

User suggested improvements now available.

It is now over two years since the first Series 3 Classifynder machines were delivered to their users and the initial user-training workshop were held at Massey University (for details refer back to Newsletter 1).

Since then we have listened to user feedback provided both at that workshop and in subsequent correspondence with users. Many improvements have been made that make the Classifynder system easier to use, give higher quality images and have also improved the capability to deal with low contrast objects and small objects of less than 5 microns in diameter.

The images shown below, and distributed through this Newsletter, illustrate the image quality improvements. In each case the images on the left were captured with Release 1 Software and a Series 3 Classifynder. The images on the right were captured with Release 2 Software and a Series 3A Classifynder.

Evident in these images is the very high contrast between the objects and the background; in the later images this gives improved edge details. We outline the specific software and hardware changes we have made that contribute to these improvements in the following sections of this newsletter.

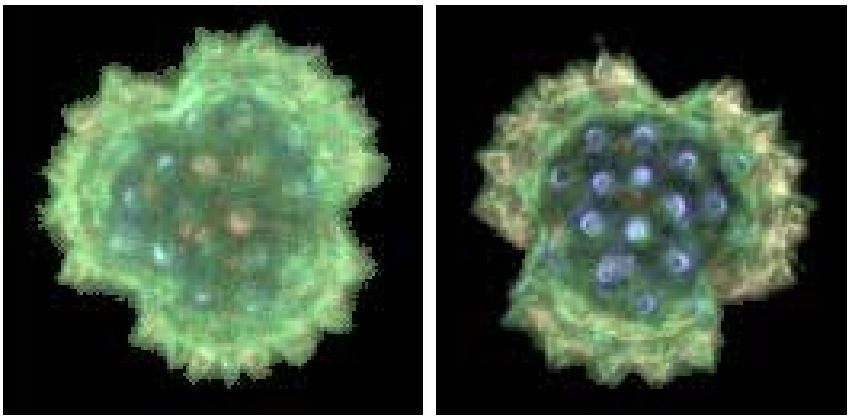


Figure 1: Olearia semidentata **(Family: Asteraceae)**

Formation of Veritaxa Limited

In January of this year (2013), agreement was reached to licence the Classifynder technology from Massey University to our new technology start up company "Veritaxa Limited".

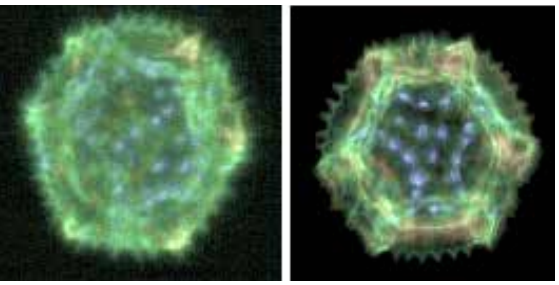


Figure 2: Taraxacum sp. (Family: Asteraceae)

Unusually for a start up company, Veritaxa will hit the ground running with existing sales and systems in use by a number of research groups in Universities around the globe including Government research organisations.

Whilst all the same people are involved as before, the shift to put the Classifynder technology on a commercial footing is a key part of the drive to make a robust, adaptable and well supported commercial system available to the wider palynological community. Veritaxa Limited will supply all new Classifynder systems in the future.

Veritaxa Limited will continue to provide support for all prototype systems sold by Massey University prior to 31 December 2012. Veritaxa expects to develop global partnerships to sell, service and support it systems in the longer term.

New training and support materials

We strongly suggest that if you decide to acquire a Classifynder system you also plan to undertake one of our user training workshops.

On the first day we give instruction on the operation and use of the system and on the second day we work with you on applying it to the tasks of interest to you.

The consistent preparation of slides is one key to successful Classifynder analysis. Standard methods consistent with most conventional acetolysed sample preparation methods are preferred, but we have developed a preference for silicon oil as the mounting medium. We can provide advice on preparation methods with tips and traps for some specific applications.

We have been developing user training videos and have a number now complete that deal with setting up a new system. While the complete set of videos are still a work in progress, if you email us we can direct you to a URL that will allow you to view and upload some of the completed training videos.

We have also developed a handy "Operators Reference Guide" based on the Release 2 software. This guide is designed as a companion for the user; it summarises the purpose, functions and operations available within each tabbed element of the user interface software. The guide also provides set up advice and details optimal PC hardware and software system requirements. Again, existing users are invited to contact us if they would like one and we will email a copy to you.

Release 2 Software

All measurements are now displayed in microns.

The initial Release 1 software presented mixed units of pixels, microns and steps. Users found this to be confusing; for this reason the Release 2 software consistently uses microns for all object measurements and for dimensioning the scaling bar and the target box.

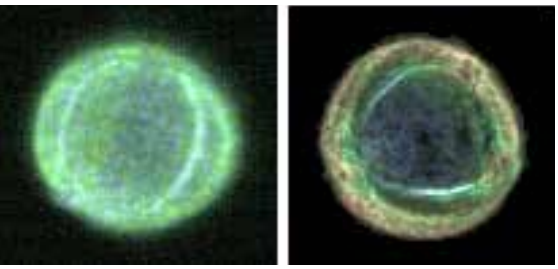


*Figure 3: **Disphyma papillatum** (Family: Aizoaceae)*

The updated software facilitates object recall and tracking.

The system now prompts the user to name each slide to be scanned. The XYZ location information of each object plus the slide name is then embedded in image of each object. This allows complete tracability down to individual objects on the slide. If when reviewing a captured image on screen, the user locates an image of special interest and wants to revisit it, the system will request that the appropriate slide be mounted on the machine. The system will then use the high-resolution camera to relocate and re-display the corresponding object for further detailed examination. This tracability is of value in forensic work and can assist relocation and re-examination of the object using a conventional microscope if required.

Manual selection feature.

Starting from the low resolution "star map" of potential pollen objects, manual object selection is now available. Once selected, objects are then located and images captured using the high-resolution camera. This facility is useful in building up reference sets for classifier training or when the taxa of interest are only present in low numbers.

A calibration wizard has been added.

This facility enables users to readily re-calibrate and establish the precise offset between the high and low-resolution microscopes. The system now also does work in the background to continually monitor the offset and make fine adjustments to slide movement behaviour.

The control screens and tabs have been cleaned up and simplified.

Hovering over a button now produces a summary explanation of what it does. This has been done to make the systems more intuitive and simpler to use.

Suggested morphology values.

Initial morphology settings are automatically calculated based on the sample as scanned These are now provided as a starting point on the morphology tab and are particularly useful in the initial use of the low-resolution microscope where the effects on varying parameters can be seen on the selection/ rejection of pollen from non-pollen objects. These settings can then be fine tuned to a task using the high-resolution microscope.

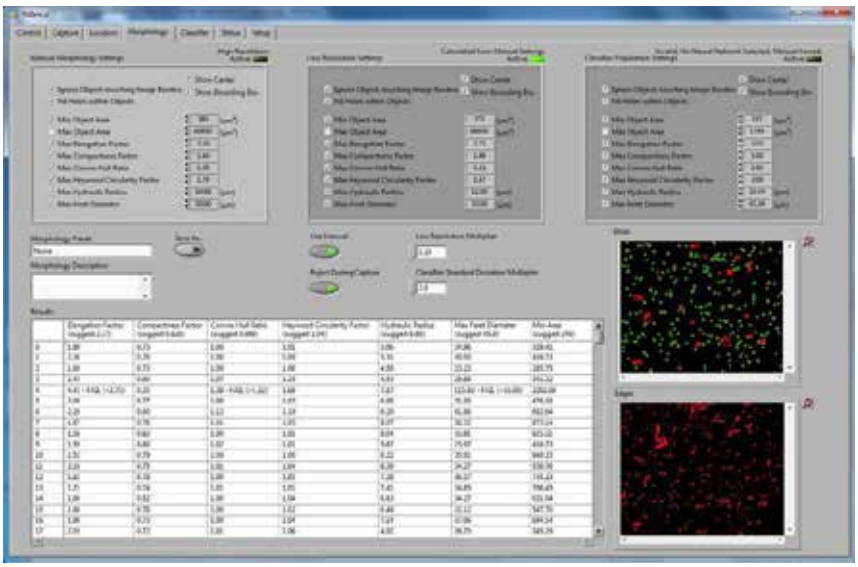


Figure 4: Release 2 Software Interface, showing the morphology tab

A morphology preset facility has been added.

This allows named sets of morphology presets to be stored recalled for subsequent use in particular tasks.

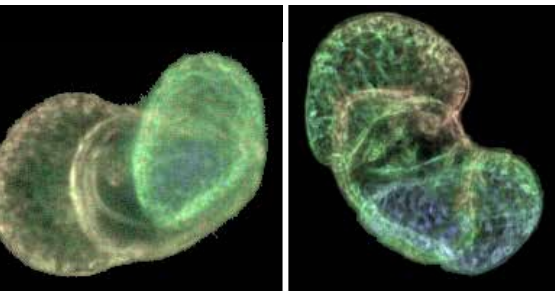


Figure 5: ***Podocarpus sp.*** (Family: Podocarpaceae)

Slide scan completion email generation.

The system software now includes a facility that, on request, will email the user when the current slide scan has been completed. This was a facility requested at the first training workshop.

Classification accuracy with fresh or undamaged acetolysed fossil pollen.

We have observed that the classification accuracy on fresh acetolysed samples depends a lot on the nature of the diversity in the sample; that is where there are two or more taxa that are very easily 'confusable' the accuracy score will be lower. The image quality and the 'representativeness' of the set of library images on which the classifier is trained can also affect the results. This has been an area of active development for us and some of our collaborators, including testing of a number of alternative mathematical classifier approaches.

The Neural Net (NN) classifier algorithm has now been enhanced by the inclusion of a facility that enables confusable pollen taxa to be found and then lumped together in a first round of classification. The confusable taxa are then subsequently discriminated and finally classified as a secondary task.

This approach results in improved classifier performance. Our testing to date indicates that with undamaged acetolysed pollen samples the accuracy of classification varies from 100% correctly classified at between four to six taxa to about 92% at 22 taxa with numerous confusable taxa present in the sample.

We have not completed any tests on samples more diverse that 22 taxa at this point. Our expectation is that the classification accuracy may well be somewhere between 87 to 90 % for the more highly diverse samples. We aim to continually extend the testing range and improve the classification accuracy for more diverse samples.

Accuracy with damaged or deformed fossil pollen:

The Classifynder will recognise and image damaged, broken or slightly degraded grains but its ability to correctly classify them to the correct taxa is currently inferior to its capabilities with undamaged pollen. The system will capture broken gains where the remaining fragment meets the low resolution morphological parameters that have been set by the user but it will also capture clumped pollen/pollen, debris/pollen or debris that also fit within those parameters.

The art is in refining these settings so that an optimal fragmented grain capture rate is obtained without accepting too many non pollen objects or clumped grains. However users are always able to correct the misclassified objects and grains (by dragging and dropping images into the correct folders on screen using their mouse) and when this is done we find that the proportional pollen counts are very close to human counts. Further, the variation in these proportional counts is less than the variation between multiple human counts of the same sample. Please see abstract SS18-006 (193) in:

http://www.psj3.org/ipc-iopc2012/Download_files/SS18.pdf,

Alternatively feel free to contact Dr Kat Holt for more information.

If you want to upgrade to the release 2 software (and we do recommend it) simply email us and we will direct you to a URL were you can download it.

Classifynder generation 3A –

and an upgrade kit for generation 3 systems:

Until mid 2012 all models distributed to users were based on our 3rd generation device. However we have made some improvements to the lighting and optics and this is now the basis of the Generation 3A Classifynder. These improvements include:

The introduction of a x40 objective.

For objects in the normal size range, say larger than 10 micron in diameter, the image produced at the camera is down sampled to give the performance of an enhanced x20 objective. For smaller objects the down sampling can be inhibited allowing the generation of more magnified images. When operated as an enhanced x20, the resultant images are compatible with existing images and libraries captured with Series 3 machines but have more fine detail present.

The effective magnification (from object to screen) on a typical PC set up is 1000x, or 2000x for the objects less than about 10 micron in diameter with down sampling turned off.

Improvements to the darkfield lighting stop and low resolution edge masking arrangements

These adjustments along with corresponding adjustments in the image processing software have resulted in the improved capability of the system to deal with low contrast and small objects.

New lighting layout and power control

Rapid advances in the technology of light emitting diodes (LED's) has enabled the quality and intensity of the lighting to be improved by redesign. This and the improved darkfield noticeably improve the edge definition of objects. Power is also delivered to the LED's on demand, rather than being continually on.

Gen 3 to Gen 3A upgrade kit

We have also produced an upgrade kit that transforms a Generation 3 Classifynder to a Generation 3A which we can fit for about US$4,000. We recommend all current Generation 3 users seriously consider the lighting and optics upgrade; we will be happy to provide a detailed quote including shipping and insurance.

We have installed one such kit and upgraded the software to release 2 on a Gen 3 system for Professor John Birks in Bergen, Norway. John's team report that they consider the upgrade to version 2 software is *"a great improvement being much faster"*. They also report that the Gen 3A lighting and optics upgrade is *"very good; the 40x optics provides better classifications and it makes great savings on electricity use"*. Power must be expensive in Bergen!

In order to gain the advantages of the improvements listed here we recommend all series 3 users consider having their machines upgraded to the series 3A configuration and apply the Release 2 software. However we will continue to support the series 3 machines in all configurations.

Future Technical Developments

New slide holder and release 3 software

To promote use with a wider range of slide configurations and tape based samples from aero samplers such as the 'Burkard' sampler, we plan to progressively change from the present double slide holder to a single slide holder that allows almost all of the area of a conventional 75 mm x 25 mm slide to be scanned. The new slide holder will feature an automated slide ejection system to facilitate future

use with robotic slide loaders. We will also introduce the software modification needed to scan the following area formats

- 10 by 10mm,
- 22 by 22mm
- 40 by 22mm
- 67 by 12mm

These areas will all be selectable within a 67 by 22mm scan area window (areas outside this are needed to hold and handle the slide). The current 'ROI' scan will still be selectable. We are working to design the new slide holder so that we can retrofit these to Gen 3 and Gen 3A upgraded units. We expect to offer the slide holder as a standard feature on all new systems. It is a work in progress at present.

Porting raw parameter data and running alternative mathematical classifier models.

The current software algorithm for grain recognition is based on a multi layer perceptron 'Neural Net' (NN) classifier operating on 43 shape, histogram, moment and texture parameters. The NN classifier is a 'forced classification' system, in that objects are classified into the category that the system considers they are the best match for based on its training from library images. Consequently there is no 'unknown' folder; the unknown taxa are dispersed among the known and correctly classified taxa. However every image is displayed in its classificatory group for the user to check and if necessary correct. This is usually a quick procedure (dragging and dropping images from folder to folder using your mouse).

The current software release provides only the native NN classifier, but we appreciate that some users would like to experiment with unforced classifiers. To support this we are applying new code that will allow users to apply other algorithms using Intel's Open CV library of statistical machine learning tools (refer http://en.wikipedia.org/wiki/OpenCV) These modular tools provide for options such as linear discriminate, decision trees, support vector and random forests classifiers. There is considerable testing to do to ensure this code operates seamlessly so its delivery is slated for a future release of the software.

However we have also developed code that allows users to export raw parameter data to excel spreadsheets that can then be used to test with alternative classifier models or used in statistical analysis programs such as 'R'. This code is available as a stand alone macro now (indeed some users have been using it to test alternative classifiers). This code will be built in and available as part of the release 3 software so that porting raw parameter and metadata data to spreadsheets can be automated.

Continued improvements to the classifier accuracy.

We have a number of potential development pathways for improving the classifier performance; in particular with damaged fossil pollen. These pathways range from enhanced use of 3D image data, use of adaptive classifiers, improving the systems knowledge of the image quality and using stochastic modelling techniques to help improve classification accuracy. These are longer term projects and if you would like to collaborate with us on any of this work we would be delighted to hear from you.

Addition of further meta-data fields:

In addition to the slide name and object location meta-data, release 3 software will provide fields for sample date and geo-location using Google Earth to populate latitude and longitude coordinates.

Removal of Matlab runtime

Over time we plan to complete code development in LabView for all tasks now performed in Matlab's run time modules. This will remove the need for users to have a Matlab licence and/or purchase additional Matlab modules. There is nothing wrong with Matlab, it's just unnecessary overhead in this case.

The introduction of a commercial version (Series 3B) hardware platform.

We have identified design and manufacturing changes that will significantly improve the fine performance and robustness of the digital slide mechanisms compared to those used in the series 3 and 3A machines. We have observed that under adverse environmental conditions, the current x-y-z slides have the potential to stick or cause what is best described as 'micro-juddering'. This can result in poor positioning repeatability and impaired focus from a loss of registration in the final Z-stack

derived images. Tweaking and a solid test and burn-in period has helped to reduce this potential problem prior to shipping, but the time taken to do so during assembly, calibration and testing is considerable.

To eliminate the possibility of such issues in future systems and to make the machines simpler to assemble and calibrate, we anticipate making a change to the type of x-y-z slides we have used to date. We consider that this improvement will give a much more consistent, precise and robust system. The changes will undoubtedly come at a price and we will be working to minimize this as we make our component selections. The 3B will have the same optics and lighting as the 3A plus the new slide holder. We expect to also take the opportunity to make a number of cosmetic enhancements while we are at it.

All current Generation 3 machines that include the Gen3A lighting and optics upgrades will produce compatible image quality, meaning you can use library images from one system on another system.

We are taking orders now for the commercial (Gen 3B) version. Call us for a quote including air courier delivery and insurance.

Applications

The Classifynder system was developed as an intelligent assistant to palynology; that is the counting and classification of pollen and spores on slides. However the system has a myriad of potential applications.

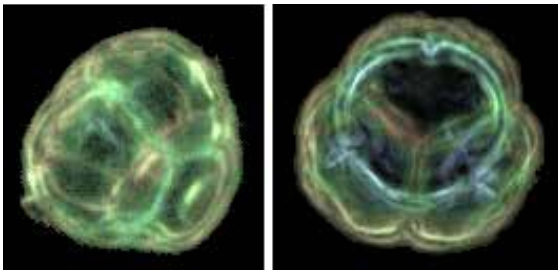


Figure 6: ***Dracophyllum arboretum*** (Family: Epacridaceae)

Possible industrial and scientific applications of the core 'Classifynder' platform include forensics, honey pollen testing, allergy studies, fruit crop pollen viability testing, seed testing, facial eczema monitoring, potato blight monitoring, stratigraphic studies in oil and gas exploration and others.

We are investigating some of these alternative applications and are keen to discuss others with potential users. Call us if you have ideas or wish to collaborate and partner with us in any way.

What are people are doing with their Classifynders?

Over the past six months, Massey University Palynologist **Dr Kat Holt** has been focussing her efforts with the Classifynder in three main areas:

- Improving the performance of the system on identification of Quaternary fossil pollen.
- Trialling the system on classification of honey pollen samples.
- Investigating the limitations of the Classifynder in scanning pollen samples collected in Burkard/Hirst airborne pollen samplers.

At the moment, efforts to improve the performance of the Classifynder with NZ fossil pollen samples revolve around optimising the training sets, as well as trialling other types of classifiers, which has been made possible with collaborative input from **Dr Ryan Lagerstrom** at CSIRO in Australia.

The honey pollen trials are a collaboration with **Dr Ian Raine** of New Zealand's Geological and Nuclear Sciences (GNS) and involve applying the Classifynder to slides of honey pollen which have previously been counted as part of a GNS honey characterisation project. So far, the Classifynder has been able to replicate the percentage pollen counts of samples with low diversity, with the advantage being that much larger counts can be performed. Trials are now progressing to samples of higher diversity.

Kat has been applying the Classifynder to tape samples collected in Burkard/Hirst samplers and is focussing on overcoming the optical issues created by slide preparations which comprise layers of three or more different substances (i.e. melinex tape, adhesive, glycerine jelly/mounting medium) all of which have different optical properties. This, combined with the abundance of fine debris (dust, soot, etc) reduces the effectiveness of the Classifynder in isolating the pollen form the background/non-pollen objects. Fuchsine staining also seems to reduce the visibility of the pollen grains. Experiments with different adhesives, mounting mediums and stains are planned, with the ultimate goal of finding a 'recipe' which works optimally with the Classifynder's dark field lighting.

Lastly, Kat has also successfully applied the Classifynder in a commercial project which involved performing a comparison of pollen concentration and pollen size for two types of kiwifruit (*Actinidia*) pollen. Pollen concentration values (grains per gram) were obtained using the standard procedure of spiking with *Lycopodium* markers, and then having the Classifynder perform the counting. Counts were something in the order of 50,000 grains per pollen type, far more than could ever be achieved by manual counting. Measurements of pollen size made by conventional methods (SEM and TL microscopy) were supplemented with size information derived from the Classifynder images by extracting the area value for each pollen grain imaged, which again provided data on a much greater number of grains than could ever be measured manually.

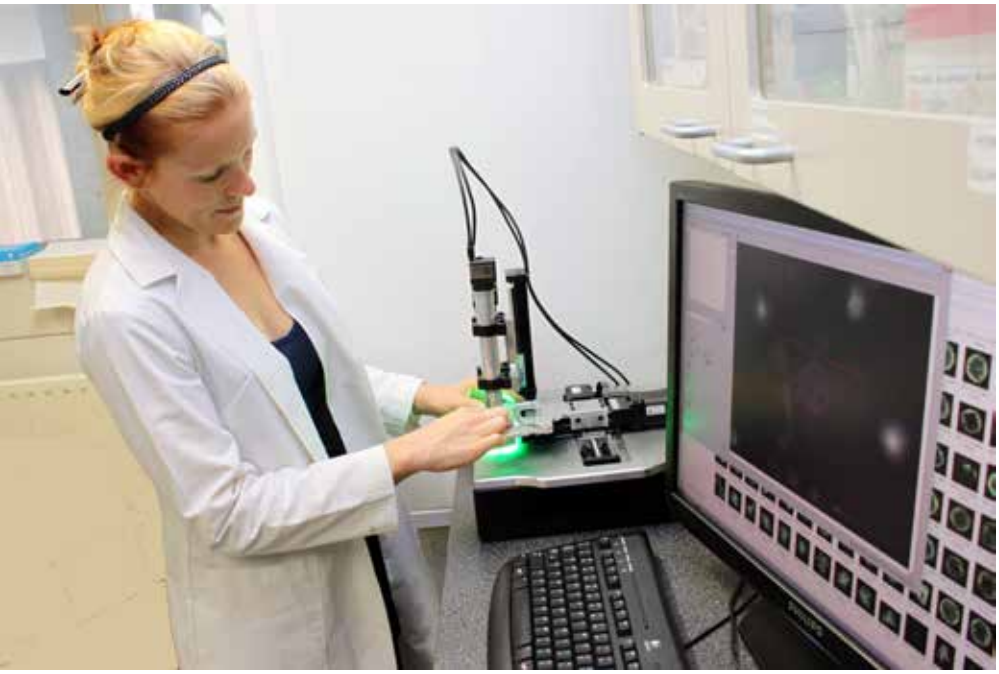


Figure 7: Dr Katherine Holt (Institute of Natural Resources, Massey University, New Zealand) using a Gen 3 Classifynder.

Dr Anna Sandiford at the Forensics Group in Auckland, New Zealand is planning to use a Classifynder trial from July this year; hopefully on a live criminal case and in parallel with conventional microscopy.

Professor Keith Bennett Head of School of Geography, Archaeology and Palaeoecology of Queens University, Belfast, Ireland has a PhD student on a project to develop the Classifynder by identifying and recording pollen taxa images to build a comparable and exchangeable reference library database, to establish protocols to produce consistent, repeatable and acceptable tested results comparable to counts completed by palynologists and to produce results that would be acceptable for publication.

Professor John Birks of the Department of Biology and the Bjerknes Centre for Climate Research at the University of Bergen, Norway has been applying a Classifynder to Norwegian Holocene lake sediments and detailed palaeoecology studies. John considers that the Classifynder has the potential to revolutionise pollen

analysis and to make essential pilot studies fast and simple.

Dr's David Lovell, Beth Mantle, Leanne Bischof, Simon Haberele and Ryan Lagerstrom of CSIRO, Australia have been involved in using the Classifynder in ecological pollination association studies, testing alternative classifiers, onion pollination trials and collaborative efforts in testing fossil pollen with Dr Kat Holt.

Dr Elisabeth Levac, Associate Professor, Dept of Environmental Studies and Geography: Departmental Chair at Bishops University in Quebec has an interest in aeropalynology and like Dr Holt is considering improved tapes and slide mounting methods.

Dr Richard Jones, College of Life and Environmental Sciences at the University of Exeter, in the UK has PhD students working on improving fossil pollen sample preparation. He is also working on potential applications in airborne pollen monitoring and has been providing us with development suggestions in this regard.

What are you doing with your Classifynder? Write to us and tell us about your projects and give us feedback and ideas for improvement.

Contact us

For general enquiries and sources of information on the Classifynder system plus a short and visually appealing video that describes some of the projects that CSIRO is currently undertaking using its Classifynder machine, visit:

www.classifynder.com

Alternatively you can contact us directly:

- Commercial Director: Jamie MacDuff **j.d.macduff@paradise.net.nz**
- Technical Director: Robert Hodgson **rmsbhodgson@inspire.net.nz**

Acknowledgements are due to Dr Katherine Holt for system testing and for identifying the pollen images in this newsletter, and to Emeritus Professor John Flenley for his editorial.