

Article (refereed) - postprint

Brennan, Georgina L.; Potter, Caitlin; de Vere, Natasha; Griffith, Gareth W.; Skjøth, Carsten A.; Osborne, Nicholas J.; Wheeler, Benedict W.; McInnes, Rachel N.; Clewlow, Yolanda; Barber, Adam; Hanlon, Helen M.; Hegarty, Matthew; Jones, Laura; Kurganskiy, Alexander; Rowney, Francis M.; Armitage, Charlotte; Adams-Groom, Beverley; Ford, Col R.; Petch, Geoff M.; Elliot, Angela; Frisk, Carl A.; Neilson, Roy; Potter, Stephen; Rafiq, Abdullah M.; Roy, David B.; Selby, Katherine; Steinberg, Natascha; Creer, Simon. 2019. **Temperate airborne grass pollen defined by spatio-temporal shifts in community composition.** *Nature Ecology & Evolution*, 3 (5). 750-754.
<https://doi.org/10.1038/s41559-019-0849-7>

© 2019 Springer Nature Publishing AG

This version available <http://nora.nerc.ac.uk/id/eprint/524587/>

NERC has developed NORA to enable users to access research outputs wholly or partially funded by NERC. Copyright and other rights for material on this site are retained by the rights owners. Users should read the terms and conditions of use of this material at <http://nora.nerc.ac.uk/policies.html#access>

This document is the author's final manuscript version of the journal article following the peer review process. Some differences between this and the publisher's version may remain. You are advised to consult the publisher's version if you wish to cite from this article.

<https://link.springer.com/>

Contact CEH NORA team at
noraceh@ceh.ac.uk

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15

Supplementary information

Title: Temperate grass flowering season defined by spatio-temporal shifts in airborne pollen communities

Authors: Georgina L. Brennan^{1*}, Caitlin Potter², Natasha de Vere^{2,3}, Gareth W. Griffith², Carsten A. Skjøth⁴, Nicholas J. Osborne^{5,6}, Benedict W. Wheeler⁵, Rachel N. McInnes⁷, Yolanda Clewlow⁷, Adam Barber⁷, Helen M. Hanlon⁷, Matthew Hegarty², Laura Jones^{1,3}, Alexander Kurganskiy⁴, Francis M. Rowney⁵, Charlotte Armitage⁸, Beverley Adams-Groom⁴, Col R. Ford³, Geoff M. Petch⁴, The PollerGEN Consortium, and Simon Creer^{1*}.

Consortium Angela Elliot⁹, Carl A. Frisk⁴, Roy Neilson¹⁰, Stephen Potter¹¹, Abdullah M. Rafiq¹, David, B. Roy¹², Katherine Selby¹³, Natascha Steinberg⁹

Corresponding authors: Simon Creer (s.creer@bangor.ac.uk) and Georgina L. Brennan (g.l.b.doonan@gmail.com)

16 **Supplementary text**

17

18 Observations of first flowering dates from a citizen science project (UKPN;
19 www.naturescalendar.org.uk) and metabarcoding data show similar sequences of seasonal
20 progression (Supplementary Figure 5). First flowering dates of each genus started almost 3-4 weeks
21 prior to the observation of peaks of grass pollen in the metabarcoding data (Supplementary Figure 5).
22 Pollen release (anthesis) occurs approximately 2-3 weeks after the production of flowering heads
23 (heading)¹, and this is reflected in the metabarcoding data suggesting that local flowering data are
24 informative for predicting the composition of airborne pollen. At the family level, the transition of the
25 dominant airborne pollen from Pinaceae, to Poaceae, to Urticaceae reflects the flowering and
26 pollinating season of their contributing flora, providing further confidence that the airborne
27 community is represented by the phenology of local flora at the ground level (Supplementary Figure
28 4). This study substantially extends research into the relationship between airborne pollen and
29 terrestrial phenology. Previous studies have suggested that patterns in airborne pollen concentrations
30 are related to phenological behaviour of dominant species² and others have demonstrated that
31 strong correlations exist between local phenology and airborne pollen concentrations (e.g.^{3,4}).
32 However, for these previous studies, it was not possible to quantify the contribution to the daily grass
33 pollen concentrations from individual genera as optical recognition of grass pollen using microscopy
34 only allow for a detection at the family level. Continuing this study over multiple years would allow us
35 to track long-term, phenological changes in airborne pollen communities and improve our ability to
36 forecast the seasonal progression of airborne pollen⁵.

37

38 **Sequencing statistics**

39 In total, 460,981 high-quality sequences were assigned to ITS2 and 1,054,333 to *rbcl*. There
40 was considerable variation in the number of reads per sample, with the total number of high
41 quality *rbcl* sequences varying from 3 to 56,981 and the number of ITS2 sequences ranging

42 from 260 to 10,180 (excluding a single sample that contained zero sequences following
43 quality control).

44

45 Excluding control samples, ITS2 reads included 179,450 grass sequences (Poaceae), assigned
46 to 13 genera. Whilst not the focus of our study, the remaining 162,540 sequences consisted
47 of 33 families and 31 genera of terrestrial plants. Of these families, 17 contained only a single
48 genus and four contained reads which could not be identified confidently to genus level.

49 Within the *rbcL* marker, 179,330 grass reads were assigned to 13 grass genera and the
50 remaining 867,823 reads belonged to 68 families and 84 genera of terrestrial plants. Of these
51 families, 33 contained only a single genus and 13 contained reads which could not be
52 identified confidently to genus level. The top three most abundant pollen belonged to the
53 families Poaceae, Pinaceae and Urticaceae, accounting for 82% and 90% of total reads, for
54 *rbcL* and ITS2 respectively (Supplementary Figure 2). Poaceae was the dominant pollen
55 during the central part of the grass pollen season (June and July) (Supplementary Figure 2),
56 where the grass pollen season is identified using the 95% method. Using this method, the
57 pollen season is defined using the period of time when 2.5% and 97.5% of the yearly total
58 airborne pollen is collected⁶. This approach is generally used in UK (e.g.⁷) and many other
59 European countries. Pinaceae was the dominant airborne pollen during late spring/early
60 summer (May to June or early July at some locations), a period typically characterised by the
61 late tree pollen season. Urticaceae was the dominant airborne pollen during the end of the
62 summer (late July to August), indicating the onset of the weed pollen season.

63

64 **ITS2 and *rbcL* detect different grass species**

65 The contrasting characteristics of the ITS2 and *rbcL* markers makes them an ideal pairing. The
66 ITS2 marker shows high specificity between species but cannot detect all plants⁸, whereas
67 *rbcL* primers are highly universal but the marker shows lower resolution between closely
68 related plants⁹. For example, we found that the Pinaceae pollen season is longer when using
69 *rbcL* compared with ITS2 (Supplementary Figure 4), and it is likely that *rbcL* is picking up a
70 broader range of species within each family, compared to ITS2.

71

72

73 Of the grass genera identified, only four were present in both ITS2 and *rbcL* datasets:
74 *Dactylis*, *Lolium/Festuca*, *Anthoxanthum*, and *Avena*. While the proportion of reads assigned
75 to *Lolium/Festuca* and *Anthoxanthum* were correlated between the two markers
76 (*Lolium/Festuca*: $t_{72}=8.6$, adjusted p-value < 0.001 , $r^2 = 0.5$, Supplementary Figure 10A;
77 *Anthoxanthum*: $t_{72}=2.9$, adjusted p-value = 0.006, $r^2 = 0.09$, Supplementary Figure 10B), this
78 was not the case for *Dactylis* or *Avena* (Supplementary Figure 10C; Supplementary Figure
79 10D). However, both of the latter species were detected at relatively low levels in both
80 datasets, potentially increasing the degree of stochasticity introduced by library
81 preparation¹⁰.

82

83 **Positive and negative controls**

84 Negative controls, with all reagents and no DNA were used to identify any cross-
85 contamination. Of the six negative controls, four contained no reads following quality control
86 filtering, one contained a single read and one contained nine reads in the *rbcL* database.
87 None of the negative controls in the ITS2 dataset contained any reads.

88

89 Two positive control samples were also included, a grass positive control (Supplementary
90 Table 1) and an exotic plant positive control (list S1). Both sets of positive controls were
91 diluted to $0.3 \text{ ng } \mu\text{l}^{-1}$, similar to that of the aerial eDNA samples.

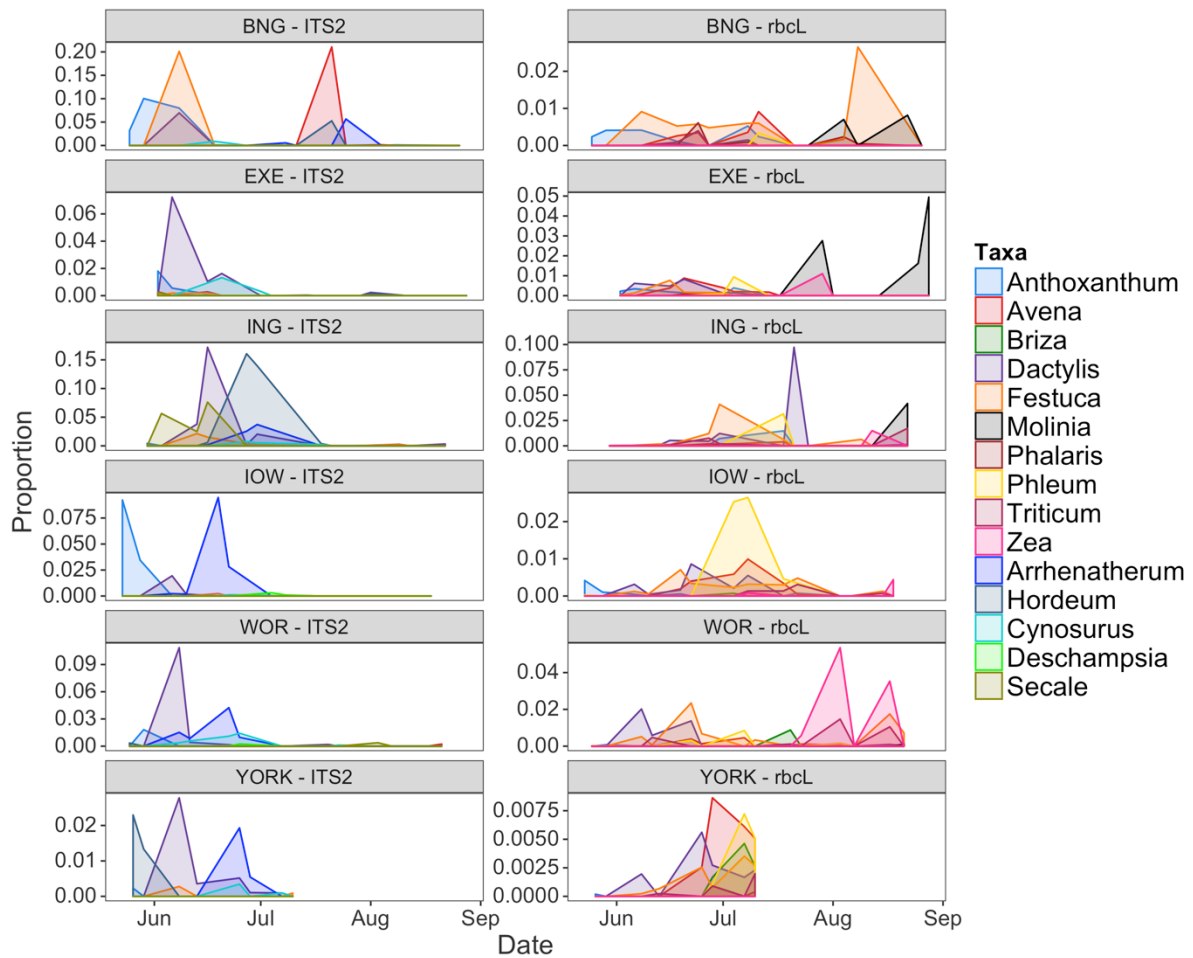
92

93 The grass positive control contained a mixture of fifty-two species of grass from herbarium
94 collections held at the National Botanic Garden of Wales. The mixture of grasses contained
95 thirty-three genera, with twenty-four of these genera represented by a single species and the
96 remaining nine genera represented by between two and five species (S1 Table). Of the
97 thirty-three genera in the grass positive control sample, four were detected across both
98 markers, eight were detected by the *rbcL* marker and twelve by the ITS2 marker. The
99 remaining sequences were too similar to be identified to genus level (27% and 37% of reads
100 in the grass positive control samples could not be reliably assigned to genus level, using ITS2

101 and *rbcL* markers respectively). However, three of the genera not detected in the positive
102 control samples, despite being included, were detected in airborne samples (*Agrostis*,
103 *Anthoxanthum*, *Alopecurus*), likely reflecting higher local abundances of airborne pollen
104 (Supplementary Table 2). The number of species in the grass positive control is much higher
105 than the number of species predicted to contribute to airborne pollen concentrations
106 according to phenological studies^{11,12}. Differences in taxon diversity between the grass
107 positive control and the airborne samples will likely lead to differences in taxonomic
108 assignment due to taxon-specific PCR amplification biases¹³⁻¹⁵. While sample coverage (i.e.
109 number of reads) obtained for the grass positive control samples was comparable to the
110 airborne samples, the high diversity of the positive control and variation in the number of
111 species between genera may have led to a higher likelihood of amplification for certain
112 genera.

113

114 In order to check for cross-contamination between samples, an exotic plant positive control
115 sample was used containing DNA extracted from twenty-one tropical tree species samples
116 held at the National Botanic Garden of Wales (list S1). None of the genera identified in this
117 positive control were present in the experimental samples.



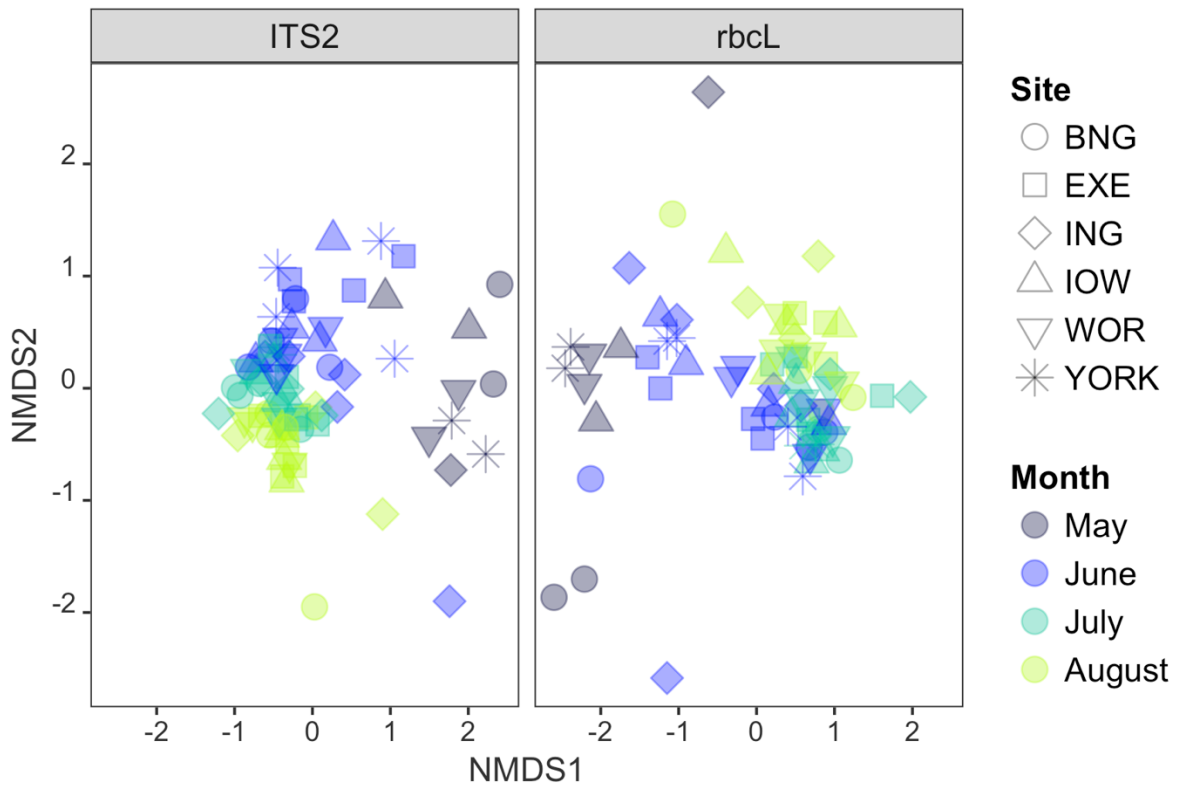
118

119 **Supplementary Figure 1** Abundance of airborne grass pollen taxa throughout the grass pollen season.

120 Abundance of rare grasses (expressed as proportion of total reads). Sampling sites are indicated in the
 121 top panel, followed by the marker used to identify grass pollen. Due to errors in sampling equipment,
 122 only 4 alternate weeks (out of a possible 7 alternate weeks) of samples were collected at the York
 123 sampling site. Note that the y axes differ between panels. Sampling sites are indicated in the right
 124 panel label abbreviated as follows: BNG = Bangor; EXE = Exeter; ING = Invergowrie; IOW = Isle of
 125 Wight; WOR = Worcester; YORK = York. A map of sampling locations and daily Poaceae pollen
 126 concentrations can be found in Figure 1.

127

128

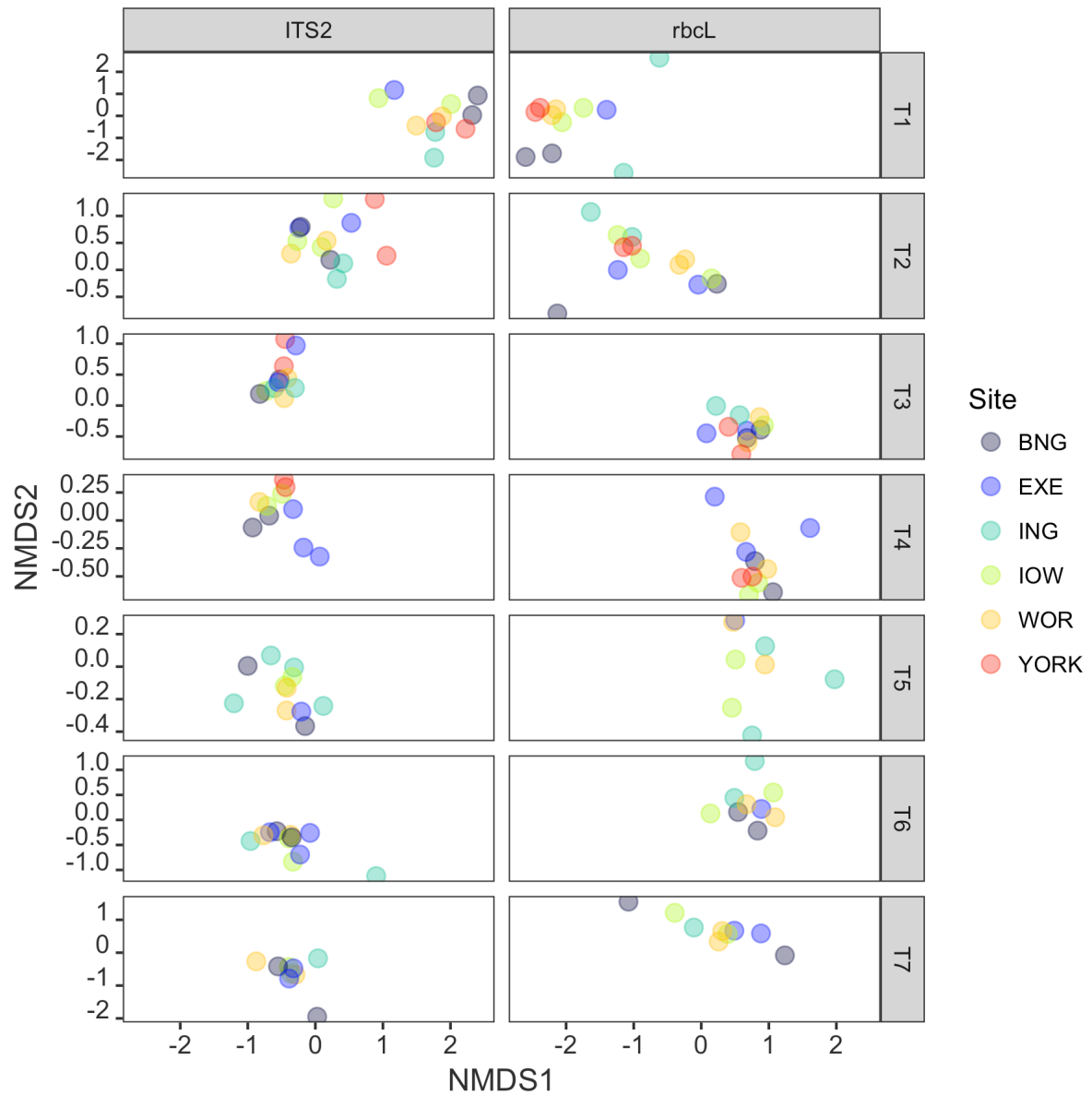


129

130 **Supplementary Figure 2** Grass community composition is structured by time across the UK. Non-
 131 metric multidimensional scaling (NMDS) ordination of grass community identified using ITS2 and *rbcL*
 132 barcode makers. Different shapes indicate the sampling location and colour of symbols indicate the
 133 month that pollen was collected. Refer to figure 2 for site name abbreviations.

134

135



136

137 **Supplementary Figure 3 Non-metric multidimensional scaling (NMDS) ordination of grass**

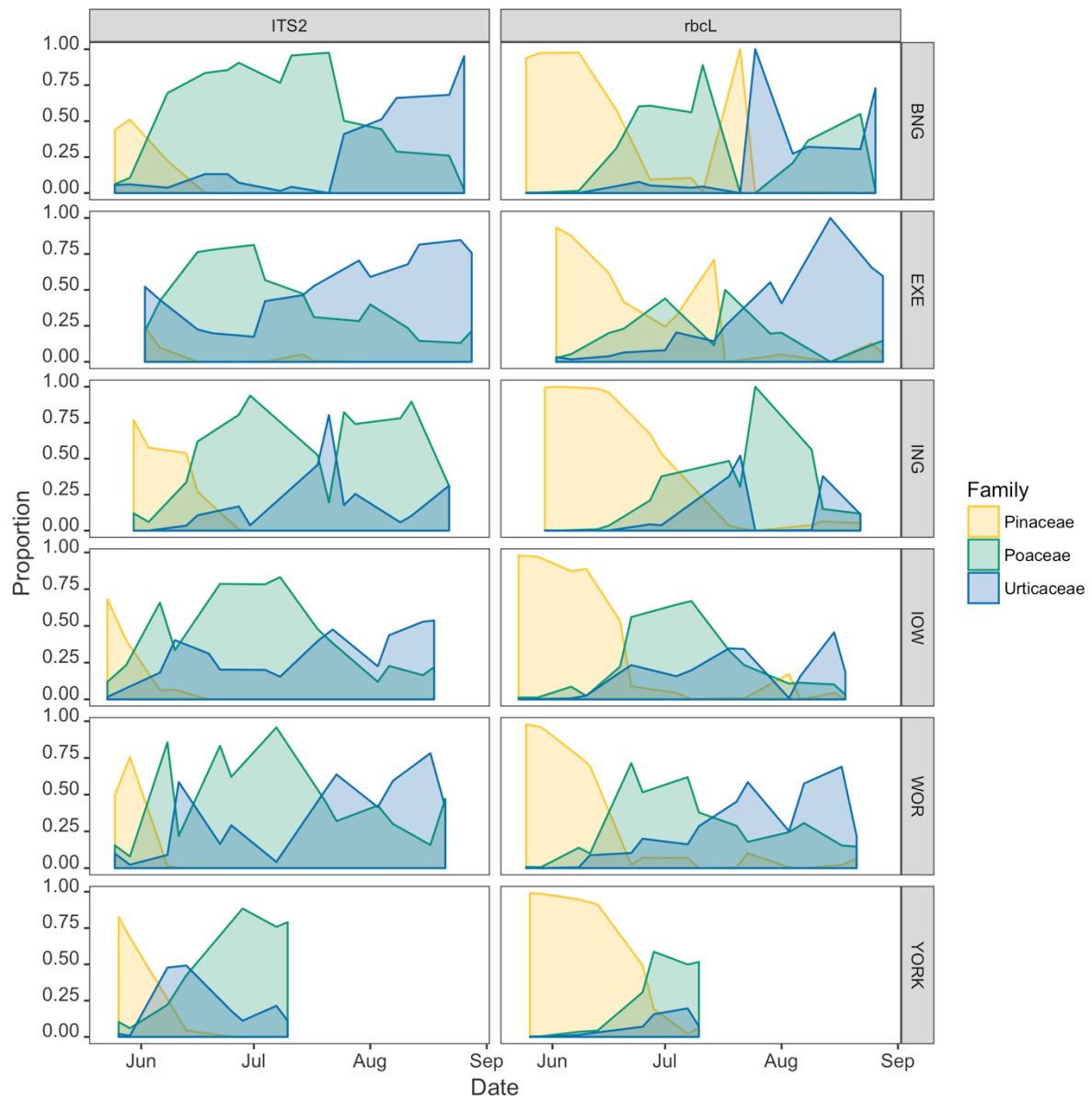
138 **community identified using ITS2 and *rbcL* barcode makers.** Each timepoint T1 to T7 indicates two

139 consecutive weeks when pollen samples were collected (May to August), in chronological order.

140 Coloured circles indicate sampling sites. Site labels are abbreviated as follows: BNG = Bangor; EXE =

141 Exeter; ING = Invergowrie; IOW = Isle of Wight; WOR = Worcester; YORK = York.

142



143

144 **Supplementary Figure 4 Abundance of the most common families of airborne pollen.** The three most

145 abundant families of airborne pollen, Pinaceae, Poaceae and Urticaceae, expressed as proportion of

146 total reads, collected between May and September. Markers used to identify pollen are stated in the

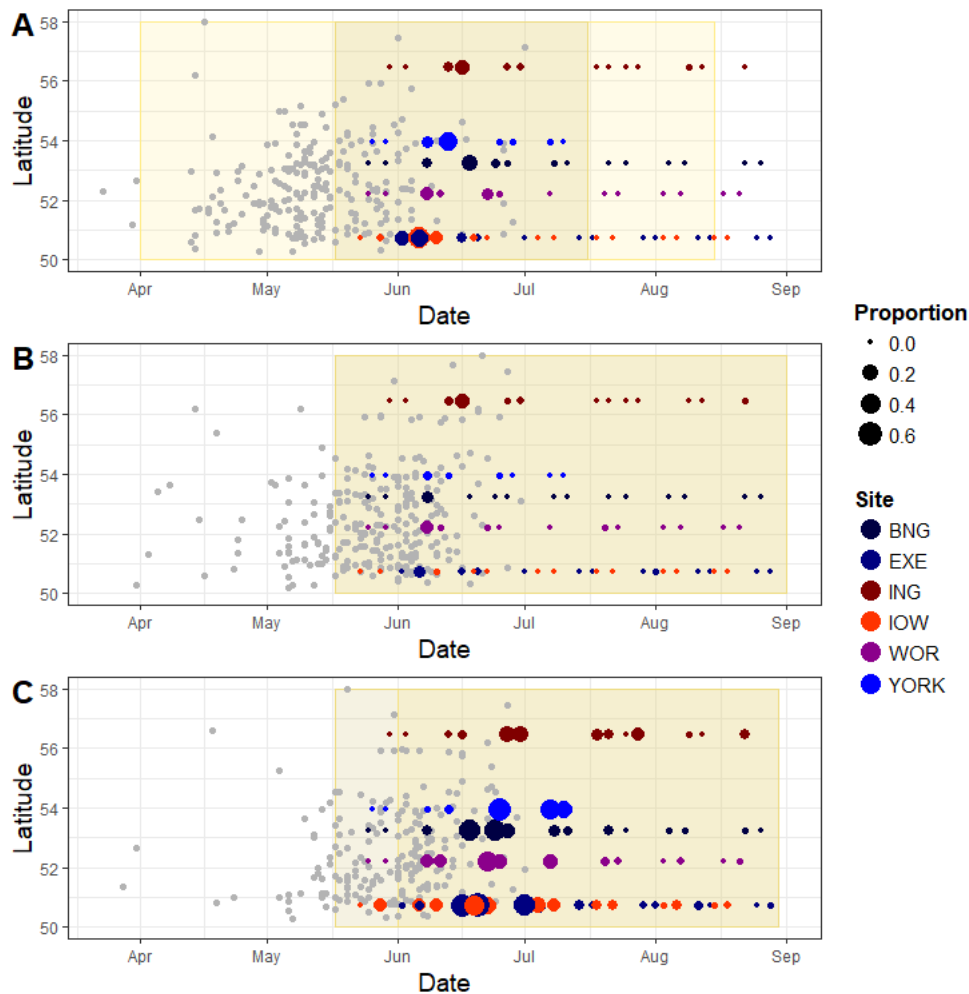
147 top panel label. Due to errors in sampling equipment, only 4 alternate weeks (out of a possible 7

148 alternate weeks) of samples were collected at the York sampling site. Sampling sites are indicated in

149 the right panel label abbreviated as follows: BNG = Bangor; EXE = Exeter; ING = Invergowrie; IOW =

150 Isle of Wight; WOR = Worcester; YORK = York. A map of sampling locations can be found in Figure 1.

151



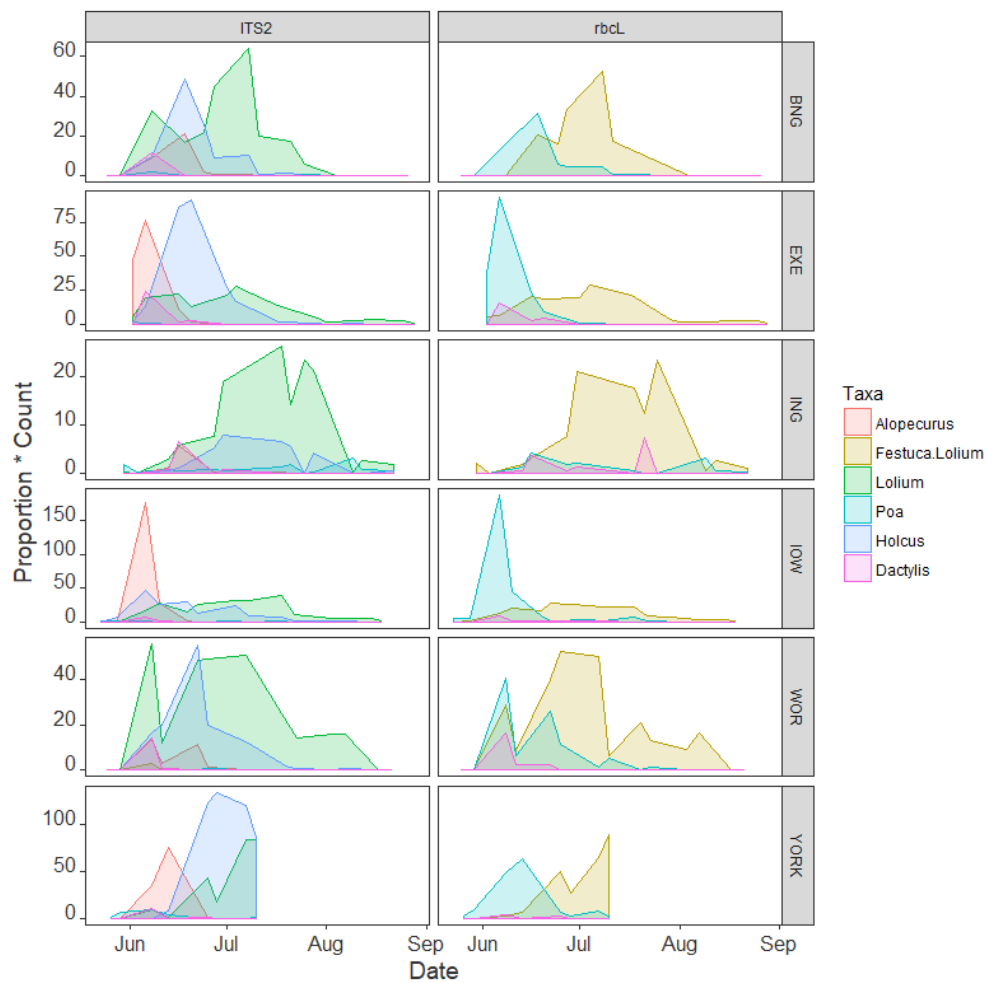
152

153 **Supplementary Figure 5 Airborne grass pollen observed 3-4 weeks after first flowering dates.**

154 Comparison of genus incidence in metabarcoding data with records of first flowering dates in 2016
 155 from the citizen science project Nature's Calendar (www.naturescalendar.org.uk) for (A) *Alopecurus*
 156 *pratensis*, (B) *Dactylis glomerata* and (C) *Holcus lanatus*. Each grey point represents the earliest time
 157 of flower heading as observed by a participant in the project. Coloured points represent
 158 metabarcoding samples, with the size of the point representing the proportion of total reads assigned
 159 to the relevant genus. Yellow shaded areas represent the expected flowering period as described in ¹⁰,
 160 with darker shades showing the 'main' flowering period.

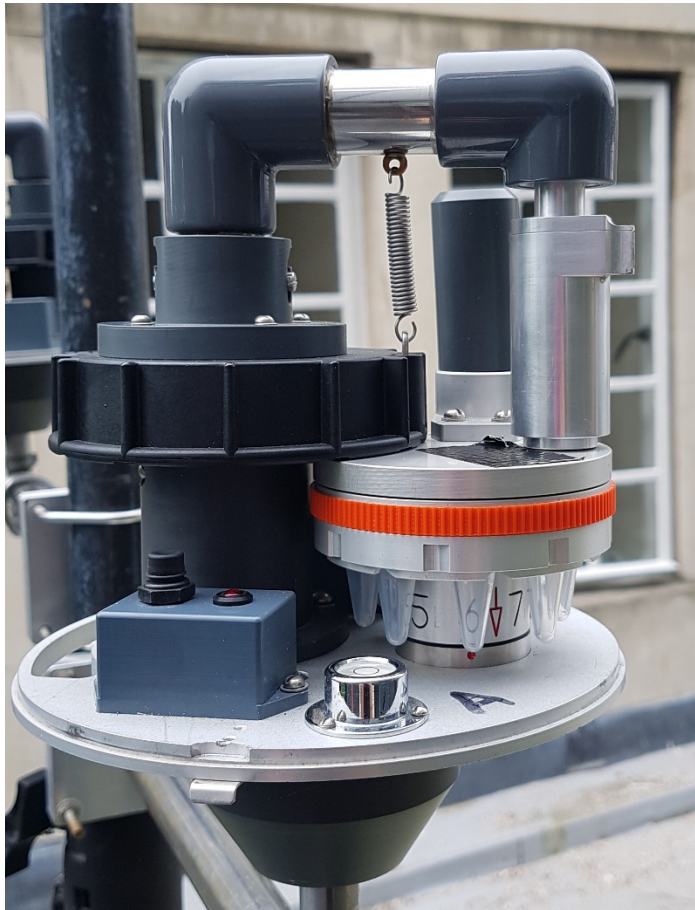
161

162



163

164 **Supplementary Figure 6** Relative abundance of the five most abundant grasses at genus level,
 165 normalized according to airborne pollen concentration data. Relative abundances were calculated as
 166 a proportion of reads assigned to Poaceae, rather than of reads as a whole, then multiplied by mean
 167 pollen concentration across the three pooled days. Markers used to identify grass pollen are stated in
 168 the top panel label. Due to errors in sampling equipment, only 4 alternate weeks (out of a possible 7
 169 alternate weeks) of samples were collected at the York sampling site. Sampling sites are indicated in
 170 the right panel label abbreviated as follows: BNG = Bangor; EXE = Exeter; ING = Invergowrie; IOW =
 171 Isle of Wight; WOR = Worcester; YORK = York.

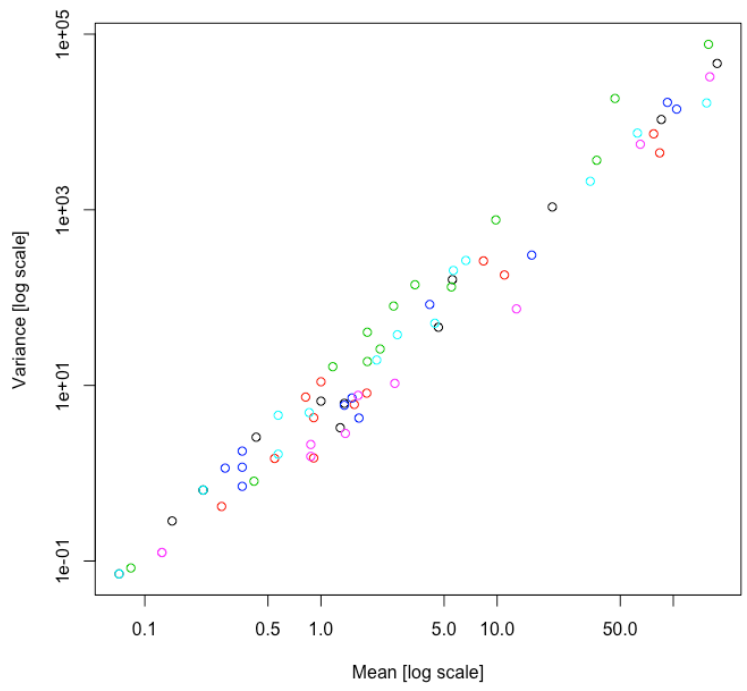


172

173 **Supplementary Figure 7** Photograph of 1.5 ml microcentrifuge tubes mounted onto carousel on
174 Burkard Automatic Multi-Vial Cyclone Sampler. Author provided.

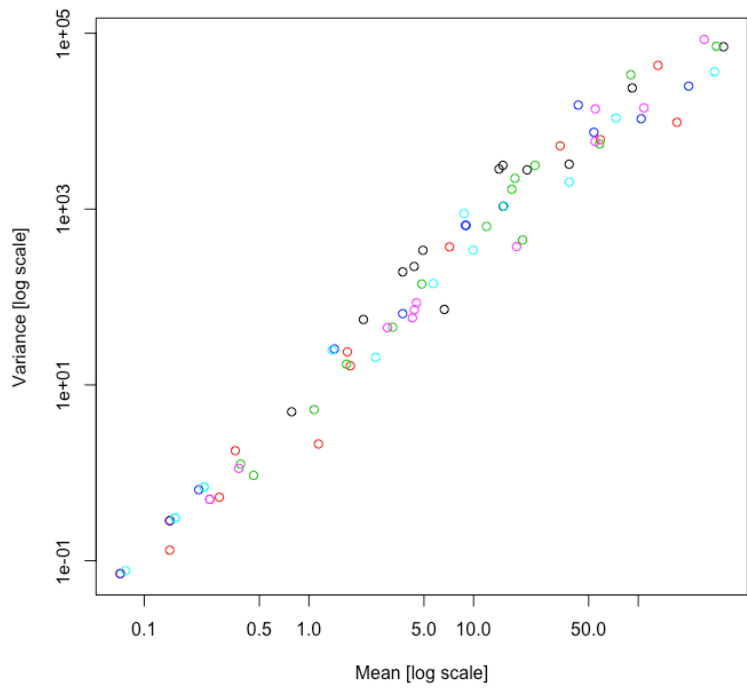
175

176 A)



177

178 B)

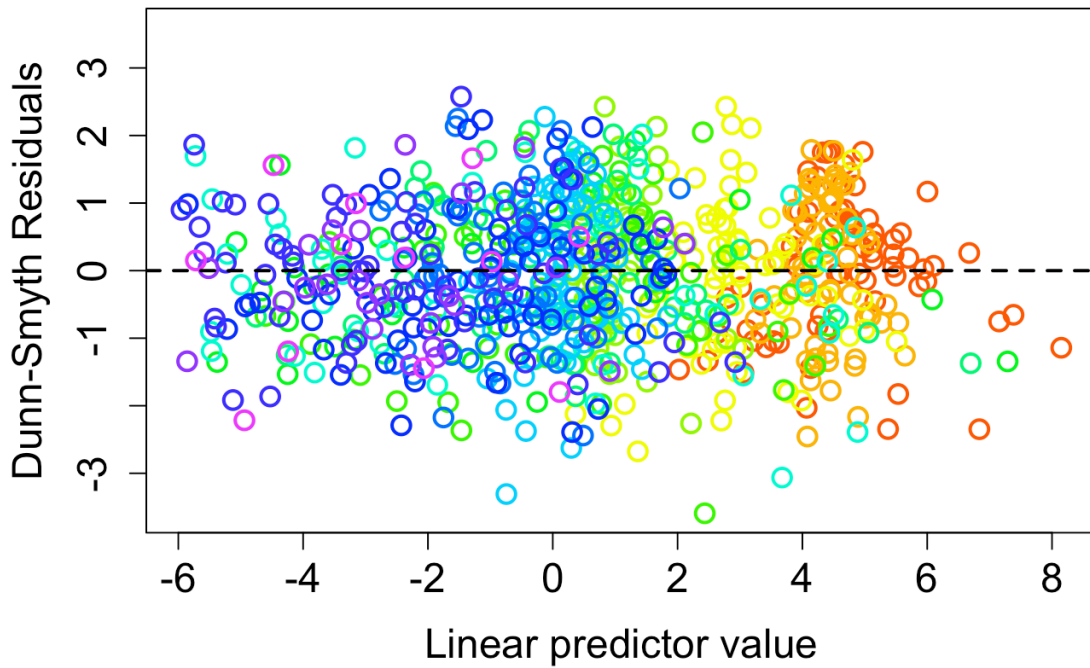


179

180

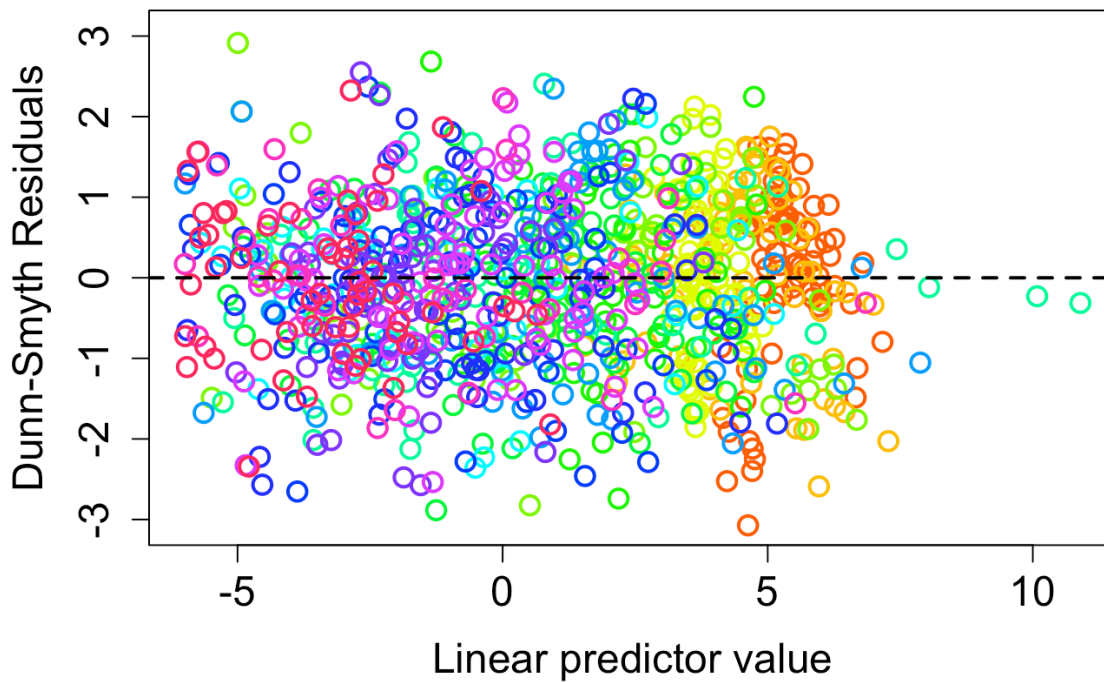
181 **Supplementary Figure 8** There is a strong relationship between the mean proportion of sequences
182 and the variance of the proportion of sequences from each sampling site using both A) *rbcL* and B)
183 ITS2 markers. Coloured circles denote sampling site. The plots were produced using the `meanvar.plot`
184 function in the `mvabund` package in R (21).

185 A)



186

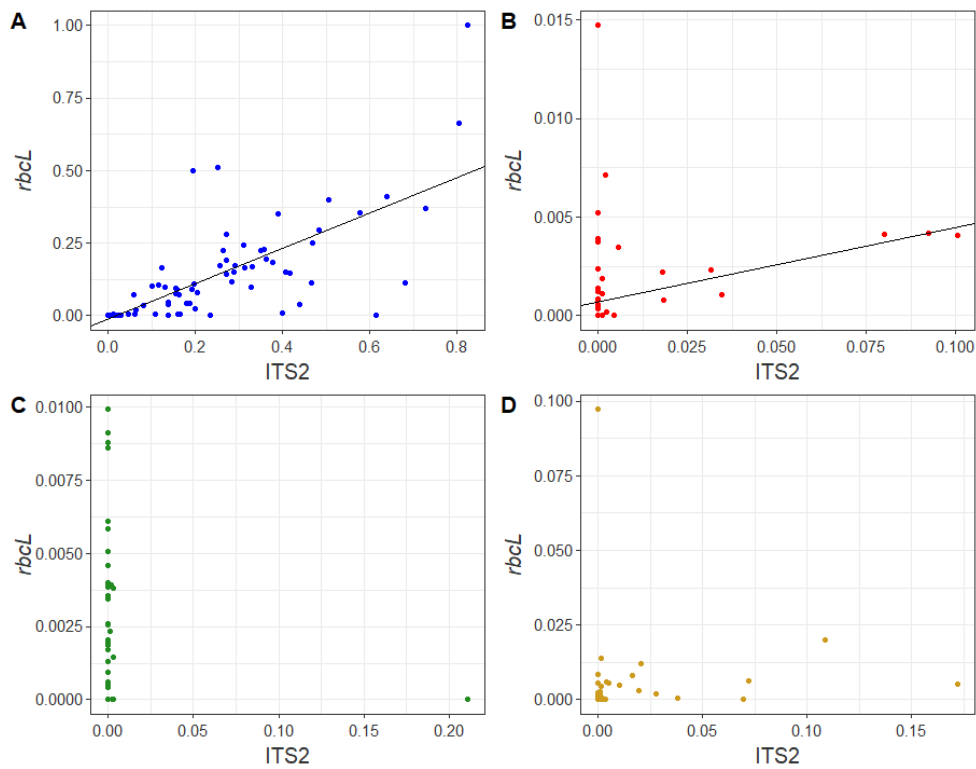
187 B)



188

189 **Supplementary Figure 9** Scatter plot of linear predictor values and the residuals output from the
190 models selected to analyse the abundance data produced by the A) *rbcL* marker and B) ITS2 marker.
191 Little association suggests that the models selected are plausible and the mean-variance assumption
192 of the negative binomial regression is correct. Coloured circles denote different genera in the

193 abundance data. The plots were produced using the plot.manyglm function in the mvabund package
194 in R (21).



195

196 **Supplementary Figure 10** Correlations between proportions of reads made up by the same genus in
197 the two marker gene datasets. All four genera present in both datasets are shown: (A)
198 *Lolium/Festuca*, (B) *Anthoxanthum*, (C) *Avena*, and (D) *Dactylis*. For cases where there was a
199 significant relationship between relative abundances in both datasets, black lines show the intercept
200 and slope.

201 List S1 Borneo plant taxa pooled for the exotic plant positive control.

202

203 *Aglaia sp.*

204 *Antidesma sp.*

205 *Baccaurea stipulata*

206 *Cynometra sp.*

207 *Dalbergia sp.*

208 *Dehaasia sp.*

209 *Dillenia excelsa*

210 *Diospyros sp.*

211 *Kleinhovia hospita*

212 *Lagerstroemia sp.*

213 *Lophopyxis sp.*

214 *Madhuca dubardii*

215 *Mallotus muticus*

216 *Microcos crassifolia*

217 *Pternandra sp.*

218 *Pterospermum macrocarpum*

219 *Syzygium sp.*

220 *Uncaria sp.*

221 *Urophyllum sp.*

222 *Vatica sp.*

223 *Xylosma sp.*

224

225 **Supplementary Table 1** Grass species pooled for the Grass Positive Control at equal volumes.

Grass positive control	Concentration of DNA (ng/μl)
<i>Agrostis canina</i>	0.121
<i>Agrostis capillaris</i>	1.29
<i>Agrostis gigantea</i>	9.48
<i>Agrostis stolonifera</i>	3.7
<i>Agrostis vinealis</i>	1.03
<i>Aira praecox</i>	0.8
<i>Alopecurus geniculatus</i>	1.18
<i>Alopecurus pratensis</i>	1.42
<i>Anisantha sterilis</i>	0.848
<i>Anthoxanthum odoratum</i>	0.804
<i>Arrhenatherum elatius</i>	1.36
<i>Brachypodium sylvaticum</i>	0.804
<i>Briza media</i>	2
<i>Bromopsis ramosa</i>	0.35
<i>Bromus hordeaceus</i>	0.098
<i>Catapodium rigidum</i>	3.96
<i>Cynosurus cristatus</i>	0.0736
<i>Dactylis glomerata</i>	13.8
<i>Danthonia decumbens</i>	1.34
<i>Deschampsia cespitosa</i>	2.52
<i>Deschampsia flexuosa</i>	0.648
<i>Elymus caninus</i>	1.62
<i>Elytrigia repens</i>	3.47
<i>Festuca arundinacea</i>	1.21
<i>Festuca gigantea</i>	1.32
<i>Festuca ovina</i>	0.592
<i>Festuca pratensis</i>	1.34
<i>Festuca rubra</i>	2.68
<i>Glyceria declinata</i>	0.226
<i>Glyceria fluitans</i>	0.892
<i>Glyceria maxima</i>	8.32
<i>Glyceria notata</i>	0.992
<i>Holcus lanatus</i>	0.42
<i>Holcus mollis</i>	below detection limit*
<i>Hordeum murinum</i>	0.476
<i>Hordeum secalinum</i>	0.416
<i>Lolium perenne</i>	0.452
<i>Milium effusum</i>	0.524
<i>Molinia caerulea</i>	2.24
<i>Nardus stricta</i>	0.246
<i>Phalaris arundinacea</i>	below detection limit*

<i>Phleum bertolonii</i>	0.444
<i>Phleum pratense</i>	18
<i>Phragmites australis</i>	13.2
<i>Poa annua</i>	0.0844
<i>Poa humilis</i>	2.37
<i>Poa pratensis</i>	1.13
<i>Poa trivialis</i>	below detection limit*
<i>Puccinellia distans</i>	11.1
<i>Trisetum flavescens</i>	0.736
<i>Triticum aestivum</i>	1.47
<i>Vulpia myuros</i>	0.199

226

227 * note these samples successfully amplified using *rbcL* and ITS2 primers shown in Supplementary

228 Table 5.

229 **Supplementary Table 2** Genera included in the grass positive control, and genera detected using
 230 metabarcoding of both marker genes in both the positive control and in actual aerial DNA extracts.
 231 Genera with a grey background were detected by at least one marker gene; genera with a white
 232 background were not.

Expected	rbcl- Control	ITS2- Control	rbcl- Samples	ITS2- Samples
<i>Agrostis</i>				<i>Agrostis</i>
<i>Aira</i>				
<i>Alopecurus</i>				<i>Alopecurus</i>
<i>Anisantha</i>				
<i>Anthoxanthum</i>			<i>Anthoxanthum</i>	<i>Anthoxanthum</i>
<i>Arrhenatherum</i>		<i>Arrhenatherum</i>		<i>Arrhenatherum</i>
<i>Avena</i>	<i>Avena</i>		<i>Avena</i>	<i>Avena</i>
<i>Brachypodium</i>				
<i>Briza</i>	<i>Briza</i>	<i>Briza/Bromus</i>	<i>Briza</i>	
<i>Bromopsis</i>				
<i>Bromus</i>		<i>Briza/Bromus</i>		
<i>Catapodium</i>				
<i>Cynosurus</i>		<i>Cynosurus</i>		<i>Cynosurus</i>
<i>Dactylis</i>	<i>Dactylis</i>	<i>Dactylis</i>	<i>Dactylis</i>	
<i>Danthonia</i>				
<i>Deschampsia</i>		<i>Deschampsia</i>		<i>Deschampsia</i>
<i>Elymus</i>				
<i>Elytrigia</i>				
<i>Festuca</i>	<i>Festuca/Lolium</i>	<i>Festuca/Lolium</i>	<i>Festuca/Lolium</i>	<i>Festuca/Lolium</i>
<i>Glyceria</i>		<i>Glyceria</i>		
<i>Holcus</i>				<i>Holcus</i>
<i>Hordeum</i>		<i>Hordeum</i>		<i>Hordeum</i>
<i>Lolium</i>		<i>Lolium</i>		<i>Lolium</i>
<i>Milium</i>				
<i>Molinia</i>	<i>Molinia</i>		<i>Molinia</i>	
<i>Nardus</i>				
<i>Phalaris</i>				
<i>Phleum</i>	<i>Phleum</i>		<i>Phleum</i>	
<i>Phragmites</i>				
<i>Poa</i>	<i>Poa</i>	<i>Poa</i>	<i>Poa</i>	<i>Poa</i>
<i>Puccinellia</i>				
<i>Trisetum</i>				
<i>Triticum</i>	<i>Triticum</i>			
<i>Vulpia</i>				

233

234 **Supplementary Table 3** Latitude and longitude of each pollen sampling site.

Site Name	Abbreviation	Latitude	Longitude
Bangor	BNG	53.2300	-4.1300
Exeter	EXE	50.7365	-3.5322
Invergowrie	ING	56.4576	-3.0687
Isle of Wight	IOW	50.7111	-1.3009
Worcestershire	WORK	52.1976	-2.2430
York	YORK	53.9484	-1.0535

235

236 **Supplementary Table 4** Sample collection dates of each sequenced air sample. Three consecutive days
 237 of air samples were pooled during DNA extraction (note that sample ING_w2_p2, three consecutive
 238 samples were unavailable due to sampling error and the next sampling day was selected for pooling).
 239 The mean pollen concentration for the three pooled days and the index i5 and i7 sequence for
 240 demultiplexing is shown here.

Sample	Index i5 and i7 Sequence	Week	Pool	Site	Collection date (2016)	Mean pollen conc. (grains m ⁻³)
BNG_w1_p1	CAAGTCGT	1	1	BNG	25 May - 28 May	61.7
BNG_w1_p2	TAACGTCG	1	2	BNG	29 May - 01 Jun	27
BNG_w2_p1	CTGTATGC	2	1	BNG	08 Jun - 11 Jun	NA
BNG_w2_p2	TGCTTGCT	2	2	BNG	18 Jun - 21 Jun	NA
BNG_w3_p1	GTAGTACC	3	1	BNG	24 Jun - 27 Jun	NA
BNG_w3_p2	AAGTCCTC	3	2	BNG	27 Jun - 30 Jun	NA
BNG_w4_p1	GCATAACG	4	1	BNG	08 Jul - 11 Jul	35.3
BNG_w4_p2	ATAGTCGG	4	2	BNG	11 Jul - 14 Jul	18.3
BNG_w5_p1	TAGGAGCT	5	1	BNG	21 Jul - 24 Jul	5.7
BNG_w5_p2	AGGTGTTG	5	2	BNG	25 Jul - 28 Jul	2
BNG_w6_p1	CATTGACG	6	1	BNG	04 Aug - 07 Aug	4.3
BNG_w6_p2	CCACAACA	6	2	BNG	08 Aug - 11 Aug	1.3
BNG_w7_p1	TCTAGGAG	7	1	BNG	22 Aug - 25 Aug	3.3
BNG_w7_p2	TTGCTTGG	7	2	BNG	26 Aug - 29 Aug	2.3
EXE_w1_p1	TGATCACG	1	1	EXE	02 Jun - 05 Jun	63
EXE_w1_p2	TCTGGACA	1	2	EXE	06 Jun - 09 Jun	139.3
EXE_w2_p1	CAGTGCTT	2	1	EXE	16 Jun - 19 Jun	126
EXE_w2_p2	ATAGTCC	2	2	EXE	20 Jun - 23 Jun	124.7
EXE_w3_p1	CTGTACCA	3	1	EXE	01 Jul - 04 Jul	52.3
EXE_w3_p2	AAGCATCG	3	2	EXE	04 Jul - 07 Jul	61.3
EXE_w4_p1	CCTGTCAA	4	1	EXE	14 Jul - 17 Jul	56
EXE_w4_p2	AATGGTCG	4	2	EXE	17 Jul - 20 Jul	21.7
EXE_w5_p1	CTCCTGAA	5	1	EXE	29 Jul - 01 Aug	7
EXE_w5_p2	GACGAAct	5	2	EXE	01 Aug - 04 Aug	2.7
EXE_w6_p1	GGTCGTAT	6	1	EXE	11 Aug - 14 Aug	2.3
EXE_w6_p2	AAGTGCAG	6	2	EXE	14 Aug - 17 Aug	3.3
EXE_w7_p1	CCATGAAC	7	1	EXE	25 Aug - 28 Aug	3
EXE_w7_p2	TACTAGCG	7	2	EXE	28 Aug - 31 Aug	0.7
ING_w1_p1	GTGATCCA	1	1	ING	30 May - 02 Jun	2
ING_w1_p2	ATAACGCC	1	2	ING	03 Jun - 06 Jun	1
ING_w2_p1	ACCATAGG	2	1	ING	13 Jun - 16 Jun	7

ING_w2_p2	AGTTCGCA	2	2	ING	16 Jun, 19 Jun, 20 Jun	19.3
ING_w3_p1	CAACTTGG	3	1	ING	27 Jun - 30 Jun	19
ING_w3_p2	CGCAATGT	3	2	ING	30 Jun - 03 Jul	38
ING_w4_p1	GGCTCAAT	4	1	ING	18 Jul - 21 Jul	67.7
ING_w4_p2	GACTTG TG	4	2	ING	21 Jul - 24 Jul	22.7
ING_w5_p1	GCTACAAC	5	1	ING	25 Jul - 28 Jul	19.7
ING_w5_p2	GGTACGAA	5	2	ING	28 Jul - 31 Jul	27.3
ING_w6_p1	ACGAACGA	6	1	ING	09 Aug - 12 Aug	3.3
ING_w6_p2	AACACTGG	6	2	ING	12 Aug - 15 Aug	3.3
ING_w7_p1	TGGATGGT	7	1	ING	22 Aug - 25 Aug	3
IOW_w1_p1	TACTGCTC	1	1	IOW	23 May - 26 May	7.3
IOW_w1_p2	CTTCGCAA	1	2	IOW	28 May - 31 May	13.7
IOW_w2_p1	GATCAAGG	2	1	IOW	06 Jun - 09 Jun	253
IOW_w2_p2	GGCGAATA	2	2	IOW	10 Jun - 13 Jun	84
IOW_w3_p1	CAACGAGT	3	1	IOW	19 Jun - 22 Jun	57.7
IOW_w3_p2	ATCGGAGA	3	2	IOW	22 Jun - 25 Jun	39.7
IOW_w4_p1	TGTTCCGT	4	1	IOW	04 Jul - 07 Jul	86
IOW_w4_p2	ATCCACGA	4	2	IOW	08 Jul - 11 Jul	52.3
IOW_w5_p1	TCACCTAG	5	1	IOW	18 Jul - 21 Jul	64.3
IOW_w5_p2	AGGATAGC	5	2	IOW	22 Jul - 25 Jul	13
IOW_w6_p1	ATGACAGG	6	1	IOW	03 Aug - 06 Aug	5
IOW_w6_p2	CCGTTATG	6	2	IOW	06 Aug - 09 Aug	6.7
IOW_w7_p1	ACCTCTTC	7	1	IOW	15 Aug - 18 Aug	4.3
IOW_w7_p2	ACAGAGGT	7	2	IOW	18 Aug - 21 Aug	2
WOR_w1_p1	CGCTACAT	1	1	WOR	25 May - 28 May	0
WOR_w1_p2	AACCAGAG	1	2	WOR	29 May - 01 Jun	0
WOR_w2_p1	GCAATTCC	2	1	WOR	08 Jun - 11 Jun	114.7
WOR_w2_p2	AGCCGTAA	2	2	WOR	11 Jun - 14 Jun	40.7
WOR_w3_p1	AACAAGGC	3	1	WOR	22 Jun - 25 Jun	131
WOR_w3_p2	GAGCAATC	3	2	WOR	25 Jun - 28 Jun	78.7
WOR_w4_p1	AGTATGCC	4	1	WOR	07 Jul - 10 Jul	76
WOR_w4_p2	TCGATGAC	4	2	WOR	10 Jul - 13 Jul	16
WOR_w5_p1	GATACCTG	5	1	WOR	20 Jul - 23 Jul	26.3
WOR_w5_p2	ACCGACAA	5	2	WOR	23 Jul - 26 Jul	16
WOR_w6_p1	ACGAATCC	6	1	WOR	03 Aug - 06 Aug	0
WOR_w6_p2	TCGAGAGT	6	2	WOR	07 Aug - 10 Aug	0
WOR_w7_p1	GTTCTTCG	7	1	WOR	17 Aug - 20 Aug	0
WOR_w7_p2	CCTTCCAT	7	2	WOR	21 Aug - 24 Aug	0
YORK_w1_p1	TCCACGTT	1	1	YORK	26 May - 29 May	3
YORK_w1_p2	TTACCGAC	1	2	YORK	29 May - 01 Jun	9.7
YORK_w2_p1	TTCGCCAT	2	1	YORK	08 Jun - 11 Jun	84.7

YORK_w2_p2	TATGGCAC	2	2	YORK	13 Jun - 16 Jun	96.7
YORK_w3_p1	CGCGTATT	3	1	YORK	25 Jun - 28 Jun	178
YORK_w3_p2	AGCCTATC	3	2	YORK	28 Jun - 01 Jul	157
YORK_w4_p1	GACACAGT	4	1	YORK	07 Jul - 10 Jul	234.3
YORK_w4_p2	GAGAGTAC	4	2	YORK	10 Jul - 13 Jul	245.3
Negative control 1	CCACTAAG	-	-	-	-	-
Negative control 2	CCACATTG	-	-	-	-	-
Negative control 3	CCGATGTA	-	-	-	-	-
Negative control 4	CTCGGTAA	-	-	-	-	-
Negative control 5	AACCGTGT	-	-	-	-	-
Negative control 6	CGGTTGTT	-	-	-	-	-
Negative control 7	CTAGCAGT	-	-	-	-	-
Negative control 8	ACAACAGC	-	-	-	-	-
Negative control 9	GATTGTCC	-	-	-	-	-
Exotic positive control	ACAGGCAT	-	-	-	-	-
Grass positive control	TTCGTACG	-	-	-	-	-

241

242 **Supplementary Table 5** Primer sequences used in library preparation. Round 1 PCR primer sequences
 243 contain forward or reverse template primer, the forward primer sequence contains a series of N's in
 244 order to improve clustering and cluster detection on MiSeq sequencing. Round 1 and round 2
 245 sequences contain complementary universal tails. Round 2 PRC primers sequences also contain the
 246 P5 or P7 Illumina adaptors and an 8 bp unique index on both the forward and reverse primers used
 247 for demultiplexing samples (see Supplementary Table 4 for index i5 and i7 sequence).

Round 1 PCR
Forward Universal Tail - NNNNNN - Template Specific Primer <i>rbclLaF</i> [CACTCTTTCCCTACACGACGCTCTCCGATCT]-[NNNNNN]-[ATGTCACCACAAACAGAGACTAAAGC]
Reverse Universal Tail - Template Specific Primer <i>rbclR506</i> [GTGACTGGAGTTCAGACGTGTGCTCTCCGATCT]-[AGGGGACGACCATACTTGTTCA]
Forward Universal Tail - NNNNNN - Template Specific Primer <i>ITS2F</i> [CACTCTTTCCCTACACGACGCTCTCCGATCT]-[NNNNNN]-[ATGCGATACTGGTGTGAAT]
Reverse Universal Tail - Template Specific Primer <i>ITS3R</i> [GTGACTGGAGTTCAGACGTGTGCTCTCCGATCT]- [GACGCTTCTCCAGACTACAAT]
Round 2 PCR
P5 Illumina adapter - i5 index - Forward Universal Tail [AATGATACGGCGACCACCGAGATCTACAC]-[i5 index]-[CACTCTTTCCCTACACGACGCTC]
P7 Illumina adapter - i7 index - Reverse Universal Tail [CAAGCAGAAGACGGCATACGAGAT]-[i7 index]-[GTGACTGGAGTTCAGACGTGTGCTC]

248

249 **Supplementary Table 6** Output from the models used to analyse the *rbcL* and ITS2 data, without the
250 York sampling site.

<i>rbcL</i>	Res.Df	Df.diff	LR	Pr(>Dev)
Time	71	1	46.71	0.001
Latitude	70	1	26.4	0.024
Longitude	69	1	27.1	0.014
Time:Latitude	68	1	47.36	0.001

ITS2	Res.Df	Df.diff	LR	Pr(>Dev)
Time	74	1	128.8	0.001
Latitude	73	1	73.2	0.001
Longitude	72	1	33.0	0.011
Time:Latitude	71	1	34.2	0.04

251

252

253 **References**

- 254 1. D’Amato, G. *et al.* in *Allergy and Allergen Immunotherapy* **62**, 261–279 (2017).
- 255 2. Peel, R. G. *et al.* Seasonal variation in diurnal atmospheric grass pollen concentration profiles.
256 *Biogeosciences* **11**, 821–832 (2014).
- 257 3. Cebrino, J., Galán, C. & Domínguez-Vilches, E. Aerobiological and phenological study of the
258 main Poaceae species in Córdoba City (Spain) and the surrounding hills. *Aerobiologia* **32**, 595–
259 606 (2016).
- 260 4. León-Ruiz, E., Alcázar, P., Domínguez-Vilches, E. & Galán, C. Study of Poaceae phenology in a
261 Mediterranean climate. Which species contribute most to airborne pollen counts?
262 doi:10.1007/s10453-010-9174-y
- 263 5. Chapman, D. S. Greater phenological sensitivity to temperature on higher Scottish mountains:
264 new insights from remote sensing. *Global Change Biology* **19**, 3463–3471 (2013).
- 265 6. Goldberg, C., Buch, H., Moseholm, L. & Weeke, E. R. Airborne Pollen Records in Denmark,
266 1977–1986. *Grana* **27**, 209–217 (1988).
- 267 7. Skjøth, C. A., Baker, P., Sadyś, M. & Adams-Groom, B. Pollen from alder (*Alnus* sp.), birch
268 (*Betula* sp.) and oak (*Quercus* sp.) in the UK originate from small woodlands. *Urban Climate*
269 **14**, 414–428 (2015).
- 270 8. Cheng, T. *et al.* Barcoding the kingdom Plantae: new PCR primers for ITS regions of plants
271 with improved universality and specificity. *Molecular Ecology Resources* **16**, 138–149 (2015).
- 272 9. Hollingsworth, M. L. *et al.* Selecting barcoding loci for plants: evaluation of seven candidate
273 loci with species-level sampling in three divergent groups of land plants. *Mol. Ecol. Resour.* **9**,
274 439–457 (2009).
- 275 10. Alberdi, A., Aizpurua, O., Gilbert, M. T. P. & Bohmann, K. Scrutinizing key steps for reliable
276 metabarcoding of environmental samples. *Methods in Ecology and Evolution* **9**, 134–147
277 (2018).
- 278 11. Ghitarrini, S., Galán, C., Frenguelli, G. & Tedeschini, E. Phenological analysis of grasses
279 (Poaceae) as a support for the dissection of their pollen season in Perugia (Central Italy).
280 *Aerobiologia* **33**, 339–349 (2017).
- 281 12. Romero-Morte, J., Rojo, J., Rivero, R., Fernández-González, F. & Pérez-Badia, R. Standardised
282 index for measuring atmospheric grass-pollen emission. *Science of the Total Environment*
283 **612**, 180–191 (2018).
- 284 13. Bista, I. *et al.* Performance of amplicon and shotgun sequencing for accurate biomass
285 estimation in invertebrate community samples. *Molecular Ecology Resources* 1–15 (2018).
- 286 14. Pawluczyk, M. *et al.* Quantitative evaluation of bias in PCR amplification and next-generation
287 sequencing derived from metabarcoding samples. *Analytical and Bioanalytical Chemistry* **407**,
288 1841–1848 (2015).
- 289 15. Suzuki, M. T. & Giovannoni, S. J. Bias caused by template annealing in the amplification of
290 mixtures of 16S rRNA genes by PCR. *Applied and Environmental Microbiology* **62**, 625–630
291 (1996).

292