Suspicious specimens: a new tool to find potentially misidentified and misnamed specimens in biological data using a case study of bryophytes

Sophie A. Roberts[1*], George R. L. Greiff[2], Katherine Slade[3], Nathan Smith[3].

[1]Edge Hill University, St Helens Road, Ormskirk, L39 4QP*

[2]University of Bristol, 24 Tyndall Avenue, Bristol, BS8 1TQ

[3]Department of Natural Sciences, Amgueddfa Cymru - National Museum Wales, Cathays Park, Cardiff, CF10 3NP

**Email for correspondence: sophiearoberts3@gmail.com*

Received: 5th Dec 2023

Accepted: 8th Dec 2024

Citation: Roberts, S. A., Greiff, G. R. L., Slade, K., and Smith, N. 2025. Suspicious specimens: a new tool to find potentially misidentified and misnamed specimens in biological data using a case study of bryophytes. *Journal of Natural Science Collections.* **13.** pp. 21-41.

Abstract

Natural history collections contain a vast quantity of biological data that provide information on past populations, the impact of invasive species or diseases, evolutionary changes, as well as the effects of climate change. Specimens which are misidentified or misnamed will produce problems for researchers, however checking identifications in large datasets is time-consuming. The new tool described here can be used to screen collection data using three analyses to generate a list of specimens that are likely to be misidentified or misnamed – termed 'suspicious specimens', flagging them for curation. The package identifies outlying biological specimens whose metadata indicates a higher risk of misidentification as well as comparing the collection dataset with a reference dataset and flagging up discrepancies. It is free to use and can be adapted for any collection of biological data. This study uses data from bryophyte specimens in National Museum Wales (NMW) and British Bryological Society (BBSUK) herbaria as a case study to demonstrate the functionality of the package. Of the 10 most suspicious species produced by the analysis and examined in this case study, 70% of the species required redeterminations, showing the effectiveness of this tool in improving the accuracy of collection records.

Keywords: botanical specimens, collections, data, mis-identification, curation, tool

Introduction

Natural history collections are an important source of information. The specimens contained vary across broad temporal and geographic ranges and often include rare and extinct species. This wealth of information has been used in a wide variety of ways by researchers to model past populations and evolutionary changes, and show responses to climate change (Andrew *et al*., 2019; Lang *et al*., 2019), past epidemics (Bieker and Martin, 2018), analysis of invasive species (Ivison *et al*., 2023) and changes in biodiversity of habitats (Mannino *et al*., 2020). These collections have also been used to detect when a new species has been introduced to an area as well as to predict species distributions (Mannino *et al*., 2020). These analyses are dependent on the accurate identification of specimens.

However, several studies have highlighted that misidentified or misnamed specimens are a consistent presence in herbarium collections. Older specimens may have information missing or be incorrectly transcribed (Mannino *et al*., 2020) as well as being named using old or contradictory taxonomic concepts (Xu *et al*., 2015). Furthermore, some misidentified specimens have been found to be unidentified species (Olds *et al*., 2023) whilst other misidentifications have been found at the genus level (Bradshaw *et al*., 2022). Misidentification also extends to voucher specimens (Łuczaj, 2010). This is particularly important as voucher specimens are used as a verifiable record of a species cited directly in scientific studies and can help resolve taxonomic issues (Bieker and Martin, 2018), thus misidentified voucher specimens are likely to propagate misidentifications in future specimen records. In a study on 4,500 specimens of African gingers (Goodwin *et al*., 2015), it was found that 58% of the specimens were misnamed. Misidentifications do not need to only be tracked once in museum collections but in the field observations as well where plant specimens are misidentified at both species and genus level (5.9% and 1.9% respectively; Scott and Hallam, 2003). Other misidentifications in the field can for example lead to invasive species such as the algae *Lophocladia lallemandii* (Montagne) F.Schmitz in the Mediterranean Sea being mismanaged, with impacts on the native ecosystem (Golo *et al*., 2023). Bias from collectors in areas where there is little interest can also produce taxonomic errors (Isaac and Pocock, 2015), and in some cases misleading and false species information can even be recorded (Pearman and Walker, 2004). Such misidentifications in the field can find their way into museum collections but could be caught beforehand.

Bryophytes are an understudied group of plants (Smith, 2020) that can be difficult to identify with some species requiring microscopy to distinguish them from others and yet they have a great abundance in the UK with about two thirds of European species existing here (Atherton *et al*., 2010). They are used as an example for this study as it is likely that the bryophyte specimens reviewed will include misidentifications and such errors are detailed here as an example case study. In this paper, an analysis of around 100,000 bryophyte specimen records from England, Wales, Scotland, and the Isle of Man – consisting of the databased portion of the NMW and the BBSUK herbaria for these regions (Thiers, accessed 2023) – has been conducted using a newly developed R Package created for this study by the authors (Roberts, 2023). Both NMW (National Museum Wales) and BBSUK (British Bryological Society UK) herbaria are held at National Museum Cardiff by Amgueddfa Cymru-Museum Wales.

An R Package is a piece of software created using the statistical coding language R and can be easily downloaded and used by anyone. These specimens are those which use the Watsonian vice county numbers 1 – 112 (Watson, 1847). Vice counties are a convenient way to section areas of Britain and Ireland for comparative analysis including historical and modern material and is still used by the BBS recording system driven by local as well as taxonomic expertise. Northern Ireland and Ireland were not included in this study due to the different vice county system requiring additional coding. The package identifies outlying biological specimens whose metadata indicates a higher risk of misidentification as well as comparing the collection dataset with a reference dataset and flagging up discrepancies. This new tool is a free and time saving method for cleaning data that can work alongside a variety of Collection Management Systems, providing the curator with an accessible method for verifying collection data with different historic data entry practices.

Analysing the distribution of locations, collections, and taxonomic species, produced 61 museum specimens that may require data verification as well as taxonomic reassessment and shows that the published distributions of some species differ substantially from the narrative offered by museum collections.

Materials and Methods

1. The NMW and BBSUK herbaria

National Museum Cardiff is part of Amgueddfa Cymru – Museum Wales and was founded in 1905 with art, geology, zoology, and botany collections currently in the museum. The total botanical collection has around 750,000 specimens including the bryophyte collection consisting of around 308,000 specimens with collections dating back to the 18th century (K. Slade, pers. comm.).

The British Bryological Society (BBS) was inaugurated in 1923, replacing the Moss Exchange Club formed in 1896 (Foster, 1979). Many of the private collections formed during this time are still part of the BBSUK herbarium which has been held at National Museum Cardiff since 1971 (Harrison, 1980). In 2001, the ownership and copyright of the BBSUK herbarium transferred to Amgueddfa Cymru (Cleal *et al*., 2022). The society compiles reliable records of bryophytes and their distributions published in census catalogues, the most recent from 2021 (Blockeel *et al*., 2021b), with an interim census released online in 2023 (Pilkington and Hodgetts, 2023) and the *Atlas of British and Irish Bryophytes* (Blockeel *et al*., 2014).

Being a voucher specimen collection, the greater accuracy of identifications allows for active research with additional voucher specimens being added frequently. The current collection houses around 47,000 specimens.

2. Analyses

The R package created as a tool for identifying outlying biological specimens conducts three separate analyses to determine specimens or observations with a high risk of misidentification due to inaccurate data verification and validation (Roberts, 2023).

2.a. Species Distributions

The first of the three analyses uses species distribution (including varieties and subspecies) from Watsonian vice county (Watson, 1847) data and compares that to published species distribution records. The biological census data for bryophytes – the British Bryological Society Census Catalogue (Blockeel *et al*., 2021b) – uses vice county records and can be used to show the known distributions of species.

Other mapping tools that use distribution datasets utilise specimen coordinates such as ModestR (García-Roselló *et al*., 2013) and DIVA-GIS (Hijmans *et al*., 2001). However, coordinates are not always available especially for older specimens and cannot be reliably retrospectively assigned. Whilst difficulties exist in assigning a vice county to specimens, particularly for specimens found on borders or for those labelled with old place names, it is nevertheless viable and has been generally carried out as standard curatorial procedure at Amgueddfa Cymru when adding specimens to the botany collection.

The R package displays the species distribution from biological specimen data onto Watsonian vice county boundaries GIS layers from the Biological Records Centre (Biological Records Centre, 2019). Specimens from Northern Ireland and Ireland, whilst available and databased, were not included in the analysis due to using a different system for vice counties (Praeger, 1896) requiring additional coding.

In this analysis, a threshold of the number of specimens in a vice county for a species is set. For example, in a smaller dataset, only one specimen found in a vice county could be suspicious as it is an anomalous result compared to the rest of the dataset. For larger datasets, the user may wish to set a higher threshold. The package produces maps for both the species distribution created from the specimen data, and census data distribution. Another tool in this analysis produces a list of specimens where the vice county it is found in is different to that of the census data.

2.b. Collectors

For the second of the three analyses, the number of collectors for a species was analysed to find any potential bias in the collection data. This analysis uses all collectors for every specimen to produce a list of collectors for each species. A threshold is set for the number of collectors that equates as being potentially suspicious. For example, if the threshold is set at one, then a list of species across all specimens with only one collector is produced. A low number of collectors is more likely to show collector bias and potential species misidentification.

It is also important to note that some taxonomic groups may only have a small number of collectors or recorders across the world. When interpreting the results of this analysis it is essential to be aware of the popularity and recording effort going into a group.

2.c. Orphan Species and Specimens

Finally, the program considered orphan species and specimens. An orphan species is one where there is only one species in the database for a given genus. Similarly, an orphan specimen is one where there is only one specimen in the database for a given species. This analysis identified genera or species with the specified number of orphan species or specimens. For example, if the threshold has been set at one, then a list of either genera with one species or species with one specimen will be produced. This method is useful for finding records of rare or under collected/ observed species as well as taxa that have been subject to excessive taxonomic splitting. Where data contains orphan species and specimens, different systems of classification could have been used. Such confusion of classification can lead to problems with identification (Christenhusz and Chase, 2018).

These methods in combination will flag up species that have either suspect distributions, biased collectors, or lack of specimen information. Suspect specimens after analysis can then be checked for their correct identification and then if relevant, sent for further verification to be recorded as new vice county records for a species. The R package can be utilised with any list of specimens that utilises Watsonian vice counties (numbered 1 – 112) for location data and can be compared with any corresponding census data. Thus, the package can be a useful tool in reviewing a broad range of biological datasets.

3. Data verification and taxonomic reassessment

After conducting the three analyses, the results were combined to produce a list of the species which have specimens most likely to be misidentified or misnamed. From this, ten species named as the most suspicious species were selected and the corresponding specimens reviewed and inspected microscopically. For each specimen, the herbarium labels were inspected for original identifications and further information about the specimen. For specimens that required taxonomic reassessment, small sections of the specimen were removed and observed microscopically using *The Moss Flora of Britain and Ireland* (Smith, 2004) and *The Liverwort Flora of the British Isles* (Paton, 1999) for species identification.

Results

NMW and BBSUK Herbaria

From each analysis, suspicious specimens were produced using set thresholds. For the NMW and BBSUK dataset the thresholds for vice county distributions were species which for any vice county had one specimen. When comparing the distribution maps, the species that differed significantly from the census data (i.e., specimens not found in vice counties adjacent to those in the census data, see Figure 2) were considered suspicious. Species with one collector were also considered suspicious and either species with one specimen or genera with one species were also deemed suspicious.

Once all three analyses had been run, the list of species was filtered to only show species that had specimens that qualified as suspicious for all three analyses. For example, having a distribution different to that of the census data, having one collector and being an orphan specimen. This produced a list of the most suspicious species having specimens potentially misidentified. The filtering was then run again for species that only qualified for two of the analyses and so on to produce a ranked list of specimens by suspiciousness (Table 1). Note that while taxonomic names on the database are currently being manually updated using Blockeel *et al.* (2021), Tropicos.org (accessed, 2023) and the United Kingdom Species Inventory (Raper 2014, last updated 12/02/2021), they may differ from currently accepted names (Katherine Slade pers.comm.).

The bryophyte specimens deemed most suspicious are listed subsequently (with accession numbers in brackets). *Neckera pennata* Hedw. (NMW C96.7.333) had a vice county vastly different to that of the census data and being an orphan specimen. *Pseudocampylium radicale* (P.Beauv.) Vanderpoorten (NMW C.2010.030.8020), *Aongstroemia longipes* (Sommerf.) Bruch & Schimp. (NMW C96.16.259), *Heterocladiella dimorpha* (Brid.) Ignatov & Fedosov (NMW C.2000.002.528), *Homomallium incurvatum* (Brid.) Loeske (NMW C96.18.127) and *Paraleucobryum longifolium* (Hedw.) Loeske (NMW C97.12.161) had a vice county vastly different to that of the census data and being an orphan species. *Philonotis tomentella* Molendo (NMW 13.68.49, 15.54.1, 20.7.m.10, 20.7.m.11, 20.7.m.12, 22.187d.977, 23.92.685, 24.457.44, 24.457.45, 25.152.4046, 25.152.4047, 25.152.4050, 25.152.4068, 40.443.46, 42.13.4, 44.265.8, 48.29.48, 64.97.488, 66.230.104, 71.1B.122) and *Riccia crystallina* L. emend Raddi (NMW C96.15.130, C96.15.2959, C96.15.2961, C96.15.2962, C96.15.2963, C96.15.2964, C96.15.2965, C96.15.2966, C96.15.2967, C96.15.2971, C96.15.2972, C96.15.2973, C96.15.2974, C96.15.2975, C96.15.2976, C96.15.2977, C96.15.2978, C97.3.1682, C97.3.1687, C97.3.1688, C.1999.028.3603, C.1999.028.3616, C.1999.028.3619, C.1999.028.3942, C.1999.028.3943, C.1999.028.3944, C.1999.028.3945, C.1999.028.3946, C.1999.028.3947, C.2000.008.186) and (BBSUK C.2001.020.8617, C.2001.020.8618, C.2001.020.8619) which had many different vice counties that were different to the census, ranking it highly as there were many specimens for this species that were found in unexpected locations. *Cirriphyllum cirrosum* (Schwaegr.) Grout (NMW C.2000.002.641) had a vice county vastly different from the census. *Plagiothecium platyphyllum* Moenk. (NMW C.2000.020.28) had a vice county different to that of the census. Of these specimens, most were from the NMW herbarium (58 specimens) and only three specimens were from the BBSUK herbarium (C.2001.020.8617, C.2001.020.8618, C.2001.020.8619), reflecting their respective levels of verification.

These top species flagged for curation were then checked against the literature and analysed microscopically to confirm if species required taxonomic reassessment.

Table 1: The ranking of bryophyte species based on the suspiciousness of specimens after running the three analyses. Species ranked from most suspicious to least suspicious based on outcome of analyses. Species names are those listed in Amgueddfa Cymru-Museum Wales Botany Collections Management System database in June 2023. The reason column dictates which analyses produced suspicious results.

Ranking	Species	Reason
Most suspicious	*Neckera pennata* Hedw.	Vice county vastly different to census Orphan specimen
	Pseudocampylium radicale (P. Beauv.) Vanderpoorten	Vice county vastly different to census Orphan species
	Philonotis tomentella Molendo *Riccia crystallina* L. emend Raddi	Many vice counties different to census
	Cirriphyllum cirrosum (Schwaegr.) Grout	Vice county vastly different to census
	Aongstroemia longipes (Sommerf.) Bruch & Schimp. *Heterocladiella dimorpha* (Brid.) Ignatov & Fedosov *Homomallium incurvatum* (Brid.) Loeske *Paraleucobryum longifolium* (Hedw.) Loeske	Vice county different to census Orphan species
Least suspicious	*Plagiothecium platyphyllum* Moenk.	Vice county different to census

1. *Neckera pennata*

(NMW C96.7.333)

The most suspicious of the moss species was *Neckera pennata* (NMW C96.7.333), which has only one specimen in the collection found in a vice county different to that of the census (Figure 1). *Neckera pennata* is a circumpolar boreal-montane species which has only been recorded once in Scotland in 1823 (Blockeel *et al*., 2014). The flagged specimen was found in VC 9 (Dorset).

When this specimen was observed under the microscope, the leaves were noted to be distinctly smooth rather than undulate (Figure 2). Undulated leaves are a feature in *N. pennata* and other *Neckera* species but not in *Neckera complanata* (Hedw.) Huebener. The specimen showed broad oblong leaves with obtuse apiculate apex and did not have a nerve present (Figure 2) This leaf shape is not like that of *N. pennata* whose leaf gradually tapers to an apex (Smith, 2004). The elongated mid-leaf cells were around 3 – 4 times as long as wide (Figure 2) whereas in *N. pennata* they are 4 – 8 times as long as wide (Smith, 2004). These characteristics, in particular the lack of undulations, points towards this specimen being *Neckera complanata*, the distribution of which includes VC 9, where this specimen was found (Figure 3; Blockeel *et al*., 2014).

Figure 2. Microscope image of (NMW C96.7.333). Image on the left shows the whole leaf missing undulations and nerve with obtuse apiculate apex. Image on the right shows the elongated mid-leaf cells that are 3 – 4 times long as wide. This description is closer to that of Neckera complanata.

Figure 1. The vice county data for Neckera pennata *NMW C96.7.333. The image on the left shows the distribution from the herbarium data: VC 9. The image on the right shows the British Bryological Society 2021 Census Catalogue (Blockeel* et al., *2021b) distribution: VC 90.*

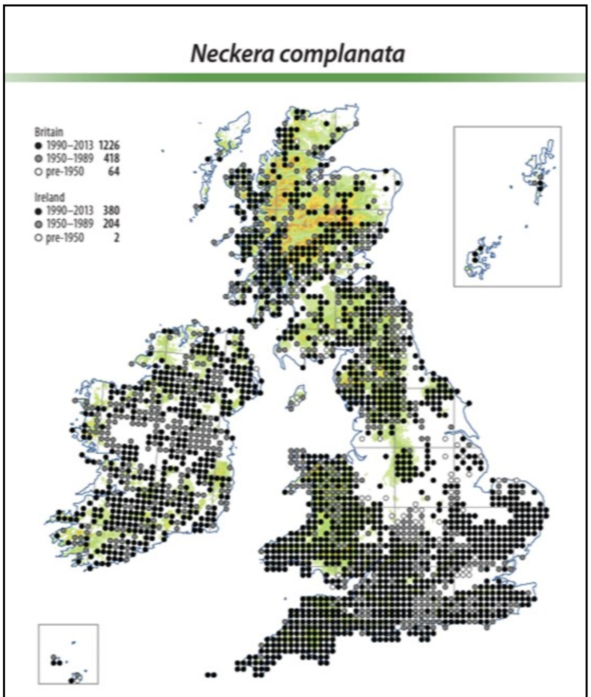


Figure 3. The distribution of Neckera complanata *from the Atlas of British Bryophytes (Blockeel* et al., *2014). This species has a large distribution and includes Dorset where the specimen NMW C96.7.333 was found.*

2. *Pseudocampylium radicale*

(NMW C.2010.030.8020)

This specimen was found in VC 13 (West Sussex), deviating from the census data which shows a more westerly distribution (Figure 4). This specimen is from a historical book of pressed bryophytes dating from the 1850s and so was only observed *in situ* under a stereo microscope so as not to damage the specimen. The original identification for this specimen was *Hypnum radicale* P. Beauv (now *Pseudocampylium radicale*). From inspection, the leaf of this specimen has a distinctive bend in the nerve which extends into the apex of the leaf like that of *Hygroamblystegium varium* (Hedw.) Mönk. (Figure 5) which can be found in West Sussex (Figure 6). The leaves of *Hygroamblystegium varium* are ovate with long acumen and stem leaves are 1.0 – 1.4 mm long (Smith, 2004). These characteristics can be seen in Figure 5.

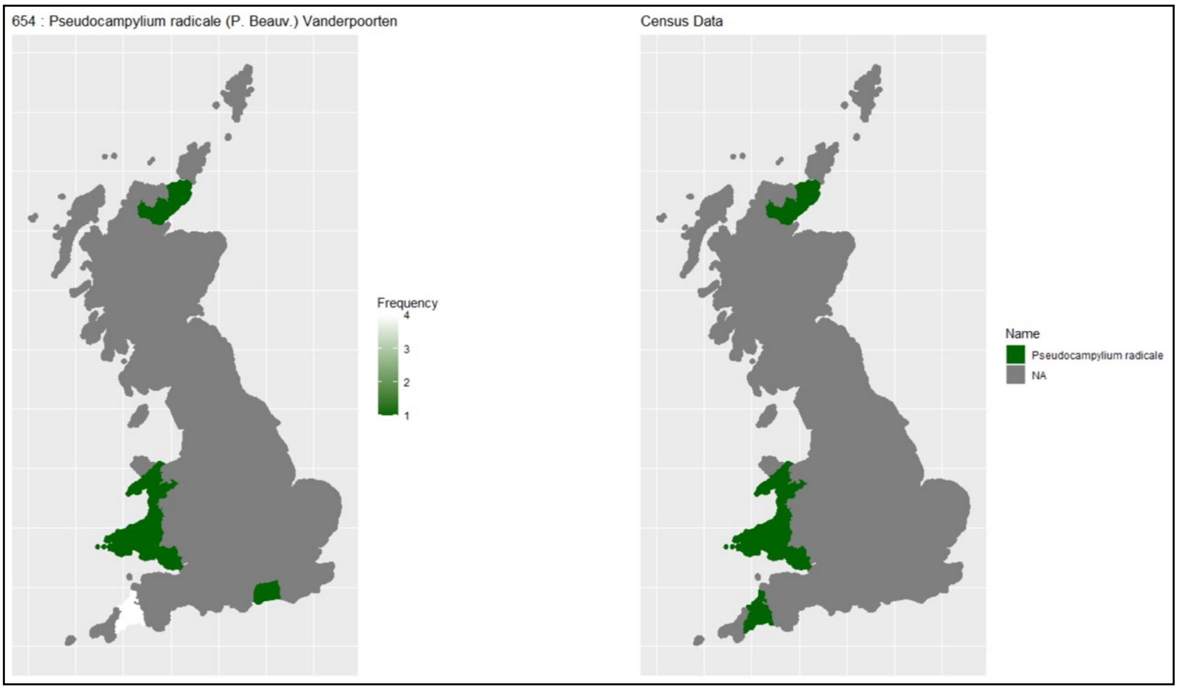


Figure 4. The vice county data for Pseudocampylium radicale *(NMW C.2010.030.8020). The image on the left shows the distribution from the herbarium data: vice counties 2, 13, 41, 44 – 46, 48, 49, 107, 109. The image on the right shows the British Bryological Society 2021 Census Catalogue (Blockeel* et al.*, 2021b) distribution: vice counties 2, 41, 44 – 46, 48, 49, 107, 109.*

Figure 5. Microscope image of (NMW C.2010.030.8020). Image shows the distinct bend in the nerve of the leaf like that of Hygroamblystegium varium.

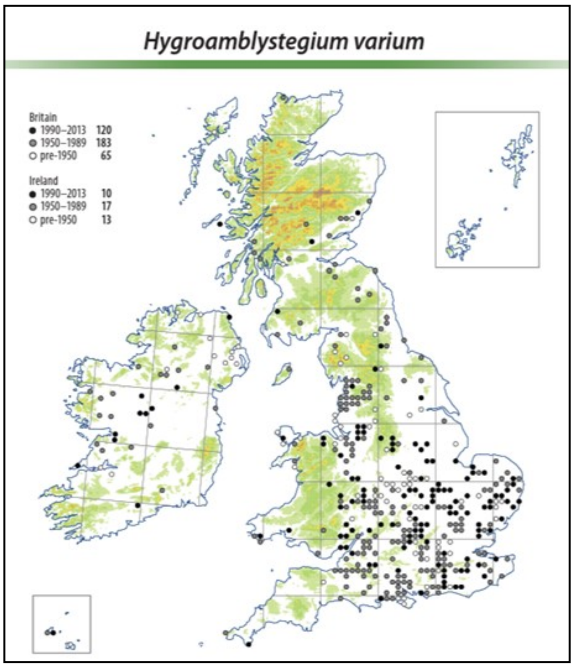


Figure 6. The distribution of Hygroamblystegium varium *from the Atlas of British Bryophytes (Blockeel* et al.*, 2014). This species has a wide distribution across England and includes West Sussex where the specimen NMW C.2010.030.8020 was found.*

3. *Philonotis tomentella*

(NMW 13.68.49, 15.54.1, 20.7.m.10 - 12, 22.187d.977, 23.92.685, 24.457.44 - 45, 25.152.4046 - 47, 25.152.4050, 25.152.4068, 40.443.46, 42.13.4, 44.265.8, 48.29.48, 64.97.488, 66.230.104, 71.1B.122)

Philonotis tomentella specimens in the collection had many vice county records that were not found in the census data (Figure 7). The vice counties recorded for *P. tomentella* not in the census data are: 5, 35, 47, 40 – 42, 46, 48, 58, 60, 62, and 87. This species has an altitudinal range of 50 – 1125 m and has been found growing in a variety of habitats on basic cliffs and sandy and peaty ground. This species is relatively scarce and closely related to *P. fontana* (Hedw.) Brid. which is a more widespread species (Blockeel *et al*., 2014). *Philonotis* is a difficult group and species can be difficult to distinguish from one another due to high levels of variations and integrations between species (Atherton *et al*., 2010; Buryová, 2004). However, there seems to be some confusion in the taxonomy of *Philonotis tomentella*. The specimen labels show them to have been originally identified as *P. fontana* and then redetermined as *P. fontana* var. *tomentella* (Molendo) A. Jaeger, before being transferred to *P. tomentella*. It is therefore likely that these specimens all belong to *P. fontana*. Determining the identity of the suspicious specimens is beyond the scope of this study.

4. *Riccia crystallina*

(NMW C96.15.130, C96.15.2959, C96.15.2961 - 67, C96.15.2971 - 78, C97.3.1682, C97.3.1687 - 88, C.1999.028.3603, C.1999.028.3616, C.1999.028.3619, C.1999.028.3942 - 47, C.2000.008.186) and (BBSUK C.2001.020.8617 - 19)

Riccia crystallina is a liverwort which has many different vice county records in the NMW and BBSUK herbaria, compared to the census data (Figure 8). The vice counties recorded for *R. crystallina* not in the census data are: 4, 6, 12, 14, 20, 22, 26, 28, 29, 32, 36, 38, 49, 55, 56, 64, 67, 83, 101 and 110. This species has a distinct ecology, growing in arable fields and sandy soil with an altitudinal range of 0 – 90 m (Blockeel *et al*., 2014). This species was split from *R. cavernosa* Hoffm. in 1966 which has a distribution more closely resembling that of the collection data (Figure 10). The herbarium packets for the mismatched specimens show that the original identifications are *R. crystallina* however many of these specimens are pre. 1966 and are likely to now be considered *R. cavernosa* (Paton, 1999). When these specimens were observed microscopically, many of the specimens resembled other *Riccia* species as the rosettes were not fused together like that of *R. crystallina* with some likely to be *R. cavernosa* whose rosettes are made up of more distinct lobes (Figure 9). As much of the

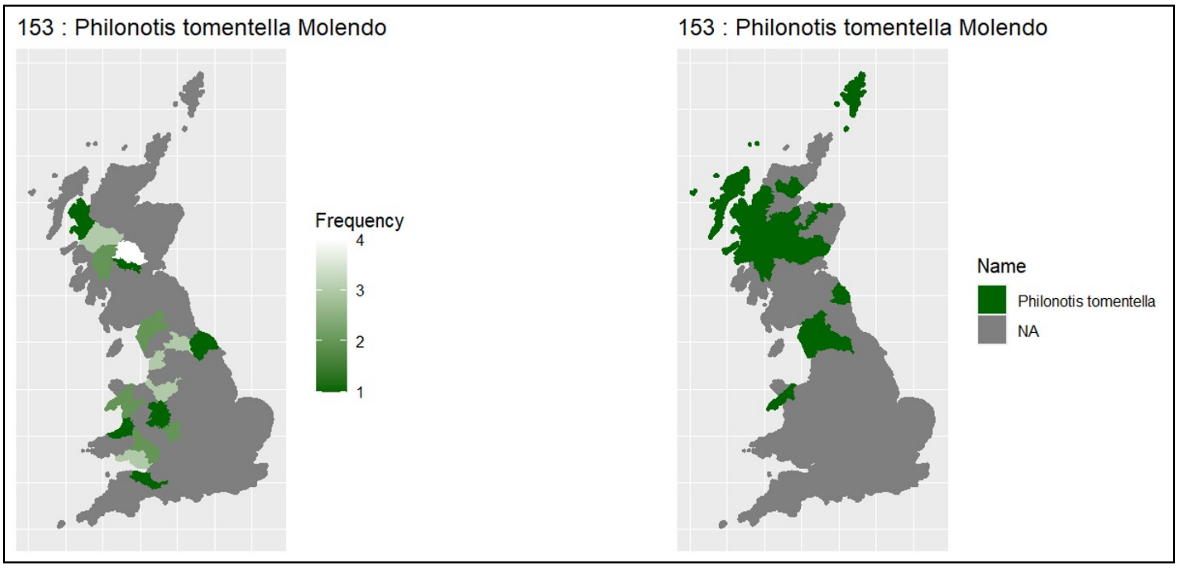


Figure 7. The vice county data for Philonotis tomentella. *The image on the left shows the distribution from the herbaria data: vice counties 5, 35, 37, 40 – 42, 46, 48, 49, 58, 60, 62, 65, 70, 87, 88, 97, 98, 104. The image on the right shows the British Bryological Society 2021 Census Catalogue (Blockeel* et al., *2021b) distribution: vice counties 49, (65, 68 – 70), 88, (89, 90), 94, 96 – 98, 103 – 105, 107, (110), 112. Bracketed vice counties are those that have not been observed since 1969.*

material for these specimens was very fragile, it was decided at this time it should not be hydrated and therefore identification to species level could not be performed for this difficult group within this study.

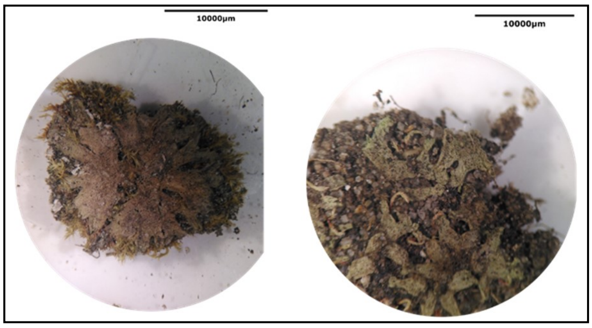


Figure 9. Examples of specimens labelled as Riccia crystallina *(NMW C96.15.2963 and NMW C96.15.261). Specimens show rosettes with more distinct lobes not fused together like that of* Riccia crystallina.

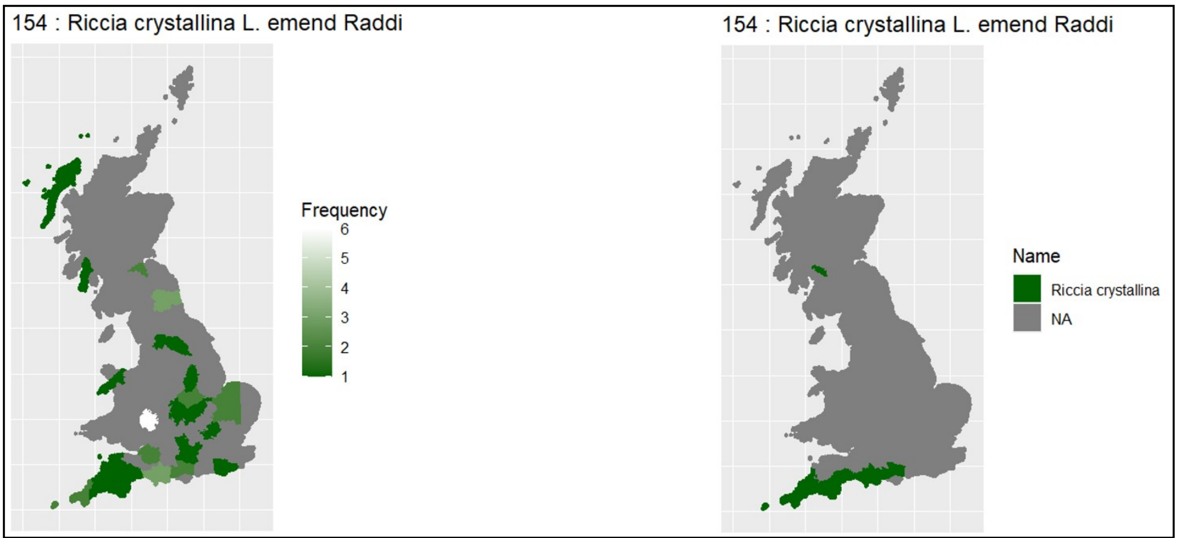


Figure 8. The vice county data for Riccia crystallina. *The image on the left shows the distribution from the herbarium data: vice counties 1 – 4, 6, 9, 11, 12, 14, 20, 22, 26, 28, 29, 32, 36, 38, 49, 55, 56, 64, 67, 83, 101, 110. The image on the right shows the British Bryological Society 2021 Census Catalogue (Blockeel* et al., *2021b) distribution: vice counties 1 – 3, 9, 11, (76). Bracketed vice counties are those that have not been observed since 1969.*

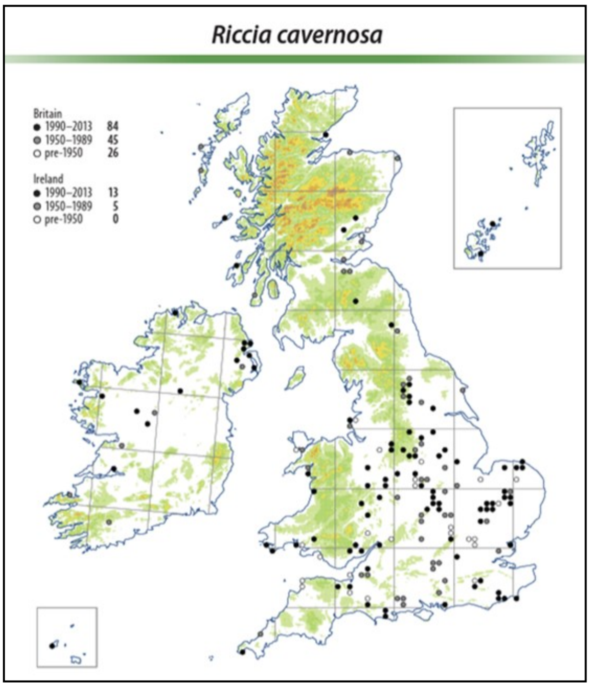


Figure 10. The distribution of Riccia cavernosa *from the Atlas of British Bryophytes (Blockeel* et al., *2014). This species has a greater distribution than* Riccia crystallina *and includes vice counties where the specimens were found.*

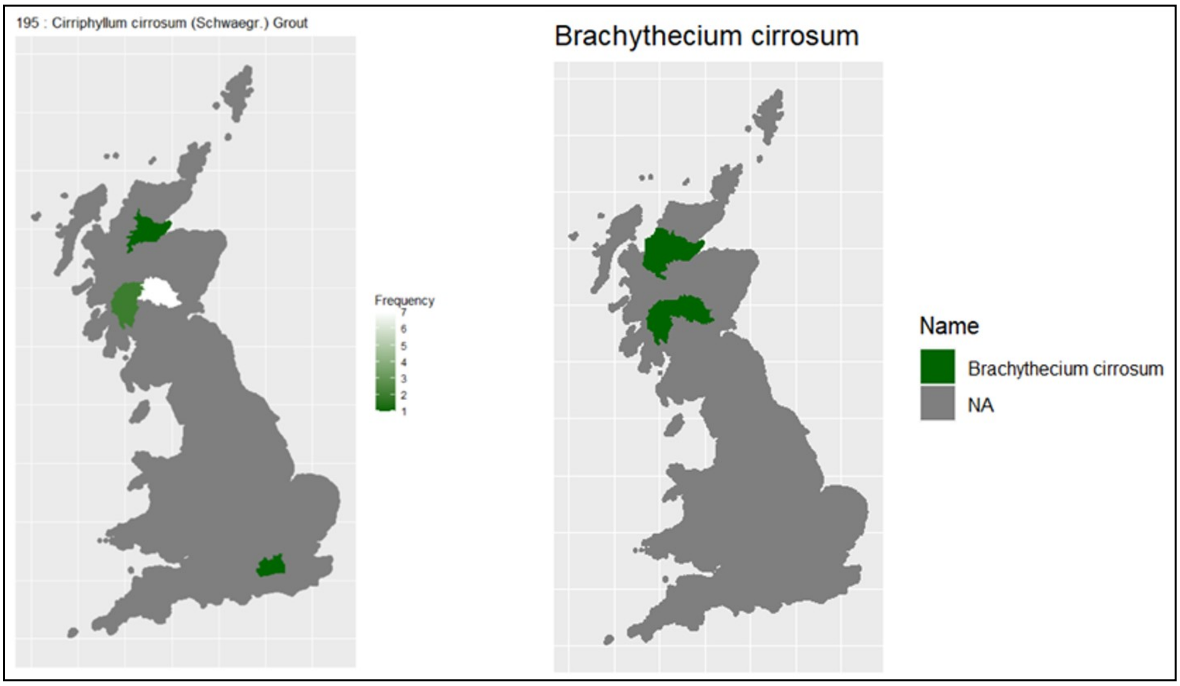
5. *Brachythecium cirrosum*

(NMW C.2000.002.641)

Brachythecium cirrosum (Schwägr.) Schimp. (synonym *Cirriphyllum cirrosum* (Schwaegr.) Grout (Flora of North America Editorial Committee, 2014)) has one specimen from a vice county not found in the census (Figure 11). *B. cirrosum* is found in Scotland on ledges or at the base of crags at higher altitudes (670 – 1070 m). This species is common in the high Arctic and found in many mountain ranges (Blockeel *et al*., 2014). However, this specimen was found in Denbies in VC 17 (Surrey) which has an altitude of around 50 m (Cucaera, accessed 2023). There is no current record of *B. cirrosum* in Surrey (Blockeel *et al*., 2021a; Gardiner, 1981). Like most species, bryophytes found at higher altitudes are likely to respond to a changing climate by shifting their elevational range usually so that they are increasingly found at higher altitudes than before (Rumpf *et al*., 2019). It therefore seems unlikely to find this species at a lower elevation than expected.

When this species was observed microscopically, it was found that the leaves have a rounded apices which tapers to a long acumen (Figure 12). The cells had a width of around 10 µm (Figure 12) and that the lower stem was pinnately branched. *Brachythecium cirrosum* has cells which are 5 – 8 µm wide and are irregularly branched (Smith, 2004). The description of this specimen closely matches that of the more common *Cirriphyllum piliferum* (Hedw.) Grout which has broad a distribution that includes VC 17 where this specimen was found (Figure 13).

Figure 12. Microscope images of (NMW C.2000.002.641). The image on the left shows the leaf shape of this specimen. The image on the right shows the elongated leaf cells around 10 µm wide. These features resemble more closely Brachythecium piliferum.



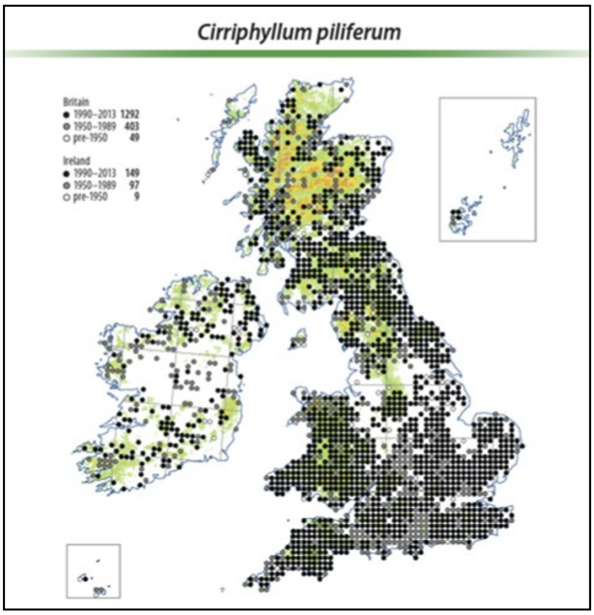


Figure 11. (Above) The vice county data for Brachythecium cirrosum *(NMW C.2000.002.641). The image on the left shows the distribution from the herbarium data: vice counties 17, 88, 98, 107. The image on the right shows the British Bryological Society 2021 Census Catalogue (Blockeel* et al.*, 2021b) distribution: vice counties 88, 98, 105, 106.*

Figure 13. The distribution of Cirriphyllum piliferum *from the Atlas of British Bryophytes (Blockeel* et al.*, 2014). This species has a large distribution and includes Surrey where the specimen NMW C.2000.002.641 was found.*

6. *Aongstroemia longipes*

(NMW C96.16.259)

This specimen had a vice county different to that of the census where it was found in VC 67 (Figure 14). *Aongstroemia longipes* is a circumpolar boreal-montane species that has only been recorded in the Scottish Highlands in Britain. Although, as it is a small species it can be easily overlooked in the field (Blockeel *et al.*, 2014). This specimen was recorded as having been found on an old lead mine waste tip in Allenheads, Northumberland and when verified microscopically it was found to be *Ditrichum plumbicola* Crundw. which is found on lead-mine spoil. The leaves of the specimen have a larger nerve than that of *Aongstroemia longipes* and are lanceolate rather than oblong-ovate (Figure 15). The leaves have a short apex compared to that of other *Ditrichum* species and are 0.4 – 0.7 mm long (Smith, 2004). The distribution of *Ditrichum plumbicola* includes South Northumberland where this specimen was found (Figure 16). *D. plumbicola* was not described as new species until 1976 (Crundwell, 1976) and this record (NMW C96.16.259) was collected in 1969 with *A. longipes* being the closest morphologically similar species. The collector of this specimen expressed doubt of the original identification on the specimen label. The earliest known record for this species was from 1914 (Blockeel *et al.*, 2014) however as this is a scarce species (Smith, 2004), this makes it an important voucher specimen, and could be an older record for this vice county.

Figure 15. Microscope images of (NMW C96.16.259). The image on the left shows the stems of the specimen. The image on the right shows the lanceolate leaf shape with wider nerve than that of Aongstroemia longipes. *These leaf characteristics resemble that of* Ditrichum plumbicola *more closely.*

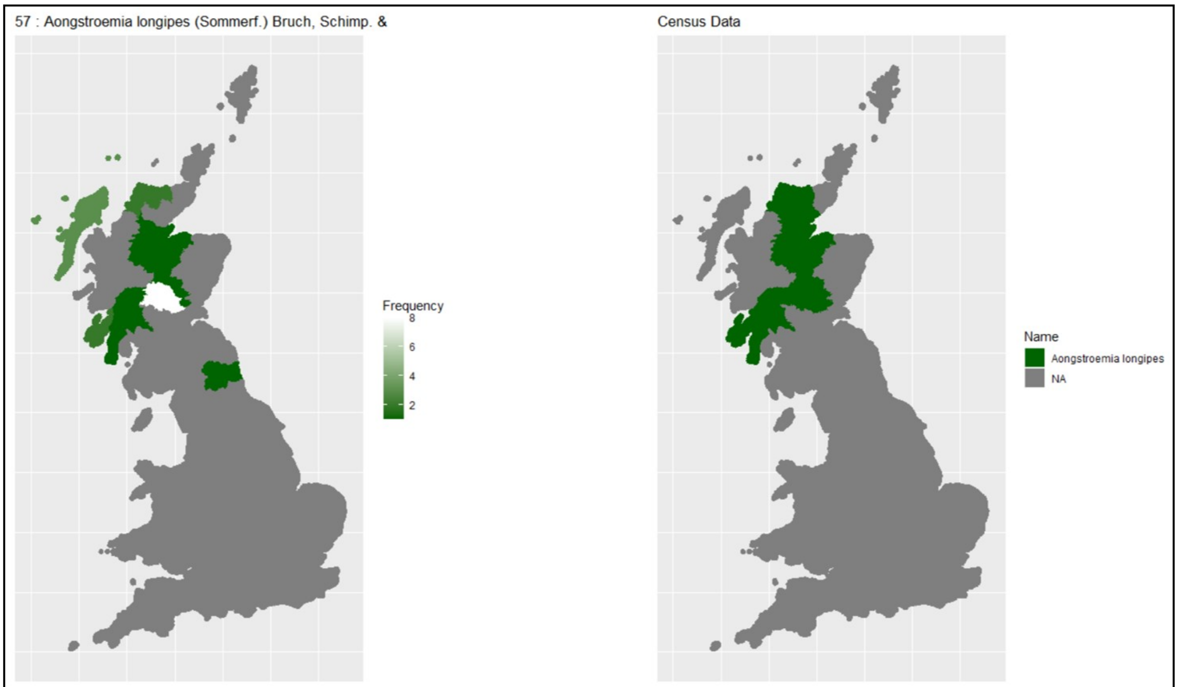


Figure 14. The vice county data for Aongstroemia longipes *(NMW C96.16.259). The image on the left shows the distribution from the herbarium data: vice counties 67, 88, 89, 95, 96, 98, 99, 101, 102, 106, 108, 110. The image on the right shows the British Bryological Society 2021 Census Catalogue (Blockeel* et al.*, 2021b) distribution: vice counties 88, 89, 95, 96, 98, 99, 101, 102, 106 – 108.*

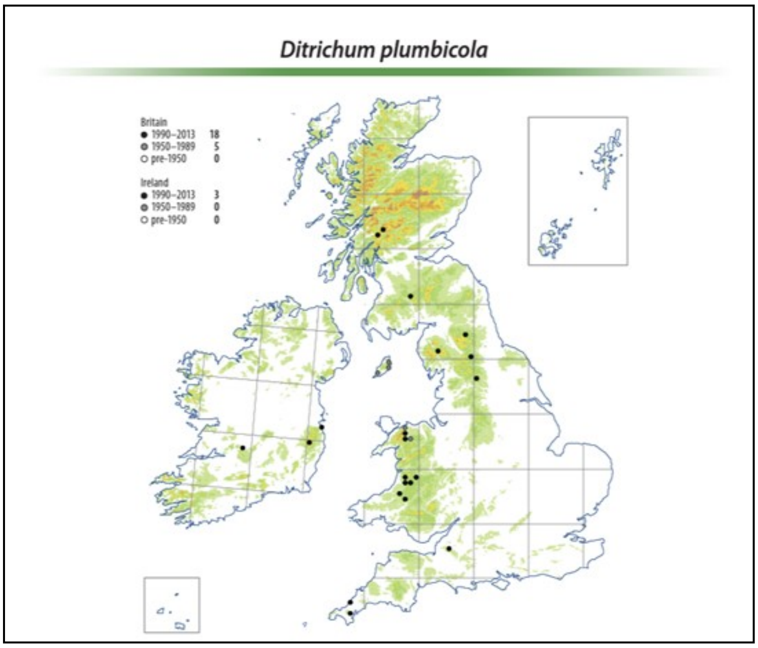


Figure 16. The distribution of Ditrichum plumbicola *from the Atlas of British Bryophytes (Blockeel* et al., *2014). This species is found on lead mine spoil and includes South Northumberland where the specimen NMW C96.16.259 was found.*

7. *Heterocladiella dimorpha*

(NMW C.2000.002.528)

This record was found in VC 73 (Kirkcudbrightshire) whereas the census data shows this species is found in the Scottish Highlands (Figure 17). The leaf shape of this specimen is not the same as *H. dimorpha* which have broadly ovate leaves with an acuminate apex (Smith, 2004). The leaves on this specimen are narrowly ovate, gradually tapering to an acute apex. Leaves are smaller than that of *H. dimorpha* with the longest being around 0.4 mm long (Figure 18) and do not show a distinct short double nerve. These leaf characteristics fit more closely with those of *Heterocladium flaccidum* (Schimp.) A.J.E. Sm. which is found in VC 73 (Figure 19).

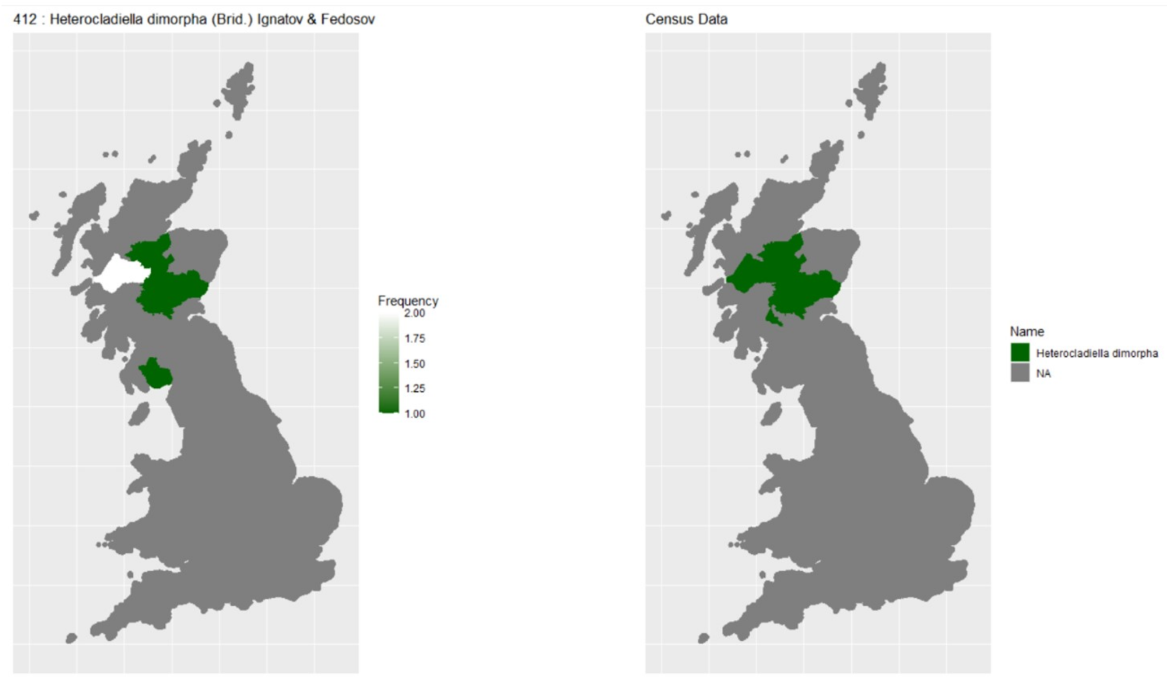


Figure 17. The vice county data for Heterocladiella dimorpha *(NMW C.2000.002.528). The image on the left shows the distribution from the herbarium data: vice counties 73, 87 – 90, 96, 97. The image on the right shows the British Bryological Society 2021 Census Catalogue (Blockeel* et al., *2021b) distribution: vice counties 87 – 90, 96, 97, 99.*

Figure 18. Microscopic image of specimen NMW C.2000.002.528. Leaves are narrowly ovate with acute apex no longer than 0.4 mm resembling those of Heterocladium flaccidum.

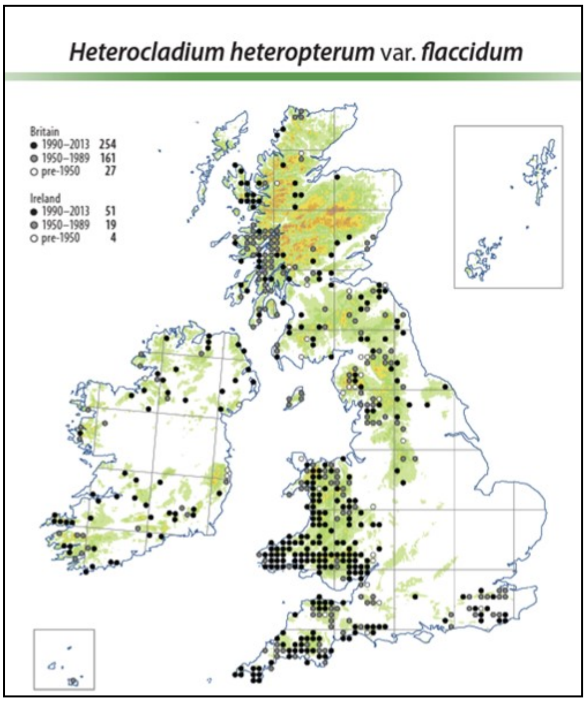


Figure 19. The distribution of Heterocladium flaccidum *(listed as* Heterocladium heteropterum *var.* flaccidum*) from the Atlas of British Bryophytes (Blockeel* et al.*, 2014). This species has a more westerly distribution but found in a variety of locations including VC 73 where NMW C.2000.002.528 was found.*

8. *Homomallium incurvatum*

(NMW C96.18.127)

This specimen was found in VC 107 (East Sutherland) in the north of Scotland, which is not recorded in the census data (Figure 20). After observing this specimen microscopically, it was found that this specimen was correctly identified, having capsules that are horizontal (Figure 21) and leaves that are lanceolate with a long acumen. This species also has distinct basal cells which are elongated but surrounded by small cells in the margin (Figure 22). The mid leaf cells are also small and rectangular (Figure 22). This is believed to be a new vice county record for this species which is Red Listed (endangered, Callaghan, 2022) and the specimen will be sent to the BBS Moss Recorder for confirmation. This is an unexpected outcome which has uncovered a very interesting record of a Red Listed species from a site not included in in the census data. The inclusion of this specimen is important as it allows the site to be targeted for future survey work for the threatened species.

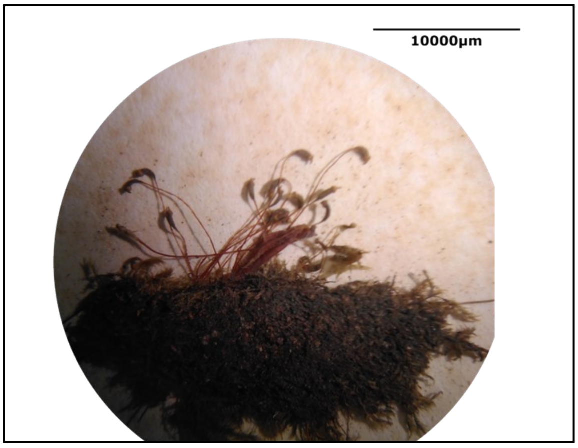


Figure 21. Microscope image of NMW C96.18.127 showing the horizontal capsules.

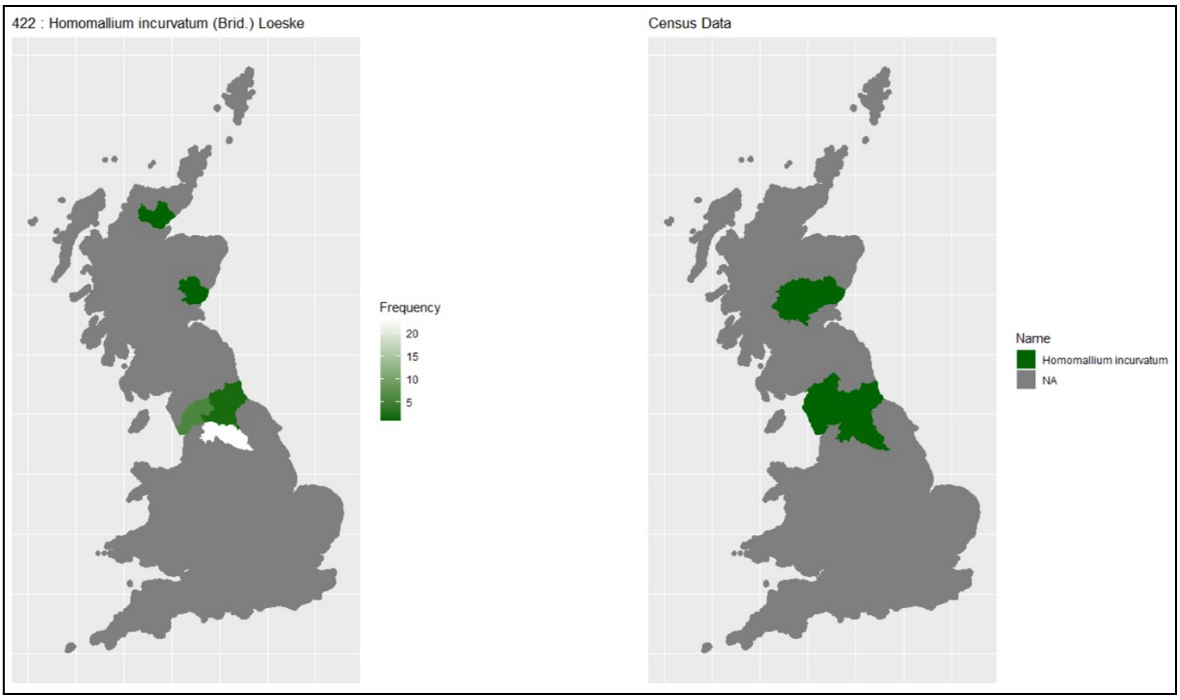


Figure 20. The vice county data for Homomallium incurvatum *(NMW C96.18.127). The image on the left shows the distribution from the herbarium data: vice counties 64 – 66, 69, 90, 107. The image on the right shows the British Bryological Society 2021 Census Catalogue (Blockeel* et al., *2021b) distribution: vice counties 64 – 66, 69, 70, 87 – 90.*

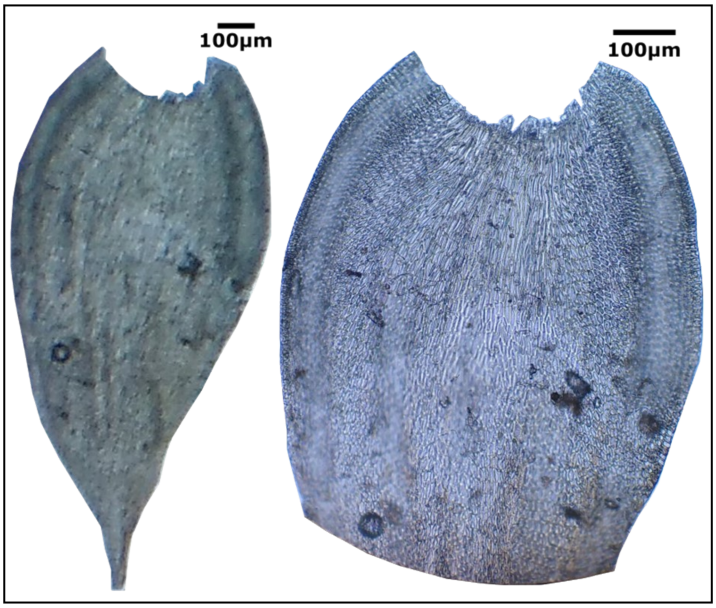


Figure 22. Microscopic image of NMW C96.18.127 leaf. Image on the left shows the lanceolate leaf shape tapering to a long acumen. Image on the right shows the distinctive elongated basal leaf cells surrounded by smaller rectangular cells. The cells in the mid leaf are rectangular rather than elongated.

9. *Paraleucobryum longifolium*

(NMW C97.12.161)

This specimen was found in VC 70 (Cumberland) whereas *Paraleucobryum longifolium* is found in the Scottish Highlands (Figure 23). Initial examination of this specimen revealed it to be a *Campylopus* species due to the long leaf shape with a wide base and tapering to a long, thin acumen (Figure 24). The width of the nerve in *P. longifolium* is greater than that seen in this specimen which is less than a third of the width of the leaf. However, it is larger than that of *Dicranum* species. The auricles of NMW C97.12.161 have a distinctive red-brown colouring and the basal cells are rectangular (Figure 24). The transverse section of the leaf shows small cells with thick walls and closely resembles the transverse section of *Campylopus flexuosus* (Hedw.) Brid. (Figure 24). This is a species that has a wide distribution including Cumberland (Figure 25).

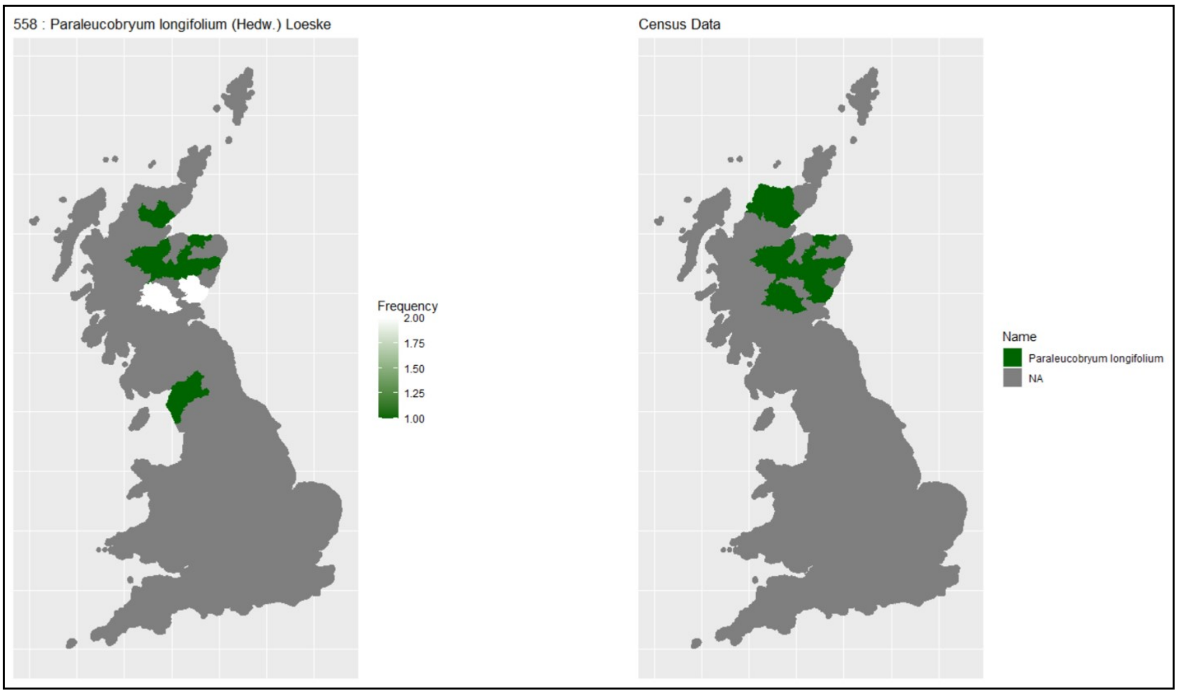


Figure 23. The vice county data for Paraleucobryum longifolium *(NMW C97.12.161). The image on the left shows the distribution from the herbarium data: vice counties 70, 88, 90, 92, 94, 96, 107. The image on the right shows the British Bryological Society 2021 Census Catalogue (Blockeel* et al.*, 2021b) distribution: vice counties 88, 90, 92, 94, 96, 107, 108.*

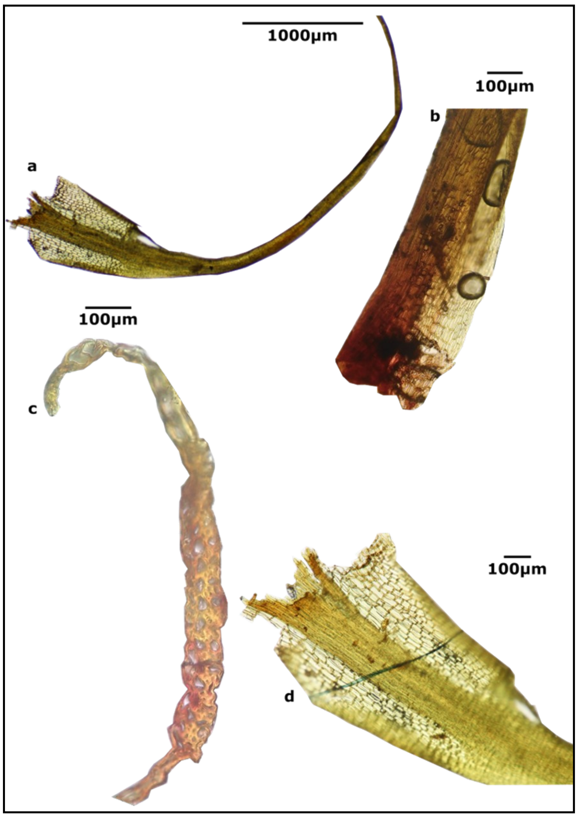


Figure 24. Microscopic images of NMW C97.12.161. a: leaf shape showing wide base and long tapering acumen, nerve less than 1/3 width of leaf. b: red-brown colouring of auricles. c: transverse section of leaf showing small cells in middle with thick cell walls (yellow in colour). d: rectangular basal leaf cells. Leaf characteristics similar to that of Campylopus flexuosus.

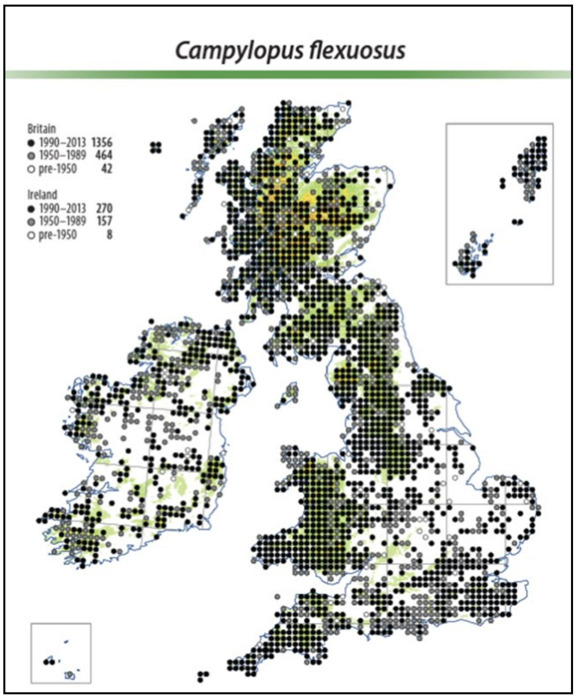


Figure 25. The distribution of Campylopus flexuosus *from the Atlas of British Bryophytes (Blockeel* et al.*, 2014). This species has a wide distribution including VC 70 where NMW C97.12.161was found.*

10. *Plagiothecium platyphyllum*

(NMW C.2000.020.28)

This specimen of *Plagiothecium platyphyllum* was found in VC 45 (Pembrokeshire) which is not recorded in the census data (Figure 26). This is a nationally scarce species (Preston, 2006) which can be found in a variety of wet habitats such as springs, rock crevices or by waterfalls in higher altitudes (480 – 870 m).

From microscopic inspection this specimen was found to be in the *Plagiothecium denticulatum* (Hedw.) Schimp. complex. This specimen has an asymmetrical, ovate-lanceolate leaf shape, elongated leaf cells and a double short nerve (Figure 27). *Plagiothecium platyphyllum* is also sharply denticulate near the apex and abruptly tapers to an acumen (Smith, 2004) which is not seen in this specimen. The double nerve of *P. denticulatum* is longer than that of *P. platyphyllum*. *Plagiothecium denticulatum* var. *denticulatum* has a wide distribution that covers Pembrokeshire where this specimen was found and is the more likely variety for this specimen to be (Figure 28) having an acute leaf shape more similar to this specimen.

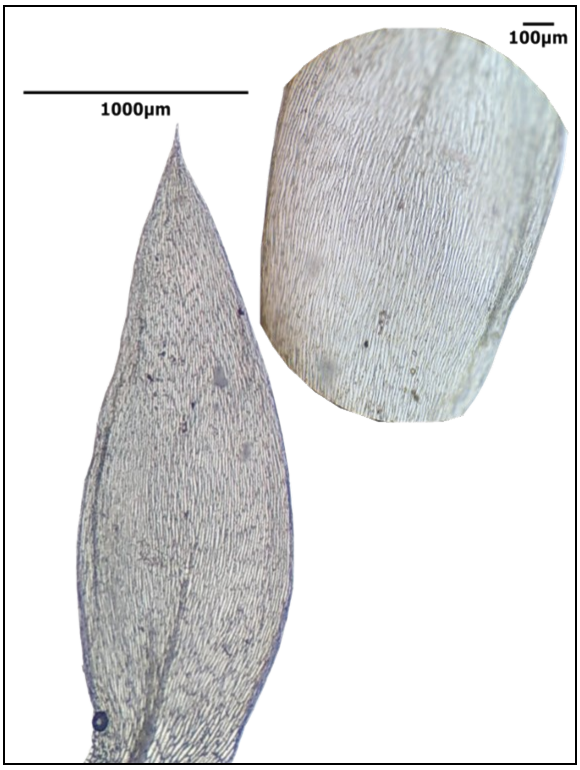


Figure 27. Microscopic images of NMW C.2000.020.28. The image on the left shows the ovate-lanceolate, asymmetrical leaf shape and double short nerve similar to that of Plagiothecium denticulatum. *The image on the right shows the elongated*

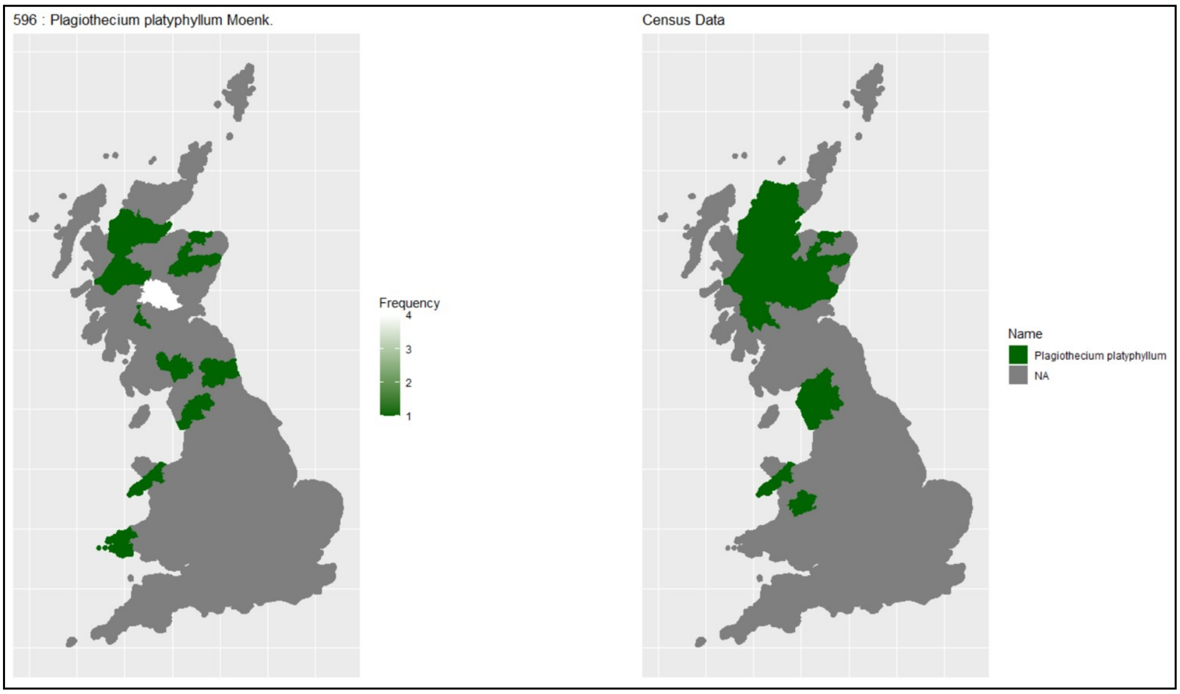


Figure 26. The vice county data for Plagiothecium platyphyllum. *The image on the left shows the distribution from the herbarium data: vice counties 45, 49, 67, 69, 72, 88, 92, 94, 97, 99, 105, 106. The image on the right shows the British Bryological Society 2021 Census Catalogue (Blockeel* et al., *2021b) distribution: vice counties 47, 49, (69), 70, 88, 89, (90), 92, 94, 96 – 99, 105 – 108. Bracketed vice counties are those that have not been observed since 1969.*

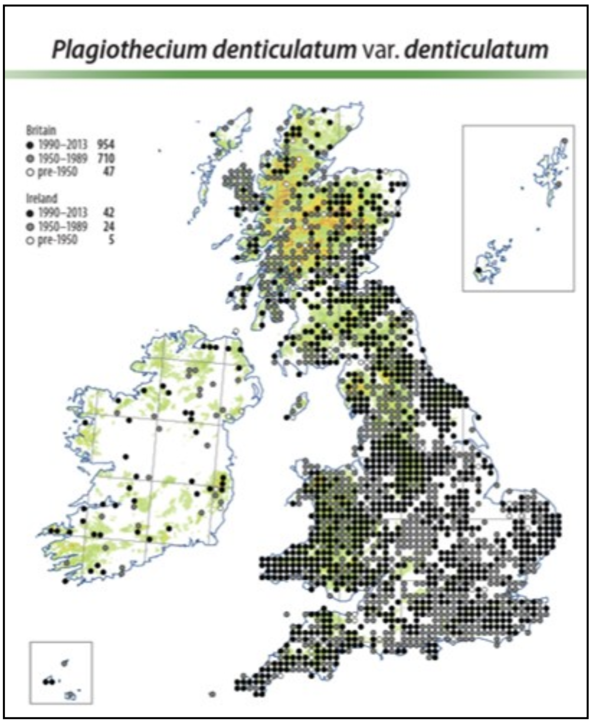


Figure 28. Plagiothecium denticulatum *var.* denticulatum *has a wide distribution including VC 45. This variety of P. denticulatum is more common than* Plagiothecium denticulatum *var.* obtusifolium *which is found in higher altitudes and is not found in Pembrokeshire.*

An overview of the results can be found in Table 2 showing that 70% of the species had specimens that had been misidentified. This can be broken down into 7 specimens requiring reidentification, 53 specimens requiring further work beyond the scope of this study and 1 specimen which had the correct identification.

Table 2. An overview of the top 10 most suspicious species flagged and reviewed and the number of specimens for that species that were suspicious and assessed microscopically. The reidentification column shows new identifications or explanations if no new identifications.

Species	**Number of Specimens**	**Reidentification**
Neckera pennata Hedw.	1	*Neckera complanata* (Hedw.) Huebener.
Pseudocampylium radicale (P. Beauv.) Vanderpoorten	1	*Hygroamblystegium varium* (Hedw.) Mönk
Philonotis tomentella Molendo	20	Taxonomic confusion that requires work beyond the scope of this study.
Riccia crystallina L. emend Raddi	33	Fragile material that could not be reidentified during this study.
Cirriphyllum cirrosum (Schwaegr.) Grout	1	*Cirriphyllum piliferum* (Hedw.)
Aongstroemia longipes (Sommerf.) Bruch & Schimp.	1	*Ditrichum plumbicola* Crundw
Heterocladiella dimorpha (Brid.) Ignatov & Fedosov	1	*Heterocladium flaccidum* (Schimp.) A.J.E. Sm.
Homomallium incurvatum (Brid.) Loeske	1	Correct identification
Paraleucobryum longifolium (Hedw.) Loeske	1	*Campylopus flexuosus* (Hedw.) Brid.
Plagiothecium platyphyllum Moenk.	1	*Plagiothecium denticulatum* (Hedw.) Schimp. complex.

Discussion

The list of suspicious specimens produced in this case study, particularly for the mosses, shows that the R package is able to detect specimens that have been misidentified, misnamed, or which have been left behind in changes of taxonomy (particularly in those resulting in splitting of a species into two or more distinct species). 70% of species reviewed in this study had specimens that were misidentified. Of the 10 species (totalling 61 specimens) determined as the most suspicious: 7 specimens were redetermined; 1 specimen was a new regional record not incorporated in the reference dataset for an endangered Red Listed species and 2 species (53 specimens) showed taxonomic reassessment was required. This was only a small subset of the possible species that could be reviewed due to time constraints and although only 11% of specimens assessed were able to be reidentified, 87% of specimens showed potential misnaming that was not possible to be rectified in the study. The only specimen to be correctly identified had been collected in a vice county not included in published data. This proves not only the effectiveness of the tool in identifying specimens labelled with incorrect or outdated specimen labels but also highlights its potential for identifying new vice county records held in collections and opens up further research possibilities of the tool into investigating and evaluating a species' distribution. For example, the specimen NMW C96.16.259 was redetermined as *Ditrichum plumbicola* and thus becomes a record from before the species was described. Even within a vice county where the species has been recorded before, the new record may have been found in a locality that is new within that vice county and can help further understanding of a species ecology and conservation needs.

Bryophytes are an understudied group being part of the 'minority taxa' that receive smaller research interest relative to their abundance (Smith, 2020). Along with other groups such as fungi, lichens, and algae, they can easily be subjected to errors especially as some species require identification microscopically (Atherton *et al*., 2010) and sometimes are only distinguishable from one another if certain morphological features are present. The case study presented here shows that bryophyte specimens had been misidentified and that some groups are difficult to reidentify without expert knowledge and time. However, they are an important group of plants that play a key role in habitat creation and improving biodiversity as well as being indicators of climate change, particularly through assessing changes in their distribution (Gignac, 2001).

The tool presented here has uses beyond bryophytes and can be used to review data for other areas of research that rely on correct identification of specimens and samples and reliable provenances. Thus, it is hoped that it will be an important tool to verify specimen data before it is shared online, particularly because whilst collections are becoming more available online, publicly available data has often been shown to be inaccurate. For example, the fungal sequences deposited in GenBank have been shown to contain a high number of misidentified taxa (Hofstetter *et al*., 2019). For the Agaricomycotina analysed in the study, it was found that around 30% of the fungal sequences in the database were misidentified. Correcting these mistakes in collection databases will ensure a higher quality and reliability of research that uses this data. The tool also has the potential to identify fraudulent records, such as those occurred in the case of Prof. John William Heslop Harrison, who purposefully and deliberately engaged in the collection and recording of specimens that he has planted on the Isle of Rum (VC 104) (Pearman and Walker, 2004).

Furthermore, providing a collection of data which is as accurate as possible is important for studies on how a changing climate is affecting species as well as research into biodiversity loss. Analysing changes in species distributions can be an effective tool, however if the data is formed from misidentified specimens this can both increase and decrease a species' distribution (Costa *et al*., 2015). Producing such distributions can show potential biodiversity hotspots as well as areas where biodiversity is low or areas where more data should be collected (Mannino *et al*., 2020; Meier and Dikow, 2004). For rarer species, distributions can be misleading as these species are more likely to be misidentified (Aubry *et al*., 2017). Species that are more common are less likely to be collected than rarer species and from areas that are easier to collect from which results in a spatial bias (Costa *et al*., 2015; Isaac and Pocock, 2015). The analysis of records presented here also shows the importance of having collections of specimens. Without this evidence, identifications could not be reassessed, biological records could not be updated and finding new regional records would not be possible.

Whilst it has already been suggested that specimen identifications are checked before research is carried out (Kitchener *et al*., 2020), it can be time-consuming particularly through the need to systematically check collections especially for larger datasets. For example, in a study by Kauserud *et al.* (2008), around 35,000 fungi specimen records were used. It would not be possible to verify the identification for all these records. Similarly, for bryophytes and lichens, this time commitment is particularly high given the need for microscopic identification that precludes automation such as automated image identification tools (Shirai *et al*., 2022), which was able to both select and correct misidentified specimens. However, Shirai *et al.* (2022) only used vascular plant specimens showing further that minority taxa are often forgotten in studies. The tool presented here presents a time-saving procedure to identify samples likely to be misidentified for further reassessment which doesn't rely on photographically identifiable macromorphological changes. The R Package can assess thousands of records at once and only those chosen are reviewed in person. Such a process is only limited by computer power and identification abilities.

The R Package presented here can be used on data of all sizes from collections and observation records of different organisms to find a selection of specimens with a high likelihood of being misidentified or misnamed, as well as detecting new vice county records. This package provides a tool for quick assessment of records which can be evaluated for importance of investigation. As the majority of the specimens reidentified were nationally scarce species, it further highlights the wider potential applications of this tool in informing species conservation measures and wider ecological policy.

References

Andrew, C., Diez, J., James, T.Y. and Kauserud, H., 2019. Fungarium specimens: a largely untapped source in global change biology and beyond. *Philosophical Transactions of the Royal Society B*, 374(1763), p.20170392.

Atherton, I., Bosanquet, S. and Lawley, M. eds., 2010. *Mosses and liverworts of Britain and Ireland: a field guide* (pp. 4-6, 636). Plymouth: British Bryological Society.

Aubry, K.B., Raley, C.M. and McKelvey, K.S., 2017. The importance of data quality for generating reliable distribution models for rare, elusive, and cryptic species. *PLoS One*, 12(6), p.e0179152.

Bieker, V.C. and Martin, M.D., 2018. Implications and future prospects for evolutionary analyses of DNA in historical herbarium collections. *Botany Letters*, 165(3-4), pp.409-418.

Biological Records Centre, 2019. *Watsonian vice county boundaries GIS layers*. [online] Available at: https://github.com/BiologicalRecordsCentre/vice-counties [Accessed 6 Jun. 2023].

Blockeel ,T.L., Bosanquet, S.D.S., Hill, M.O. & Preston, C.D. eds., 2014. *Atlas of British and Irish Bryophytes*. Newbury: Pisces Publications.

Blockeel, T.L., Bell, N.E., Hill, M.O., Hodgetts, N.G., Long, D.G., Pilkington, S.L. and Rothero, G.P., 2021a. A new checklist of the bryophytes of Britain and Ireland, 2020. *Journal of Bryology*, 43(1), pp.1-51.

Blockeel, T.L., Hodgetts, N.G., Pilkington, S.L., Pescott, O.L. 2021b. *A Census Catalogue of British and Irish Bryophytes 2021*. Southampton: British Bryological Society.

Bradshaw, A.J., Backman, T.A., Ramírez-Cruz, V., Forrister, D.L., Winter, J.M., Guzmán-Dávalos, L., Furci, G., Stamets, P. and Dentinger, B.T., 2022. DNA Authentication and Chemical Analysis of Psilocybe Mushrooms Reveal Widespread Misdeterminations in Fungaria and Inconsistencies in Metabolites. *Applied and Environmental Microbiology*, 88(24), pp.e01498-22.

Buryová, B., 2004. Genetic variation in two closely related species of Philonotis based on isozymes. *The Bryologist*, 107(3), pp.316-327.

Callaghan, D.A., 2022. A new IUCN Red List of the bryophytes of Britain, 2023. *Journal of Bryology*, 44(4), pp.271-389.

Christenhusz, M.J. and Chase, M.W., 2018. PPG recognises too many fern genera. *Taxon*, 67(3), pp.481-487.

Cleal, C.J., Pardoe, H.S., Slade, K., Whyman, S., Tangney, R.S. and Jüttner, I., 2022. The Welsh National Herbarium. *Botany Letters*, 169(1), pp.3-17.

Costa, H., Foody, G.M., Jiménez, S. and Silva, L., 2015. Impacts of species misidentification on species distribution modeling with presence-only data. *ISPRS International Journal of Geo-Information*, 4(4), pp.2496-2518.

Crundwell, A.C., 1976. Ditrichum plumbicola, a new species from lead-mine waste. *Journal of Bryology*, 9(2), pp.167-169.

Cucaera, (updated continuously). *Cucaera Grid Reference Plotter*. [online] Available at: https://www.cucaera.co.uk/grp/ [Accessed 19 Oct. 2023].

Flora of North America Editorial Committee, 2014. Bryophyta, part 2. 28: i–xxii, 1–702. In: *Flora of North America North of Mexico*. New York: Oxford University Press.

Foster, W.D., 1979. The History of the Moss Exchange Club. British Bryological Society Bulletin: 33, 19-26.

García-Roselló, E., Guisande, C., González-Dacosta, J., Heine, J., Pelayo-Villamil, P., Manjarrás-Hernández, A., Vaamonde, A. and Granado-Lorencio, C., 2013. ModestR: a software tool for managing and analyzing species distribution map databases. *Ecography*, 36(11), pp.1202-1207.

Gardiner, J.C., 1981. A bryophyte flora of Surrey. *Journal of Bryology*, 11(4), pp.747-841.

Gignac, L.D., 2001. Bryophytes as indicators of climate change. *The bryologist*, 104(3), pp.410-420.

Golo, R., Vergés, A., Díaz-Tapia, P. and Cebrian, E., 2023. Implications of taxonomic misidentification for future invasion predictions: Evidence from one of the most harmful invasive marine algae. *Marine Pollution Bulletin*, 191, p.114970.

Goodwin, Z.A., Harris, D.J., Filer, D., Wood, J.R. and Scotland, R.W., 2015. Widespread mistaken identity in tropical plant collections. *Current biology*, 25(22), pp.R1066-R1067.

Harrison, S.G., 1980. The Herbarium of the British Bryological Society, BBSUK. British Bryological Society Bulletin: 36, p.50.

Hijmans, R.J., Guarino, L., Cruz, M. and Rojas, E., 2001. Computer tools for spatial analysis of plant genetic resources data: 1. DIVA-GIS. *Plant genetic resources newsletter*, pp.15-19.

Hofstetter, V., Buyck, B., Eyssartier, G., Schnee, S. and Gindro, K., 2019. The unbearable lightness of sequenced-based identification. *Fungal Diversity*, 96(1), pp.243-284.

Isaac, N.J. and Pocock, M.J., 2015. Bias and information in biological records. *Biological Journal of the Linnean Society*, 115(3), pp.522-531.

Ivison, K., Speed, J.D., Prestø, T. and Dawson, W., 2023. Testing enemy release of non-native plants across time and space using herbarium specimens in Norway. *Journal of Ecology*, 111(2), pp.300-313.

Kauserud, H., Stige, L.C., Vik, J.O., Økland, R.H., Høiland, K. and Stenseth, N.C., 2008. Mushroom fruiting and climate change. *Proceedings of the National Academy of Sciences*, 105(10), pp.3811-3814.

Kitchener, A.C., Machado, F.A., Hayssen, V., Moehlman, P.D. and Viranta, S., 2020. Consequences of the misidentification of museum specimens: the taxonomic status of Canis lupaster soudanicus. *Journal of Mammalogy*, 101(4), pp.1148-1150.

Lang, P.L., Willems, F.M., Scheepens, J.F., Burbano, H.A. and Bossdorf, O., 2019. Using herbaria to study global environmental change. *New phytologist*, 221(1), pp.110-122.

Łuczaj, Ł.J., 2010. Plant identification credibility in ethnobotany: a closer look at Polish ethnographic studies. *Journal of ethnobiology and ethnomedicine*, 6(1), pp.1-16.

Mannino, A.M., Armeli Minicante, S. and Rodríguez-Prieto, C., 2020. Phycological Herbaria as a useful tool to monitor long-term changes of macroalgae diversity: Some case studies from the Mediterranean Sea. *Diversity*, 12(8), p.309.

Meier, R. and Dikow, T., 2004. Significance of specimen databases from taxonomic revisions for estimating and mapping the global species diversity of invertebrates and repatriating reliable specimen data. *Conservation Biology*, 18 (2), pp.478-488.

Olds, C.G., Berta-Thompson, J.W., Loucks, J.J., Levy, R.A. and Wilson, A.W., 2023. Applying a modified metabarcoding approach for the sequencing of macrofungal specimens from fungarium collections. *Applications in Plant Sciences*, 11(1), p.e11508.

Paton, J.A., 1999. *The Liverwort Flora of the British Isles*. 1st ed. Colchester: Harley Books.

Pearman, D.A. and Walker, K.J., 2004. An examination of JW Heslop Harrison's unconfirmed plant records from Rum. *Watsonia*, 25(1), pp.45-64.

Pilkington, S.L. and Hodgetts N.G., 2023. British Bryological Society Interim Census Catalogue 2023. Including VC records compiled to the end of 2022. Unpublished. Available at: https://www.britishbryologicalsociety.org.uk/publications/census-catalogue/

Praeger, R.L., 1896. On the botanical subdivision of Ireland. *The Irish Naturalist*, 5(2), pp.29-38.

Preston, C.D., 2006. A revised list of nationally scarce bryophytes. *Field bryology*, 90, pp.22-30.

Raper, C., 2014. Dataset: UK species inventory—simplified copy. Natural History Museum Data Portal. Available at: https://data.nhm.ac.uk/dataset/uk-species-inventory-simplified-copy [Accessed 07 Nov. 23].

Roberts, S., 2023. SuspectSpecimens: Find Potentially Misidentified Specimens. R package version 0.3.0. Available at: https://github.com/SophieARoberts/SuspectSpecimensPackage [Accessed 18 Oct. 2023].

Rumpf, S.B., Hülber, K., Wessely, J., Willner, W., Moser, D., Gattringer, A., Klonner, G., Zimmermann, N.E. and Dullinger, S., 2019. Extinction debts and colonization credits of non-forest plants in the European Alps. *Nature Communications*, 10(1), p.4293.

Scott, W.A. and Hallam, C.J., 2003. Assessing species misidentification rates through quality assurance of vegetation monitoring. *Plant Ecology*, 165(1), pp.101-115.

Shirai, M., Takano, A., Kurosawa, T., Inoue, M., Tagane, S., Tanimoto, T., Koganeyama, T., Sato, H., Terasawa, T., Horie, T. and Mandai, I., 2022. Development of a system for the automated identification of herbarium specimens with high accuracy. *Scientific Reports*, 12(1), p.8066.

Smith, A.J.E., 2004. *The Moss Flora of Britain and Ireland*. 2nd ed. Cambridge: Cambridge university press.

Smith, N., 2020. Minority taxa, marginalised collections: A focus on fungi. *Journal of Natural Science Collections*, 7, pp.49-58.

Thiers, B.M., (updated continuously). *Index Herbariorum: a global directory of public herbaria and associated staff. New York Garden's Virtual Herbarium*. [online] Available at: https://sweetgum.nybg.org/science/ih/ [Accessed 19 Sep. 23].

Tropicos.org, (updated continuously). *Missouri Botanical Garden Tropicos database*. [online] Available at: https://tropicos.org [Accessed 19 Sep. 2023].

Watson, H.C., 1847. *Cybele Britannica: or british plants and their geographical relations* (Vol. 4). Longman, & Company.

Xu, C., Dong, W., Shi, S., Cheng, T., Li, C., Liu, Y., Wu, P., Wu, H., Gao, P. and Zhou, S., 2015. Accelerating plant DNA barcode reference library construction using herbarium specimens: improved experimental techniques. *Molecular Ecology Resources*, 15(6), pp.1366-1374.