WHAT'S ON THE MENU: *DROSERA ROTUNDIFOLIA* DIET DETERMINATION USING DNA DATA

by

Barbora Lekesyte[1,2], Stephen Kett[2] and Martijn J.T.N. Timmermans[2,3]

[2]Department of Natural Sciences, Middlesex University, The Burroughs, Hendon, London, NW4 4BT

[3]Department of Life Science, Natural History Museum, Cromwell Road, London, SW7 5BD

[1]Corresponding author, e-mail: barbora.lekesyte@gmail.com

ABSTRACT

The round-leaved sundew, *Drosera rotundifolia*, is a carnivorous plant species. On Lundy it is found in the nutrient-poor bog environments of Pondsbury and the northernmost quarry, where it supplements its diet with invertebrate prey. To gain insight into the diet of these two sundew populations a metabarcoding approach was trialled. This is, to our knowledge, the first study to use DNA barcodes to identify *Drosera* prey. At each site, a $0.25m^2$ quadrat was placed in a representative *Drosera* patch and two days' worth of prey were collected. To identify prey items, Cytochrome c oxidase subunit I (COX1) sequences were obtained and compared to the Barcode of Life database. This revealed that Lundy sundews have a mixed diet. In total at least 20 different prey taxa were detected in the two $0.25m^2$ areas sampled. Sixteen taxa could be identified to species, indicating that metabarcoding permits accurate species level identification of sundew prey items. The majority of prey taxa were dipterans (two-winged flies), of which several have previously been reported on Lundy. Most prey taxa were detected in only one of the two quadrats examined (Jaccard's index of Similarity=0.01; 'dissimilar'). This might indicate that the two *Drosera* populations feed on distinct prey communities, but more research is needed to confirm this.

Keywords: *Lundy, carnivorous plants, sundew Drosera, DNA barcodes, prey taxa, Diptera*

INTRODUCTION

Carnivorous plants of the genus *Drosera* (sundews) are typically found in nutrient poor environments (Ellison and Gotelli, 2001). They thrive under these deprived conditions by supplementing their diet with arthropod prey (Millett *et al.*, 2003). Prey are caught and digested with modified leaves ('blades'). Blades possess large numbers of glandular 'hairs' that secrete drops of viscous adhesive solution. When arthropods contact with

these drops they are trapped and die (Adlassnig *et al.*, 2010) (Figure 1). Digestive enzymes produced by the plants then dissolve prey items, releasing nutrients to be absorbed by the plant (Adamec, 2002).

Figure 1: *Drosera rotundifolia* with prey items. © S. Kett

The genus *Drosera* has attracted scientific attention since the eighteenth century, most of which focussed on benefits of prey capture on plant growth and survival (e.g. 'Botany', 1874; Darwin, 1875; Roth, 1782; Thum, 1988; Hooker, 1916). For example, Thum (1988) showed that artificially increased food supply *Drosera* increases dry weight, flower and leaf number and the overall trapping area of individual plants. Equally, plant traits (Foot *et al.*, 2014) and microhabitat (Thum, 1986) have also been shown to affect prey capture efficiency and diet composition.

Investigating natural *Drosera* diet, via morphological identification of prey, however, is often hampered by the rapid digestion of prey tissue. To overcome this difficulty a DNA barcoding approach to identify prey was trialled. DNA barcodes are standardised genetic markers used for taxonomic identification, ideally to species level (Hebert *et al.*, 2003). DNA sequences are obtained from specimens and then compared to sequences from accurately identified and vouchered specimens in a reference database. Matches between 'unknown' DNA sequences and sequences in the database result in a positive identification for specimens of interest.

This study focussed on the round-leaved sundew, *D. rotundifolia* L., Lundy's only carnivorous plant species. Samples were taken from two populations (Figure 2). One population is found at the edge of Lundy's largest pond, Pondsbury (51°10'38"N, 4°40'12"W). Much of the surface vegetation in this area is *Sphagnum* bog with frequent tussocks of *Juncus* sp. The other population is found in the northernmost quarry (51°10'45"N, 4°39'53"W). Here vegetation is characterised by *Sphagnum* and other plants adapted to acid, poorly drained soils.

This study aimed to test whether sundew prey items can be identified to species level using molecular barcoding and to compare obtained identifications to existing Lundy species records.
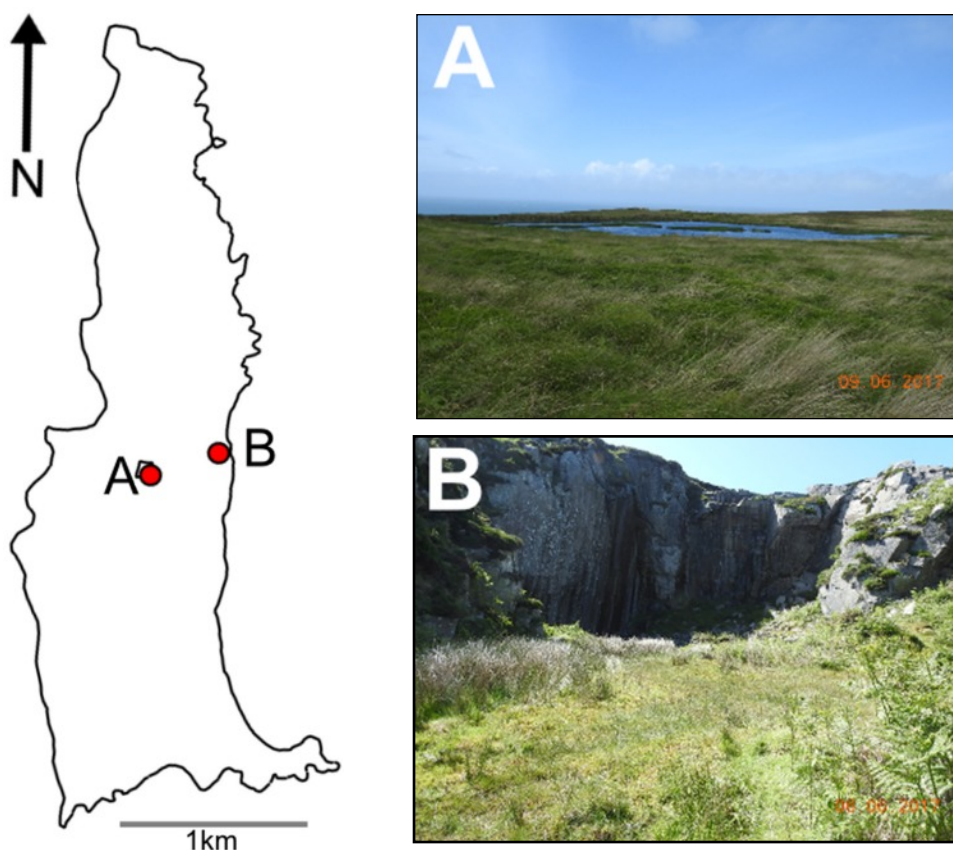
Figure 2: The two Lundy *Drosera rotundifolia* populations sampled.
A) Pondsbury, B) the northernmost quarry. © B. Lekesyte

MATERIAL AND METHODS
Field work methods
In June 2016, 0.5×0.5 m quadrats were established within the Pondsbury and Quarry *D. rotundifolia* populations. Quadrats were placed in locations judged 'typical' of a dense *Drosera* 'patch'. Flags were used to indicate the four quadrat corners to permit relocation of each quadrat. On the first day of the experiment, plant blades were 'cleaned' using forceps to remove all prey items. To determine prey composition, prey were collected two days after cleaning occurred. Collected prey items were stored in tubes of absolute ethanol.

Laboratory methods
For each quadrat, prey samples were pooled in a single tube. DNA extractions were performed on these pooled samples. Ethanol was removed by pipetting. A heating block (56°C) was used to evaporate residual ethanol. DNA extractions used the Blood and Tissue Kit (Qiagen) and followed manufacturer's recommendations, except that double volumes were used for buffer ATL, buffer AL and 100% ethanol. Extracted DNA was subsequently sent to NatureMetrics Ltd for metabarcoding. Metabarcoding followed NatureMetrics Ltd standard procedures. In brief, a short fragment of the cytochrome oxidase c subunit 1 (COI) barcode was amplified using primer Fol-degen-rev 5'-

TANACYTCNGGRTGNCCRAARAAYCA-3' (Yang *et al.* 2012) combined with Leray primer mlCOIintF: 5'-GGWACWGGWTGAACWGTWTAYCCYCC-3' (Leray *et al.* 2013) or combined with primer 'Short2' 5'-CCNGAYATRGCNTTYCCNCG-3' (NatureMetrics Ltd, pers. comm.) (Figure 3). All PCR reactions were performed in triplicate. PCR products were purified and quantified (Qubit high sensitivity kit). PCR products for the same site (quadrat) were pooled and Next Generation Sequencing (NGS) libraries were prepared as specified by Illumina for amplicon sequencing on the Illumina MiSeq System (Illumina Inc. 2013) and sequenced using an Illumina MiSeq 2×300 kit.
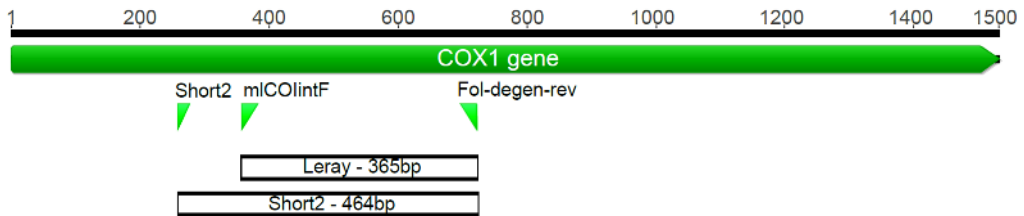

Figure 3: Two fragments of the COX1 gene were amplified using PCR. Primers Short2 and Fol-degen-rev amplify a 365bp fragment. Primers mlCOIintF and Fol-degen-rev a 464bp fragment. Positions of the three primers (green triangles) on the COX1 gene sequence (green bar) are given

Bioinformatic methods
Raw sequencing reads for each site were stitched using PEAR (Zhang *et al.*, 2014) and subsequently split by forward primer sequence using cutadapt (Martin, 2011). This step also trimmed uninformative PCR primer sequences. Low quality sequences were removed using the prinseq-lite Perl script (Schmieder and Edwards, 2011), removing all sequences that contained at least a single 'N', had a single position with a Phred quality below 20 and an average Phred quality below 30. Sequences were then converted to FASTA format using fq2fa (Peng *et al.*, 2012). Operational Taxonomic Units (OTUs) were constructed from these files using the UPARSE pipe-line (Edgar, 2013). Sequences were de-replicated (merging all exact duplicates) and singletons (sequences that were observed once only) were removed. Remaining sequences were clustered at 97% similarity in USEARCH (Edgar, 2010) to generate OTUs and all sequences were subsequently assigned to each of the different OTUs (again at 97% similarity). OTUs with less than 10 sequences for both quadrats combined were discarded. To identify OTUs in the final dataset, sequences were compared to the Barcode of Life (BOLD) database (http://www.boldsystems.org/). Identifications were compared to the Diptera checklist for Lundy (Lane, 1977) and various other sources (Figure 4). Prey taxon approximate sizes were obtained from a variety of generic sources.

Jaccard's Index of Similarity
Jaccard's Index of Similarity was used to determine overall similarity of composition between the two sets of identified prey taxa, from the Pondsbury and the quarry sites. It was applied only where prey taxon presence could be unequivocally determined, e.g. if a genus occurs in both prey sets, it was not possible to determine without species identification whether a species level difference occurred between the two taxa.
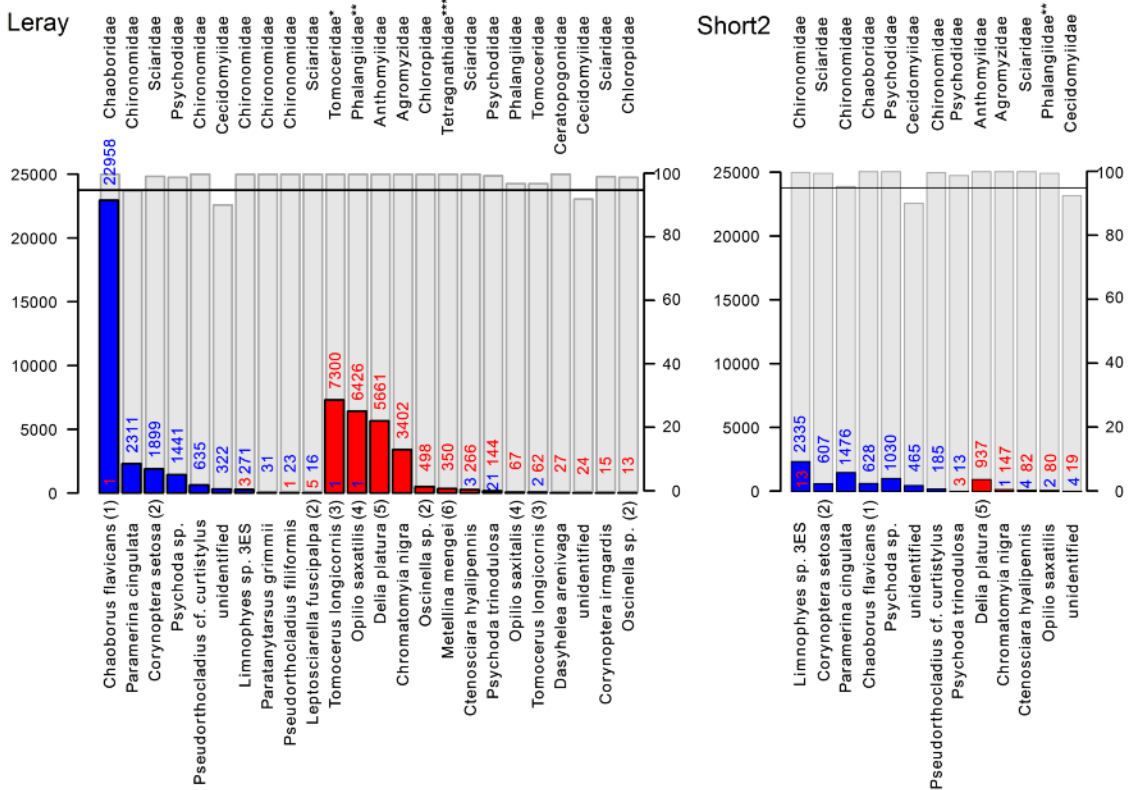
Figure 4: Taxa observed at two sites on Lundy using: Left) Leray (2013) primers and Right) Short2 (NatureMetrics Ltd., pers. comm.) primers. Left vertical axis: Number of sequences observed for a specific taxon. Blue (Pondsbury) and red (Quarry) bars and blue and red numbers on the graph represent number of sequences observed. Right axis: % similarity (grey bar) to a reference sequence in the BOLD database. Horizontal line indicates 97% similarity. Family names are given above graph. All families belong to Diptera, except *) Entomobryomorpha (Collembola), **) Opiliones (Arachnida) and ***) Araneae (Arachnida). Species names (if available) are given below graph. Numbers between parentheses refer to: 1) (Lane, 1978) 2) (Menzel et al., 2006) 3) (Smith and Nunny, 2012) 4) (Parsons, 1988) 5) (Parsons, 1996) 6) srs.britishspiders.org.uk (last accessed 07/07/2017. Note that Tomocerus longicornis, Opilio saxitalis and Oscinella sp. are represented by two Leray OTU

RESULTS

Drosera populations

The Pondsbury quadrat contained 91 *D. rotundifolia* plants and the quarry quadrat 66. Within the two-day sampling period plants in the Pondsbury and quarry quadrats caught totals of 44 and 83 prey items respectively.

Molecular identification of prey items

In total, 234,058 raw paired-end reads were obtained. There were 162,345 paired-end reads for the Pondsbury sample and 71,713 paired-end reads for the quarry sample. Reads were merged and 94.2% of the Pondsbury and 96.8% of the quarry samples could

be combined into single contiguous sequences. These datasets were subsequently split by PCR primer sequence (i.e. Leray or Short2), resulting in two datasets for the quarry sample (Leray: 58866 sequences; Short2: 10222 sequences) and two datasets for the Pondsbury sample (Leray: 79316 sequences; Short2: 73110 sequences). Sequences were clustered per primer pair to construct OTUs. OTUs with 10 or more associated sequences are given in Figure 4.

Figure 4 shows that several OTUs are represented by a large number of sequences in one sample and by a very small number in the other (e.g. *Tomocerus longicornis*: 7300 sequences in quarry sample, 1 sequence in Pondsbury sample). It is currently unclear whether such very low numbers are a genuine reflection of the presence of a species at the respective quadrat, or whether they are merely a result of cross-contamination. Cross-contamination between samples is a well-known risk when working with environmental DNA (Thomsen and Willerslev, 2015). To avoid overestimation of the number of prey species at each site, such cases were treated as cross-contamination.

Identifications were obtained using the Barcode of Life Database (http://www.boldsystems.org) search engine. This revealed that all Pondsbury OTUs are dipterans (most from the family Chironomidae). Dipterans were also most abundant in the quarry datasets, but two species of Arachnida and one collembolan species were also observed.

With the Leray primer greater diversity was revealed than with the Short2 primer (Figure 4). All OTUs observed using the Short2 primer, are also observed using the Leray primer, but not *vice versa*. The Leray primer revealed three more species for Pondsbury and six more species for the quarry.

At least 12 different arthropod families were detected as prey items. Almost all of these families consist of relatively 'small' species (≤5mm) and the majority have an adult flying stage (Table 1). Three of the 12 families are associated specifically with aquatic/bog environments and two with 'decaying/rotting organic matter' (Table 1). It must be noted that several non-arthropod OTUs were observed that are not shown in Figure 4. These included a nematode, two fungi and bacteria (Rickettsiales).

There was a clear distinction between prey items taken by the two *Drosera* populations (Jaccard's index of Similarity=0.01; 'dissimilar'), with most (19/21) prey taxa found in only one of the two quadrats examined.

DISCUSSION

This trial suggests metabarcoding permits accurate species level identification of *Drosera* prey items. Reliable identifications (>97% sequence similarity with a BOLD database entry) were obtained for a total of 16 species. Some OTUs could not be identified to species level, because they are currently not represented in the BOLD database. With an ever-growing BOLD database, however, the number of such unidentifiable taxa is likely to decrease over time.

Results indicate that Lundy *Drosera* have an eclectic diet. Analysis of just two days' worth of prey reveals that at least 10 different arthropod species were caught at each 0.5×0.5 m quadrat. A large proportion of these prey species have been reported to occur on the island (Figure 4). However, for at least some of the species this might be the first Lundy record. This suggests that metabarcoding of *Drosera* prey could be a reliable and

Table 1: Arthropod families detected using the metabarcoding approach.
Flying: taxa with flying adult stage. Length data approximate and from a variety of
generic sources

Family	Comments	Flying	Diptera	Length (mm)
Chaoboridae	Chaoboridae: non-biting, mosquito-like, larvae aquatic and predatory upon zooplankton	Y	Y	≤10
Chironomidae	Chironomidae: non-biting midges; larvae in water/wet habitat or decaying matter	Y	Y	≤10
Sciaridae	Sciaridae breed in all sorts of rotting matter and fungi	Y	Y	≤7
Psychodidae	Psychodidae:owl midges, swarm over breed sites, larvae feed on decaying matter, esp. in water	Y	Y	≤3
Cecidomyiidae	Cecidomyiidae: gall midges, 600+ spp, larvae not all in galls, some eat aphids, for example	Y	Y	≤5
Tomoceridae	Collembola with long antennae	N	N	≤5
Phalangiidae	Harvestman – feed on small invertebrates (perhaps stealing from *Drosera*?)	N	N	≤5
Anthomyiidae	Anthomyiidae: flies; larvae feed on decaying matter	Y	Y	≤3
Agromyzidae	Agromyzidae: like miniature houseflies, larvae=leaf and stem-miners	Y	Y	≤6.5
Chloropidae	Chloropidae: small to minute flies; larvae mainly plant feeders, *Oscinella*=frit-fly ~1.5mm	Y	Y	1.5
Tetragnathidae	Spiders with elongated body	N	N	2-23
Ceratopogonidae	Ceratopogonidae: tiny biting midges, larvae live in water/swamp often with much organic matter	Y	Y	≤5

(relatively) non-invasive technique for community analyses and assessment of taxon presence. It certainly offers significant advantages of accuracy and ease compared to species level identification of semi-digested arthropod fragments via more traditional, morphological techniques.

Dipterans (two-winged flies) dominate the prey samples. Nine of the twelve observed families belong to this Order. This confirms that dipterans form a large proportion of sundew prey, suggesting that the Order forms an important source of nutrients for Lundy *D. rotundifolia*. Similar overrepresentations of dipterans are reported by Ellison and Gotelli (2001) and Foot *et al*. (2014). The latter study investigated the attraction efficiency of *D. rotundifolia* blades and reported that 57% of captured prey items were dipterans. The authors point out that Diptera are most likely not actively attracted (or deterred) by red *D. rotundifolia* blades because Diptera lack red receptors (Foot *et al*., 2014). Thus, the apparent dominance of dipteran prey might simply reflect their relative abundance in boggy environments.

Our analyses revealed low overlap of prey taxa between the two sites. This could indicate the local invertebrate communities to be very different too, possibly as a result of dissimilar microenvironmental conditions. Pondsbury and North Quarry are, after all, very different environments in that Pondsbury is open and exposed to both sun and wind whilst North Quarry is sheltered from the prevailing wind and receives less

sunlight (Figure 2). Even so, although intriguing, the observed prey taxon dissimilarity may simply represent an artefact of only sampling one time-point and the high dipteran diversity on Lundy. More detailed investigations (including replication of quadrats over space and time) are needed to determine whether the two *Drosera* populations feed on genuinely distinct arthropod communities.

Most prey items were heavily degraded. It is probable that this was reflected at the DNA level, with 'older' prey items containing DNA of lesser quality than 'newly' captured ones. Differences in prey DNA quality might have introduced bias, with 'older' prey being underrepresented or even missing from the final sequencing dataset. Metabarcoding is prone to other biases, including relative specimen size (larger specimens contribute more DNA to the pool than smaller ones) and primer efficiency (some specimens in the pool will PCR amplify better than others). Because of such biases the sequence numbers given in Figure 4 cannot be extrapolated to biomass or number of individuals caught. They merely indicate that a taxon is present on the island and in the habitat sampled.

It is also important to note that metabarcoding is an indirect method of community assessment. It is well known that cross-contamination among samples can introduce noise and the method therefore never provides 100% certainty that a species was present, even when uttermost care is taken. Such noise can lead to the incorrect conclusion that a species inhabited a site, whereas it actually did not. This type of error is more likely for species detected with low numbers of sequences. In typical metabarcoding experiments, dozens of samples are run in parallel. To minimise the chance of falsely concluding any of the species reported here stems from contamination from samples run in parallel, sequencing reads were compared among all samples processed in the same sequencing run (NatureMetrics Ltd, pers. comm.). Based on this comparison it seems unlikely that any of the species was absent from the site for which we report it.

This study opens opportunities for future research. Accurate identification of prey to species level will permit high resolution analysis of environmental effects upon *Drosera* diet. Questions that might be posed include: Do diets of *Drosera* populations inhabiting contrasting habitats differ? Do *Drosera* prey taxa vary according to season or even depending on whether it is night or day? Are some arthropod species more attracted by the plants than others and if so, why?

These are some of the questions that might be addressed using Lundy *Drosera* populations as a model system. Whatever the answers may be there is no doubt that these predatory little plants will both invite and repay research for many years to come.

REFERENCES

Adamec, L., 2002. Leaf absorption of mineral nutrients in carnivorous plants stimulates root nutrient uptake. *New Phytologist.* 155, 89-100. doi:10.1046/j.1469-8137.2002.00441.x

Adlassnig, W., Lendl, T., Peroutka, M., Lang, I., 2010. Deadly Glue – Adhesive Traps of Carnivorous Plants, in: von Byern, J., Grunwald, I. (Eds.), *Biological Adhesive Systems: From Nature to Technical and Medical Application*. Springer Vienna, Vienna, pp. 15-28

Botany, 1874. *American Naturalist*. 8, 55-57. doi:10.1086/271258

Darwin, C., 1875. *Insectivorous plants*, first edition. ed. John Murray, United Kingdom

Edgar, R.C., 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Natural. Methods* 10, 996-998. doi:10.1038/nmeth.2604

Edgar, R.C., 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics,* 26, 2460-2461. doi:10.1093/bioinformatics/btq461

Ellison, A.M., Gotelli, N.J., 2001. Evolutionary ecology of carnivorous plants. *Trends in Ecology & Evolution*. 16, 623-629. doi:10.1016/S0169-5347(01)02269-8

Foot, G., Rice, S.P., Millett, J., 2014. Red trap colour of the carnivorous plant *Drosera rotundifolia* does not serve a prey attraction or camouflage function. *Biological Letters*. 10

Hebert, P.D.N., Cywinska, A., Ball, S.L., deWaard, J.R., 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences,* 270, 313-321

Hooker, H.D., 1916. Physiological Observations on *Drosera rotundifolia*. *Bulletin of the Torrey Botanical Club*, 43, 1-27. doi:10.2307/2479710

Lane, R., 1978. The Diptera (two-winged flies) of Lundy island. *Annual Report of the Lundy Field Society 1977*, 28,15-31

Leray, M., Yang, J.Y., Meyer, C.P., Mills, S.C., Agudelo, N., Ranwez, V., Boehm, J.T., Machida, R.J., 2013. A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. *Frontiers in Zoology*, 10, 34. doi:10.1186/1742-9994-10-34

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal*, 17, No 1 Gener. Seq. Data Anal.

Menzel, F., Smith, J.E., Chandler, P.J., 2006. The sciarid fauna of the British Isles (Diptera: Sciaridae), including descriptions of six new species. *Zoological Journal of the Linnean Society,* 146, 1-147. doi:10.1111/j.1096-3642.2006.00190.x

Millett, J., Jones, R.I., Waldron, S., 2003. The contribution of insect prey to the total nitrogen content of sundews (*Drosera* spp.) determined in situ by stable isotope analysis. *New Phytologist*. 158, 527-534. doi:10.1046/j.1469-8137.2003.00763.x

Parsons, A.J., 1996. Invertebrates (excluding Lepidoptera) on Lundy 1995. *Annual Report of the Lundy Field Society 1995*, 46, 87-91

Parsons, A.J., 1988. Invertebrates on Lundy, 1987. *Annual Report of the Lundy Field Society 1987*, 38, 52-53

Peng, Y., Leung, H.C.M., Yiu, S.M., Chin, F.Y.L., 2012. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics*. 28, 1420-1428. doi:10.1093/bioinformatics/bts174

Roth, A.W., 1782. Von der Reizbarkeit der Blatter des sogenannten Sonnenthaues (*Drosera rotundifolia, longifolia*), *Beytrage zur Botanik, Erster theil*.

Schmieder, R., Edwards, R., 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* ,27, 863-864. doi:10.1093/bioinformatics/btr026

Smith, P., Nunny, R., 2012. Mapping of sedimentary marine biotopes around Lundy, UK. *Journal of the Lundy Field Society*. 3, 41-74

Thomsen, P.F., Willerslev, E., 2015. Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. Special Issue Article: Environmental DNA, *Biological Conservation*, 183, 4-18. doi:10.1016/j.biocon.2014.11.019

Thum, M., 1988. The significance of carnivory for the fitness of *Drosera* in its natural habitat.1. The reactions of *Drosera intermedia* and *Drosera rotundifolia*. *Oecologia,* 75, 472-480. doi:10.1007/BF00376954

Thum, M., 1986. Segregation of habitat and prey in two sympatric carnivorous plant species, *Drosera rotundifolia* and *Drosera intermedia*. *Oecologia*, 70, 601-605. doi:10.1007/BF00379912

Yang, C., Wang, X., Miller, J.A., de Blécourt, M., Ji, Y., Yang, C., Harrison, R.D., Yu, D.W., 2014. Using metabarcoding to ask if easily collected soil and leaf-litter samples can be used as a general biodiversity indicator. *Ecological Indicators,* 46, 379-389

Zhang, J., Kobert, K., Flouri, T., Stamatakis, A., 2014. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics,* 30, 614-620. doi:10.1093/bioinformatics/btt593