

**This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.**

**Author(s):** Zhang, Hai-Yang; Bissett, Andrew; Aguilar-Trigueros, Carlos A.; Liu, Hong-Wei; Powell, Jeff R.

**Title:** Fungal genome size and composition reflect ecological strategies along soil fertility gradients

**Year:** 2023

**Version:** Accepted version (Final draft)

**Copyright:** © 2023 Wiley-Blackwell

**Rights:** In Copyright

**Rights url:** <http://rightsstatements.org/page/InC/1.0/?language=en>

**Please cite the original version:**

Zhang, H., Bissett, A., Aguilar-Trigueros, C. A., Liu, H., & Powell, J. R. (2023). Fungal genome size and composition reflect ecological strategies along soil fertility gradients. *Ecology Letters*, 26(7), 1108-1118. <https://doi.org/10.1111/ele.14224>

# ECOLOGY LETTERS

## Fungal genome size and composition reflect ecological strategies along soil fertility gradients

Journal:	<i>Ecology Letters</i>
Manuscript ID	ELE-01110-2022.R4
Manuscript Type:	Letter
Date Submitted by the Author:	09-Mar-2023
Complete List of Authors:	Zhang, Hai-Yang; College of Life Sciences, Hebei University; Western Sydney University Hawkesbury Institute for the Environment, School of Science Bissett, Andrew; CSIRO, Plant Industry Aguilar-Trigueros, Carlos; Free University of Berlin; Berlin-Brandenburg Institute of Advanced Biodiversity Research Liu, Hongwei; Western Sydney University Hawkesbury Institute for the Environment Powell, Jeff; Hawkesbury Institute for the Environment, Soil Biology and Genomics

1  
2  
3 **1 Fungal genome size and composition reflect ecological strategies along soil fertility**  
4 **2 gradients**

5  
6 3 Authors: Hai-Yang Zhang<sup>1, 2\*</sup>, Andrew Bissett<sup>3</sup>, Carlos A. Aguilar-Trigueros<sup>4, 5</sup>, Hong-Wei  
7 4 Liu<sup>2</sup>, Jeff R. Powell<sup>2</sup>

8  
9 5 Affiliations:

10 6 <sup>1</sup>College of Life Sciences, Hebei University, Baoding, China

11 7 <sup>2</sup>Hawkesbury Institute for the Environment, Western Sydney University, Penrith, New South  
12 8 Wales, Australia

13 9 <sup>3</sup>Oceans and Atmosphere, CSIRO, Hobart, TAS 7000, Australia

14 10 <sup>4</sup>Department of Biological and Environmental Science, University of Jyväskylä, P.O. Box 35,  
15 11 FI-40014, Finland

16 12 <sup>5</sup>Freie Universität Berlin, Institute of Biology, 14195 Berlin, Germany

17 13 e-mail address of all authors:

18 14 [Haiyang.zhang@westernsydney.edu.au](mailto:Haiyang.zhang@westernsydney.edu.au); [Andrew.Bissett@csiro.au](mailto:Andrew.Bissett@csiro.au); [calgit@gmail.com](mailto:calgit@gmail.com);

19 15 [hongwei.liu@westernsydney.edu.au](mailto:hongwei.liu@westernsydney.edu.au); [jeff.powell@westernsydney.edu.au](mailto:jeff.powell@westernsydney.edu.au)

20 16 **Data accessibility statement:** All datasets, codes, and associated metadata underlying the  
21 17 study are available at Figshare (<https://doi.org/10.6084/m9.figshare.20277759.v4>).

22 18 **Author Contributions**

23 19 HZ and JP conceived the study and analysed the data. HZ compiled the fungal traits dataset.  
24 20 AB led the AMI project and JP performed taxonomic assignments for zOTUs. HZ performed  
25 21 the statistical analyses with help from JP. HZ, HL and JP wrote the manuscript, and all co-  
26 22 authors provided input on subsequent drafts.

27 23 **Running title:** Fungal genomic traits shift with nutrient limitation

28 24 **Keywords:** Genome size diversity, fungal base-pair, mycelium chemistry, species  
29 25 distribution, fungal guild, pathogen, symbiotic fungi, nutrient limitation

30 26 **Type of article:** *Letter*

31 27 **Word count**

32 28 **Abstract:** 148. **Text body** 4992: (Introduction: 1032; Methods 1520; Results: 811;  
33 29 Discussion: 1629). **Number of references:** 65. **Figures and tables:** figure: 4; table:1.

34 30 **Supplementary information:** figure: 7; table: 6.

35 31 \*Corresponding author:

36 32 Dr. Haiyang Zhang, Tel: +61 2 4570 1082;

1  
2  
3 33 E-mail: [zhanghaiyang798@gmail.com](mailto:zhanghaiyang798@gmail.com);  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 34 **Abstract**  
4

5 35 Genomic traits reflect the evolutionary processes that have led to ecological variation among  
6 36 extant organisms, including variation in how they acquire and use resources. Soil fungi have  
7 37 diverse nutritional strategies and exhibit extensive variation in fitness along resource  
8 38 gradients. We tested for trade-offs in genomic traits with mycelial nutritional traits and  
9 39 hypothesize that such trade-offs differ among fungal guilds as they reflect contrasting  
10 40 resource exploitation and habitat preferences. We found species with large genomes  
11 41 exhibited nutrient-poor mycelium and low GC content. These patterns were observed across  
12 42 fungal guilds but with varying explanatory power. We then matched trait data to fungal  
13 43 species observed in 463 Australian grassland, woodland and forest soil samples. Fungi with  
14 44 large genomes and lower GC content dominated in nutrient-poor soils, associated with shifts  
15 45 in guild composition and with species turnover within guilds. These findings highlight  
16 46 fundamental mechanisms that underpin successful ecological strategies for soil fungi.  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## 47 **Introduction**

48 A pervasive challenge in ecology is to understand and predict how organisms adapt to  
49 their local environments and respond to environmental changes, with recent efforts  
50 attempting to address this challenge with microbial organisms (Widder *et al.* 2016; Fierer  
51 2017). There has been growing interest in using trait-based approaches to study microbial  
52 ecological strategies and their responses to local and global environmental factors due to the  
53 direct ecological implications of functional traits on organism's fitness and tolerance to biotic  
54 and abiotic factors (Kraft *et al.* 2008; Martiny *et al.* 2015). This is particularly the case for  
55 fungi (Aguilar-Trigueros *et al.* 2014; Pöhlme *et al.* 2020; Zanne *et al.* 2020), which are  
56 ubiquitous organisms that control critical aspects of ecosystem functioning including plant  
57 nutrition, host health and fitness and nutrient cycling in terrestrial ecosystems (Tedersoo *et al.*  
58 2014; Naranjo-Ortiz & Gabaldón 2019). Most trait-based studies of fungi have targeted  
59 phenotypic traits related to mycelium construction (e.g., hyphal extension and branching  
60 rates, hyphal chemistry; Camenzind *et al.* 2020; Camenzind *et al.* 2021), reproduction (e.g.,  
61 spore size and sporulation behaviour; Aguilar-Trigueros *et al.* 2019; Chan *et al.* 2019, 2020)  
62 and resource uptake (gene expression of enzymatic pathways; Talbot *et al.* 2015). However,  
63 these traits have been studied largely (if not entirely) under highly controlled conditions,  
64 which limits our understanding of how fungal traits can shape species distributions in their  
65 natural environment.

66 All living organisms require nutrients to grow and reproduce. Genomic traits such as  
67 genome size (DNA content of the entire genome) and genomic DNA base composition of  
68 guanine-cytosine (GC) content in organisms are associated with their adaptation strategies to  
69 changes in nutrient availability in the environment (Giovannoni *et al.* 2014; Shenhav & Zeevi  
70 2020). Nucleic acids are among the most cellular nutrient-rich molecules, with carbon (C) :  
71 nitrogen (N) : phosphorus (P) stoichiometry being 12 : 4 : 1 (Sterner & Elser 2002). Thus,  
72 species with large genomes, which are more demanding and costly to build and maintain than  
73 small genomes, are expected to be less competitive when N and P (and in some cases other  
74 nutrients) are limited (Leitch & Leitch 2013). Evidence from plant studies suggests that  
75 species with large genomes are more likely to successfully compete and dominate in natural  
76 plant communities when levels of resources such as N and P are high in soil (Šmarda *et al.*  
77 2013; Guignard *et al.* 2016). The same patterns exist in bacteria: species with small genomes  
78 possess growth advantages in nutrient-depleted environments due to their reduced nutrient  
79 requirements than species with large genomes (Giovannoni *et al.* 2014). These studies  
80 highlight that mapping microbial species and their genomic traits along resource gradients

1  
2  
3 81 can provide insight into their eco-evolutionary adaptations to resource limitations (Gudelj *et*  
4 82 *al.* 2010; Barberán *et al.* 2014).

5  
6 83 Trait-trait correlations reflect fundamental constraints related to organisms' ecological  
7 84 strategies (Westoby *et al.* 2002; Reich *et al.* 2003) because one must optimize their  
8 85 performance by strategically allocating limited resources to different features (Bazzaz &  
9 86 Grace 1997). These trait trade-offs can exist at the genome level. For example, nutrient  
10 87 limitation can select for genomes with low GC content due to the adenine-thymine (AT) base  
11 88 pair having a lower nitrogen-to-carbon ratio (N: C= 7: 10, =0.7) than the GC pair (N: C= 8: 9,  
12 89 ~0.9). Therefore, organisms with large genomes may adapt to higher N requirements by  
13 90 reducing the use of the relatively N-rich GC base pair (Kelly 2018). In addition, tissue  
14 91 nutrient concentration reflects an organism's nutrient demand and utilization (Aerts & Chapin  
15 92 III 1999) and determines organism's growth efficiency (Sinsabaugh *et al.* 2013). A link  
16 93 between genomic traits and tissue nutrient concentrations can be expected if species with  
17 94 different genome sizes have different nutrient requirements, exploit nutrients with varying  
18 95 efficiency, or store or convert nutrients to biomass at different rates (Faizullah *et al.* 2021).  
19 96 Besides the potential linkage between GC content and N economy, others have found a direct  
20 97 connection of GC content with recombination of homologous chromosomes in organisms,  
21 98 along with other factors such as mutation rates, natural selection, and genetic drift can  
22 99 influence GC content of genomes. But how nutrient availability is involved in the GC content  
23 100 of soil fungi is less investigated.

24 101 A striking property of fungi is their diversity of lifestyles reflected by different guilds  
25 102 (e.g. mycorrhizal, pathogenic, saprotrophic guilds). Each guild exhibits specific trophic  
26 103 behaviours, i.e., resource exploitation and habitat or host preferences (Johnson *et al.* 2013;  
27 104 Naranjo-Ortiz & Gabaldón 2019), which might affect their sensitivity to changes in resource  
28 105 availability (Maaroufi *et al.* 2019; Lekberg *et al.* 2021). The fitness of free-living (e.g.,  
29 106 saprotrophic) fungi is expected to be directly linked to abiotic resource conditions, while the  
30 107 fitness of symbiotic (e.g., mycorrhizal and pathogenic) fungi is linked to nutrients and carbon  
31 108 source-sink dynamics during interactions with the host (and, for mycorrhizal guilds,  
32 109 competition of soil resources with saprotrophs). In addition, recent studies suggested that  
33 110 genome size varies among fungi with different lifestyles (Spanu 2012; Miyauchi *et al.* 2020).  
34 111 Spanu (2012) suggested that the acquisition of biotrophy in plant pathogens was associated  
35 112 with an expanded genome. Genome size variation during plant-host interactions or specific  
36 113 sources in nutrient acquisition among different fungal guilds might determine the strength of  
37 114 the correlation between genome size and nutrient demand for growth. Thus, considering

1  
2  
3 115 fungal guilds help to understand how trophic behaviour shapes an organism's genome traits  
4  
5 116 and nutrient adaptation (Bahram & Netherway 2021).  
6

7 117 Here we compiled species-level data on two genomic traits (genome size from 2,437  
8  
9 118 and GC content from 1,276 fungal species) and estimated their inter-correlations and  
10  
11 119 relationships with fungal nutrient concentrations to assess their ecological trade-offs and  
12  
13 120 adaptative strategies among and within fungal guilds. We also investigated the distribution of  
14  
15 121 these genomic traits in relation to nutrient availability in 463 soil samples taken from  
16  
17 122 grassland, woodland and forest ecosystems over natural soil nutrient gradients across  
18  
19 123 Australia. We hypothesized that nutrient-depleted environments favour fungi with smaller  
20  
21 124 genomes and lower GC content. We also hypothesised that such environmental selection  
22  
23 125 would be stronger on genomic traits for plant symbiotic fungi than saprotrophs given that  
24  
25 126 plant symbiotic fungi will respond to changes in resource availability as well as to how their  
26  
27 127 plant hosts respond to those changes (eg., greater resource-sharing or stronger competition  
28  
29 128 for limited resources).  
30

129

## 30 130 **Material and Methods**

### 31 131 *Fungal genomic and nutritional traits*

32  
33 132 We retrieved fungal genome size (Mbp, 1C-values) and GC content (%) data from the  
34  
35 133 National Center for Biotechnology Information (NCBI) genome database  
36  
37 134 (<https://www.ncbi.nlm.nih.gov/genome/>), Mycocosm (Grigoriev *et al.* 2014)  
38  
39 135 (<https://mycocosm.jgi.doe.gov/mycocosm/home>) and Fungal Genome Size Database  
40  
41 136 (<http://www.zbi.ee/fungal-genomesize>) (Kullman *et al.* 2005). We assigned the most recently  
42  
43 137 accepted taxonomic names from Catalogue of Life (<https://www.catalogueoflife.org/>) to our  
44  
45 138 database by comparing accepted and synonym names using the *cp\_nu\_suggest* function from  
46  
47 139 the *rcol* R package (Chamberlain 2021). After checking synonymic names, the genomic traits  
48  
49 140 were averaged for the same species from the three resources. We kept the genomic trait data  
50  
51 141 for species that could be aligned with fungal guild information derived from the FUNGuild  
52  
53 142 database with a confidence level of “probable” or “highly probable” (Nguyen *et al.* 2016).  
54  
55 143 This resulted in our dataset containing 2,437 and 1,276 guild-annotated species with genome  
56  
57 144 size and GC content data, respectively; we obtained genome size data for all 1,276 species  
58  
59 145 with GC content data. For both genomic traits, the guilds included were 7 Arbuscular  
60  
146 mycorrhizal (AM), 78 ectomycorrhizal (EcM), 317 plant pathogenic and 874 saprotrophic  
147 species.



1  
2  
3 148 We retrieved fungal trait data related to growth and nutrition from the Fungal  
4  
5 149 Functional Trait database (Fun<sup>Fun</sup>) (Zanne *et al.* 2020), including N and P concentrations in  
6  
7 150 fungal mycelial tissues obtained from the database but originally derived from Zhang and  
8  
9 151 Elser (2017). There were 534 observations for 280 species in mycelium N and 654  
10  
11 152 observations for 275 species in mycelium P. For species with multiple observations, we  
12  
13 153 calculated a mean trait value. After cross-referencing each database, we found matching  
14  
15 154 information for genome size with 70 and 85 species for mycelium N and P, respectively. For  
16  
17 155 matches with GC content, we found 32 and 47 species with estimates of mycelium N and P,  
18  
19 156 respectively. The numbers of observations at the guild-level for the overlap between genomic  
20  
21 157 traits and fungal nutrients were supplied in Table 1. The majority of the observations with  
22  
23 158 overlapping data were from ECM and saprotrophic fungi, followed by plant pathogenic  
24  
25 159 fungi.

#### 26 160 *Study sites and soil sampling across Australia*

27  
28 161 In this study, we used data generated from soil samples collected across Australia  
29  
30 162 during the period from 2005 to 2018 as part of the Australian Microbiome Initiative (AMI;  
31  
32 163 formerly the Biomes of Soil Environments [BASE] project; Bissett *et al.* 2016). We obtained  
33  
34 164 the sequencing data from <https://www.australianmicrobiome.com/> on 27 August 2020. A  
35  
36 165 total of 463 soil samples from 409 sites were included in this study, selected based on the  
37  
38 166 criteria described below (see Appendix S1 for a list of sample numbers used here). The soil  
39  
40 167 samples were collected according to the methods described in Bissett *et al.* (2016). Briefly,  
41  
42 168 composite soil samples were generated from 9-25 soil cores collected to a depth of 10 cm  
43  
44 169 within 25 × 25 m plots (details see Fig. S1). Each soil sample was separated into two  
45  
46 170 subsamples. The first subsample was frozen and transported to the Adelaide node of the  
47  
48 171 Australian Genome Research Facility (AGRF) laboratories for DNA extraction and  
49  
50 172 sequencing. The second subsample was air-dried for soil available N and P measurements  
51  
52 173 (Bissett *et al.* 2016). Note that we used soil available nutrients (inorganic form) as the proxy  
53  
54 174 for indicating fungal nutrient limitation although some fungi, particularly saprotrophs but  
55  
56 175 also EcM fungi, can also access organically bound soil nutrients in addition to those in  
57  
58 176 inorganic forms. In addition, biotrophic pathogens can obtain essential nutrients from host  
59  
60 177 tissues, which will vary to a certain extent but not entirely with soil nutrient availability.  
178 Thus, estimates of relationships may lack precision for these guilds, and further studies could

179 also benefit from estimating soil total N and P as well as autecological studies of how free-  
180 living and symbiotic fungi acquire and allocate N and P from different source pools.

### 181 *Fungal DNA extraction, sequencing and bioinformatics*

182 Genomic DNA extraction and bioinformatic analysis were conducted according to  
183 methods described in Bissett *et al.* (2016). Briefly, soil gDNA was extracted in triplicate  
184 using MoBio PowerSoil extraction kits (MO BIO Laboratories Inc., USA) following the  
185 manufacturer's instructions. DNA amplicons targeting the fungal ITS1-5.8S-ITS2 region  
186 were prepared and sequenced at the Australian Genome Research Facility (Melbourne,  
187 Australia) and Ramaciotti Centre for Genomics (Sydney, Australia). ITS amplicons were  
188 sequenced on the Illumina MiSEQ platform with MiSeq Reagent Kit v3 600 cycle chemistry,  
189 to produce 300 bp paired-end reads. The sequenced region (ITS1-5.8S-ITS2) is 550bp on  
190 average in fungi but can be much larger (Nilsson *et al.* 2015), thus for many reads there was  
191 not sufficient (or any) overlap between the forward and reverse pair of each template  
192 sequence to merge them into a single sequence read. To ensure that those long fungal ITS1-  
193 5.8S-ITS2 sequences were not excluded from our analysis, the ITS1 and ITS2 regions were  
194 separately extracted from forward and reverse reads, respectively, using ITSx  
195 (Bengtsson-Palme *et al.* 2013) and both regions were processed. Zero-radius operational  
196 taxonomic units (zOTUs) were generated from identified ITS regions and frequencies of  
197 zOTUs within each sample were determined. Taxonomic and guild identities of zOTUs were  
198 assigned and matched with fungal genomic traits (Appendix S2). Eventually, there were 225  
199 fungal species with genome size data occurring in 463 soil samples and 145 species with GC  
200 content data in 460 soil samples across Australia (Fig. S1-S2).

201

### 202 *Statistical analyses*

203 To assess correlations among traits, we used the phylogenetic generalized least  
204 squares (PGLS) method to account for the shared evolutionary histories among fungal  
205 species. We constructed the fungal phylogenetic tree using the fungal mega-phylogeny  
206 published by Li *et al.* (2021) (time-calibrated tree) as the backbone. For those genera and  
207 species that were absent from the meta-phylogeny, we used V.PhyloMaker  
208 (<https://github.com/jinyizju/>) to add them to their respective families (in the case of genera)  
209 and genera (in the case of species) in the mega-phylogeny under Scenario 3 (Jin & Qian  
210 2019). The polytomies were resolved by *multi2di* function in the R package ape (Paradis &  
211 Schliep 2019). We estimated differences in genome size and GC content among fungal  
212 guilds, as a fixed effect, by fitting PGLS-based models using the *phylolm* R package (Tung

1  
2  
3 213 Ho & Ané 2014). In addition, we examined the relationships between (i) fungal genome size  
4 214 and GC content; and (ii) each of genome size and GC content versus each of mycelial N  
5 215 concentration and mycelial P concentration. To determine differences in resource acquisition  
6 216 strategies among guilds, we tested these relationships separately across four guilds: AM,  
7 217 EcM, pathogens and saprotrophic fungi. These PGLS-regressions were performed using the  
8 218 *phylolm* function. For the correlations in (ii), we also calculated the independent variation  
9 219 explained by each predictor ( $R^2$ ) using the R package *rr2* (Ives & Li 2018).

15 220 We evaluated distributions of fungal genomic traits in soils and their correlations with  
16 221 soil available nutrients. We calculated the community-weighted mean genome size (CWM-  
17 222 GS) and GC content (CWM-GC) using the fungal species-sample table generated after  
18 223 summing read counts associated with zOTUs assigned to the same species. To generate  
19 224 CWM, weighted averages of genomic traits were calculated for each soil sample using the  
20 225 relative abundance of each fungal species as weights. CWM was calculated across the entire  
21 226 fungal community and individually for subsets of species assigned to each of three major  
22 227 fungal guilds (AM fungi were represented by too few species for a robust analysis to be  
23 228 performed, thus were not included) and correlated with soil available nutrients. Note that  
24 229 when calculating CWM, values can be potentially affected by coverage in each sample, i.e.,  
25 230 the proportion of zOTU counts that could be assigned to a species for which trait data were  
26 231 available and, therefore, were used to calculate CWM. To assess whether samples with very  
27 232 low species coverage may cause spurious correlations, we analysed the data twice using the  
28 233 following strategies: (i) using all soil samples for which trait data could be assigned to at least  
29 234 one zOTU, and (ii) excluding those soil samples for which the percentage of zOTU read  
30 235 counts that could be assigned trait data was lower than the median proportion across all  
31 236 samples.

36 237 Linear mixed-effects models were used to evaluate the effects of soil available  
37 238 nutrients within each vegetation category on (i) CWM-GS and CWM-GC at the community  
38 239 level, (ii) relative abundance of each fungal guild, and (iii) CWM-GS and CWM-GC at the  
39 240 fungal guild level (e.g., CWM was calculated for each of the four fungal guilds, separately).  
40 241 For all mixed-models, we included one of each soil nutrient ( $\text{NH}_4^+$ ,  $\text{NO}_3^-$ ,  $\text{PO}_4^-$ ) as fixed  
41 242 effects and the ITS fragment used to generate zOTUs (ITS1 and ITS2) as random effects, to  
42 243 account for potential biases between the two regions in their ability to detect fungal species  
43 244 and, thus, result in different members of the fungal community being detected (Bazzicalupo  
44 245 *et al.* 2013). Before the modelling using functions in the *lme4* package (Bates *et al.* 2015),  
45 246 soil nutrients and CWM-GS values were log-transformed (natural logarithms). Dispersion

1  
2  
3 247 was checked using the *DHARMA* package (Hartig 2022). While determining statistical  
4  
5 248 significance, Kenward-Roger degrees of freedom were calculated using the *Anova* function  
6  
7 249 from the *car* package (Fox *et al.* 2012). Marginal  $R^2$  values (variance explained only by fixed  
8  
9 250 effects) were calculated using the *r.squaredGLMM* function from the *MuMIn* package  
10  
11 251 (Barton 2015). All the analyses were performed using R version 4.0 (R Core Team 2020).  
12

252

## 13 253 **Results**

### 14 254 *Genomic, growth and nutritional traits correlations across fungal guilds*

15 255 Fungal genome size was negatively correlated with GC content across all four guilds  
16  
17 256 (Fig. 1), with AM and ECM fungi having significantly larger genomes and lower GC  
18  
19 257 contents on average than pathogenic and saprotrophic fungi (Table S1). Furthermore, this  
20  
21 258 pattern was still observed even when we restricted the analysis to only the subset of species  
22  
23 259 observed within the Australian Microbiome database (Fig. S3). When using all species for  
24  
25 260 which we collected trait data, we observed relatively strong negative correlations for EcM ( $r^2$   
26  
27 261 = 0.23,  $P < 0.001$ ; slope 95% confidence interval (CI) was [-0.061, -0.025]; Fig. 1b) and  
28  
29 262 plant pathogens ( $r^2 = 0.32$ ,  $P < 0.001$ ; 95% CI [-0.021, -0.015]; Fig. 1c) and a much weaker  
30  
31 263 negative correlation for saprotrophs ( $r^2 = 0.03$ ,  $P < 0.001$ ; 95% CI [-0.007, -0.003]; Fig. 1d).

32 264 Fungal genome size negatively correlated with tissue N ( $r^2 = 0.17$ ,  $P = 0.002$ ; slope  
33  
34 265 95% CI was [-4.143, -1.175]; Fig. 2a) and P concentrations ( $r^2 = 0.14$ ,  $P = 0.01$ ; slope 95%  
35  
36 266 CI [-0.244, -0.043]; Fig. 2b). Furthermore, fungal GC content was negatively correlated with  
37  
38 267 N ( $r^2 = 0.14$ ,  $P = 0.009$ ; slope 95% CI [-1.278, -0.200]; Fig. 2c) and P concentrations ( $r^2 =$   
39  
40 268 0.10,  $P = 0.028$ ; slope 95% CI [-0.054, -0.004]; Fig. 2d) in fungal tissues. When comparing  
41  
42 269 the relative importance of the genomic traits to explain variation in growth and nutritional  
43  
44 270 traits, genome size explained greater variation than GC content in predicting concentrations  
45  
46 271 of N and P (Table S2).

47 272 However, the correlational patterns between genomic traits and nutritional traits were  
48  
49 273 also dependent on specific fungal guilds: the negative correlation between genome size and  
50  
51 274 fungal N was only observed for EcM and saprotrophic fungi; while the negative correlation  
52  
53 275 of genome size and P was only observed among saprotrophs (Fig. 2ab). In contrast, the  
54  
55 276 negative correlations between GC content and fungal nutrients were not significant within  
56  
57 277 most guilds (Fig. 2cd), except for that with fungal P for saprotrophic fungi.  
58

278

### 58 279 *Biogeography of fungal traits at the community level*

59

60

1  
2  
3 280 On average, 3% of sequence reads per sample could be assigned trait data (range:  
4 281 0.002% to 72%). This is equivalent to a range from 18 to 464220 sequence reads (mean:  
5 282 16970) being assigned trait data in each sample (Appendix S3). Contrary to our initial  
6 283 hypothesis, we found that nutrient-depleted soils favoured fungi with larger genomes. This  
7 284 pattern was explained, in part by variation in soil N and P availability, with the response  
8 285 being most consistent and with the largest effect sizes in forest samples. In general, CWM-  
9 286 GS was negatively correlated with soil available  $\text{NH}_4^+$  and  $\text{NO}_3^-$ , with correlations being  
10 287 strongest for forests ( $\text{NH}_4^+$ :  $r^2 = 0.22$ ,  $P < 0.001$ ;  $\text{NO}_3^-$ :  $r^2 = 0.22$ ,  $P < 0.001$ ), weak but  
11 288 significant for woodlands ( $\text{NH}_4^+$ :  $r^2 = 0.02$ ,  $P < 0.001$ ;  $\text{NO}_3^-$ :  $r^2 = 0.07$ ,  $P < 0.001$ ), and not  
12 289 observed for grasslands (Fig. 3abc). Similarly, CWM-GS was negatively correlated with soil  
13 290 available  $\text{PO}_4^-$ , although this correlation was only significant and relatively weak for forests  
14 291 ( $r^2 = 0.05$ ,  $P < 0.001$ ; Fig. 3d). In contrast, CWM-GC was positively correlated with soil  
15 292 available nutrients as we expected in our initial hypothesis, with most of these correlations  
16 293 being significant (Fig. 3e-g), except for correlations with soil available  $\text{NH}_4^+$  and  $\text{NO}_3^-$  in  
17 294 grasslands (Fig. 3ef; Table S3). We also observed that these relationships between trait  
18 295 CWMs and soil available nutrients were generally consistent when we limited our analyses to  
19 296 samples in which trait data could be assigned to a *relatively* high proportion of community  
20 297 members (Table S3, which compares the analysis using all samples to one using only those  
21 298 for which a proportion of reads that could be assigned trait data was higher than the median  
22 299 value for all samples (1.2%)).

300

### 301 *Guild-specific biogeographical patterns of fungal traits*

302 The response patterns of genomic traits along the soil nutrient gradient at the  
303 community level can be explained by both shifts in the relative abundance among different  
304 fungal guilds and by species turnover within specific fungal guilds. For guild composition,  
305 EcM fungi exhibited generally negative responses (Fig. 4a), and plant-pathogenic (Fig. 4d)  
306 and saprotrophic fungi (Fig. 4g; Fig. S4; Table S4) both exhibited generally positive  
307 responses.

308 For CWM-GS responses to increased soil nutrients, most effect sizes were neutral for  
309 EcM, except for two negative responses ( $\text{NO}_3^-$  in the woodland and  $\text{PO}_4^-$  in the forest; Fig.  
310 4b). In contrast, associations with CWM-GS were positive for plant pathogens (Fig. 4e),  
311 while negative for saprotrophic fungi (Fig. 4h; Fig. S5; Table S5). For CWM-GC responses  
312 to increased soil nutrients, most effect sizes were neutral for ECM fungi, except for two  
313 positive responses (response to  $\text{NO}_3^-$  in the woodland and  $\text{PO}_4^-$  in the forest; Fig. 4c). Most

1  
2  
3 314 associations with CWM-GC were positive for plant pathogens (Fig. 4f), while both positive  
4  
5 315 and negative responses existed for saprotrophic fungi (Fig. 4i; Fig S6; Table S6).  
6  
7 316

## 8 317 **Discussion**

9  
10 318 Our study demonstrates that genomic traits can be potentially used to explain  
11  
12 319 variations in fungal functional traits and biogeographical patterns of fungi along large  
13  
14 320 environmental gradients. First, fungi with large genomes produced nutrient-poor tissue and  
15  
16 321 adjusted their nucleotide composition by reducing the frequency of the expensive GC base-  
17  
18 322 pair. Second, soils across Australia with low levels of nutrient availability favoured fungi  
19  
20 323 with large genomes and low GC content. Such genomic patterns at the community level  
21  
22 324 along nutrient gradients were generally consistent across different forms/types of nutrients  
23  
24 325 and in different vegetations. Finally, partially supporting our hypothesis, the environmental  
25  
26 326 selection on genomic traits along the nutrient gradient would be stronger in symbiotic fungi  
27  
28 327 than saprotrophic fungi, but this was only true for pathogenic but not for mycorrhizal fungi.  
29  
30 328 Together, our results reveal fundamental mechanisms that underpin genome size and  
31  
32 329 nucleotide selection in soil fungal communities along soil nutrient gradients, although in  
33  
34 330 some cases the explanatory power was weak. This is not surprising given that tissue nutrient  
35  
36 331 concentration and soil mineral nutrient stocks only partially reflect the range of biological  
37  
38 332 and environmental controls on fungal distributions in soils (Tedersoo *et al.* 2014).  
39  
40 333

### 41 334 *Trade-offs among fungal genomic, growth and nutritional traits*

42 335 The negative correlation between genome size and GC content in fungal genomes,  
43  
44 336 especially for pathogenic fungi, supports the existence of a close relationship between the  
45  
46 337 size and composition of fungal genomes. Previous studies found that community-averaged  
47  
48 338 GC content and genome size in bacteria were negatively correlated in soil but, in the same  
49  
50 339 study, these two genomic traits were positively correlated under marine conditions (Chuckran  
51  
52 340 *et al.* 2021; Chuckran *et al.* 2022). Thus, their results support our findings and suggest similar  
53  
54 341 selection pressures on genome size and GC content for soil bacteria and soil fungi. In bacteria  
55  
56 342 and archaea, several environmental factors in addition to nutrient availability are known to  
57  
58 343 drive the selection on genome size and genomic GC-content, mainly including environmental  
59  
60 344 selection (such as growth temperature and the availability of oxygen; see Foerstner *et al.*  
61  
62 345 (2005); Sabath *et al.* (2013)) or GC-biased gene conversion (which favours G/C nucleotides  
63  
64 346 during DNA recombination; see Webster and Hurst (2012) and Lassalle *et al.* (2015)). In our  
65  
66 347 study, we mainly focused on nutrients due to their importance for genome construction

1  
2  
3 348 during the growth of fungi, but acknowledge that further variation in these genomic traits  
4  
5 349 could be predicted with additional environmental drivers. In fungi, we found fungal genome  
6  
7 350 sizes were negatively correlated with fungal N and P concentrations, suggesting that species  
8  
9 351 with small genomes produced nutrient-rich tissue. We propose that this fungal behaviour has  
10  
11 352 a similar function as in plants, for which the plant leaf economic spectrum suggests that  
12  
13 353 nutrient-rich tissue is associated with high nutrient demand during rapid growth, and that  
14  
15 354 small genomes facilitate a faster cell division and therefore a higher growth rate (Rayburn *et*  
16  
17 355 *al.* 1994; Knight *et al.* 2005).

17 356 For bacterial and archaeal communities, inconsistent patterns have been described  
18  
19 357 when linking genome size to copiotrophy/oligotrophy in previous studies, with both  
20  
21 358 supporting (Liu *et al.* 2023) and non-supporting (Westoby *et al.* 2021) evidence being  
22  
23 359 revealed. We previously found that bacterial genome sizes weakly but significantly correlated  
24  
25 360 with soil carbon and P concentrations, in that bacteria with small genomes exhibited fewer  
26  
27 361 negative responses to increasing soil carbon or more positive responses to increasing soil P,  
28  
29 362 suggesting that higher soil fertility favours bacteria with smaller genomes (Liu *et al.* 2023).  
30  
31 363 Fungi are generally considered to be more oligotrophic than bacteria, but it may be possible  
32  
33 364 to assess these relationships in targeted studies of fungi exhibiting generally more  
34  
35 365 copiotrophic strategies, such as those in the Zygomycota (Ho *et al.* 2017).

36 366

### 367 *Nutrient-depleted soil environments favour fungi with large genomes and low GC content*

37 368 CWM-GS significantly increased, and CWM-GC content decreased, in response to  
38  
39 369 decreasing soil available nutrients, indicating that fungi with large genomes and low GC  
40  
41 370 content dominated in nutrient-depleted sites. These patterns were generally found in each  
42  
43 371 vegetation type (grassland, woodland and forest) or soil nutrient type (NH<sub>4</sub><sup>+</sup>, NO<sub>3</sub><sup>-</sup> and PO<sub>4</sub><sup>-</sup>),  
44  
45 372 but were more pronounced in forest ecosystems, possibly due to the high prevalence of EcM  
46  
47 373 fungi in forests and presumably a tighter relationship with available N (Table S4). Additional  
48  
49 374 analyses showed weak but significant correlations between the relative abundance of EcM  
50  
51 375 fungi and their genomic traits at the community level, where CWM-GS was increased ( $r^2 =$   
52  
53 376  $0.02$ ,  $P < 0.0001$ ) and CWM-GC content decreased ( $r^2 = 0.03$ ,  $P < 0.0001$ ) when the relative  
54  
55 377 abundance of EcM fungi increased.

55 378 The opposite responses between CWM-GS and CWM-GC content along these  
56  
57 379 nutrient gradients may be explained by the trade-off between the two genomic traits: fungi  
58  
59 380 with large genomes tended likely to reduce the cost of genome construction by lowering the  
60  
381 GC content. Genome size increases in low-nutrient environments can reflect several potential

1  
2  
3 382 mechanisms to cope with these conditions. For example, undergoing gene duplication can  
4  
5 383 facilitate production of the enzymes and proteins required to cope with nutrient stress  
6  
7 384 (Konstantinidis & Tiedje 2004; Giovannoni *et al.* 2005), suggesting a mechanism for EcM  
8  
9 385 and saprotrophic fungi that produce extracellular hydrolytic enzymes for nutrient  
10  
11 386 mineralization. Such mechanism aligns with an active and fitness-enhancing response of an  
12  
13 387 organism capable of enhanced resource acquisition. As at a certain point it may no longer be  
14  
15 388 adaptive to have a large genome under nutrient-limiting conditions (given that it is  
16  
17 389 accompanied by a high genome construction cost), reducing GC content to build large  
18  
19 390 genomes can offset the cost because a reduction in GC content decreases the amount of  
20  
21 391 nitrogen required for DNA synthesis. These findings suggest that shifts in size and  
22  
23 392 composition of organisms' genomes can be an important evolutionary and ecological strategy  
24  
25 393 to adapt to their local nutritional environments.

26 394 To our knowledge, our study provided the first link between fungal genomic traits and  
27  
28 395 fungal distribution patterns across large nutrient gradients. Previous investigations on this  
29  
30 396 topic have focused on genome size patterns in plants (Pellicer *et al.* 2018) and bacteria  
31  
32 397 (Chuckran *et al.* 2021). For plants, positive associations between CWM-GS and soil nutrients  
33  
34 398 were reported (Šmarda *et al.* 2013; Guignard *et al.* 2016). Plants with large genomes are  
35  
36 399 more likely to dominate in communities where nutrient availability is high in soil, probably  
37  
38 400 because plants with large genomes require more nutrients to build genomes (Pellicer *et al.*  
39  
40 401 2018). For bacteria, inconsistent patterns have been described previously as both positive and  
41  
42 402 negative correlations between genome size and nutrients were revealed. For example,  
43  
44 403 bacteria with smaller genomes were associated with more harsh or nutrient-limited  
45  
46 404 environments (Chuckran *et al.* 2022). In contrast, whole-genome shotgun sequencing data of  
47  
48 405 *Lactobacillus* (phylum Firmicutes) suggested that deep, nutrient-depleted marine  
49  
50 406 environments were dominated by bacteria with large genomes (Makarova *et al.* 2006).

51 407

#### 52 408 *Guild-level responses along soil nutrient gradients*

53 409 Correlations between genomic traits and soil nutrient availabilities at the whole fungal  
54  
55 410 community level could be derived from two major processes: (i) shifts in fungal guild  
56  
57 411 structure and (ii) species turnover within fungal guilds. For the first process, the relative  
58  
59 412 abundance for fungi with relatively large genomes (i.e., EcM; Fig. 1) significantly decreased  
60  
413 while the relative abundance of fungi with relatively small genomes (i.e., plant pathogenic  
414 and saprotrophic fungi) significantly increased when soil nutrients increased to high levels;  
415 such shifts in fungal guild structure resulted in a lower CWM-GS under nutrient-enriched



1  
2  
3 416 sites. For the second process, the strength of such negative correlations varied depending on  
4  
5 417 vegetation type, nutrient form and fungal guild. For example, when observed to be  
6  
7 418 statistically significant, genomic traits of saprotrophs tended to exhibit relatively weak  
8  
9 419 correlations across many combinations of vegetation types and nutrient forms compared with  
10  
11 420 EcM and pathogenic fungi. The latter two groups also exhibited noticeable variations, with  
12  
13 421 EcM fungi showing significant correlations ranging from weak to strong in woodlands and  
14  
15 422 forests, and pathogenic fungi generally showing stronger correlations in grasslands. The  
16  
17 423 distinct responses among guilds highlight the importance of considering different fungal  
18  
19 424 groups independently when investigating their genomic trait patterns along environmental  
20  
21 425 gradients.

22 426 Pathogens stood out because their CWM-GS and CWM-GC content was the only  
23  
24 427 guild that mainly showed positive responses to soil fertility, i.e., supporting our original  
25  
26 428 hypothesis that nutrient-depleted environments favour fungal species with smaller genomes  
27  
28 429 and lower GC content. The nitrogen disease hypothesis states that plant growth at high N  
29  
30 430 availability may result in increased plant susceptibility to pathogens because of increased  
31  
32 431 foliar nitrogen concentrations (Mitchell *et al.* 2003). Our results expand this theory by  
33  
34 432 showing that pathogens with large genomes and higher GC content could be particularly  
35  
36 433 favoured in nutrient-enriched environments. Spanu and Kämper (2010) proposed that  
37  
38 434 acquisition of biotrophy in plant fungal pathogens is associated with an expanded genome,  
39  
40 435 with obligate pathogenic fungi such as powdery mildews and rust fungi having greater  
41  
42 436 genome size than necrotrophic pathogens. However, in our study, the proportion of DNA  
43  
44 437 reads that could be assigned to biotrophic fungi, including rust and mildew fungi, was low  
45  
46 438 (0.2% on average) among all pathogens. Thus, the changes in genomic size along soil  
47  
48 439 nutrients should be also driven by other pathogen types (e.g. necrotrophic pathogens).  
49  
50 440 Variation in genome size can also be related to polyploidy, the amount of transposable  
51  
52 441 elements and the potential function of predicted genes encoding secreted proteins and other  
53  
54 442 effectors in pathogens (Lo Presti *et al.* 2015; Lorrain *et al.* 2019). We found that P  
55  
56 443 availability had stronger explanatory power than N availability in explaining genome GC  
57  
58 444 content for pathogens in grassland and woodland soils. This suggests that other, non-C/N-  
59  
60 445 related, mechanisms may have been involved. In addition, P availability does influence plant  
61  
62 446 and fungal growth and ecophysiologicals, and acquisition of limited P can indirectly affect  
63  
64 447 organisms' growing environments, modifying pH, redox state or availability of other  
65  
66 448 nutrients availability, all of which could lead to environmental selection on GC content

1  
2  
3 449 (Foerstner et al. 2005). Determining which of these drivers is responsible for shaping genome  
4  
5 450 GC content in pathogenic (and other) fungi warrants future research.  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 451 **Acknowledgements**  
4

5 452 We would like to acknowledge the contribution of the Australian Microbiome and Biomes of  
6  
7 453 Australian Soil Environments (BASE) consortiums in the generation of data used in this  
8  
9 454 publication. They are supported by funding from the CSIRO, Director of National Parks,  
10  
11 455 through the Bush Blitz program funded by the Australian Government and BHP,  
12  
13 456 Bioplatforms Australia and the Integrated Marine Observing System (IMOS) through the  
14  
15 457 Australian Government's National Collaborative Research Infrastructure Strategy (NCRIS).  
16  
17 458 We are grateful to all of those who contributed samples and contextual data to BASE and  
18  
19 459 AM (<https://www.australianmicrobiome.com/>). This work was supported by funding from the  
20  
21 460 Australian Research Council (FT190100590).

22 461 **Competing interests:**

23 462 The authors declare no competing interests.

24 463 **Data Doi:** <https://doi.org/10.6084/m9.figshare.20277759.v4>.  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

464 **Reference**

- 465 Aerts, R. & Chapin III, F.S. (1999). The mineral nutrition of wild plants revisited: a re-  
466 evaluation of processes and patterns. In: *Adv. Ecol. Res.* Elsevier, pp. 1-67.
- 467 Aguilar-Trigueros, C.A., Hempel, S., Powell, J.R., Cornwell, W.K. & Rillig, M.C. (2019).  
468 Bridging reproductive and microbial ecology: a case study in arbuscular mycorrhizal fungi.  
469 *ISME J.*, 13, 873-884.
- 470 Aguilar-Trigueros, C.A., Powell, J.R., Anderson, I.C., Antonovics, J. & Rillig, M.C. (2014).  
471 Ecological understanding of root-infecting fungi using trait-based approaches. *Trends Plant*  
472 *Sci.*, 19, 432-438.
- 473 Bahram, M. & Netherway, T. (2021). Fungi as mediators linking organisms and ecosystems.  
474 *FEMS Microbiol. Rev.*, 46.
- 475 Barberán, A., Ramirez, K.S., Leff, J.W., Bradford, M.A., Wall, D.H. & Fierer, N. (2014).  
476 Why are some microbes more ubiquitous than others? Predicting the habitat breadth of soil  
477 bacteria. *Ecol. Lett.*, 17, 794-802.
- 478 Barton, K. (2015). MuMIn: Model selection and model averaging based on information  
479 criteria (AICc and alike). <https://cran.r-project.org/web/packages/MuMIn/index.html>
- 480 Bates, D., Mächler, M., Bolker, B.M. & Walker, S.C. (2015). Fitting linear mixed-effects  
481 models using lme4. *J. Stat. Softw.*, 67, 1-48.
- 482 Bazzaz, F.A. & Grace, J. (1997). *Plant resource allocation*. Elsevier.
- 483 Bazzicalupo, A.L., Bálint, M. & Schmitt, I. (2013). Comparison of ITS1 and ITS2 rDNA in  
484 454 sequencing of hyperdiverse fungal communities. *Fungal Ecol*, 6, 102-109.
- 485 Bengtsson-Palme, J., Ryberg, M., Hartmann, M., Branco, S., Wang, Z., Godhe, A. *et al.*  
486 (2013). Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS  
487 sequences of fungi and other eukaryotes for analysis of environmental sequencing data.  
488 *Methods Ecol. Evol.*, 4, 914-919.
- 489 Bissett, A., Fitzgerald, A., Meintjes, T., Mele, P.M., Reith, F., Dennis, P.G. *et al.* (2016).  
490 Introducing BASE: the Biomes of Australian Soil Environments soil microbial diversity  
491 database. *Gigascience*, 5, s13742-13016-10126-13745.
- 492 Camenzind, T., Lehmann, A., Ahland, J., Rumpel, S. & Rillig, M.C. (2020). Trait-based  
493 approaches reveal fungal adaptations to nutrient-limiting conditions. *Environ. Microbiol.*, 22,  
494 3548-3560.
- 495 Camenzind, T., Philipp Grenz, K., Lehmann, J. & Rillig, M.C. (2021). Soil fungal mycelia  
496 have unexpectedly flexible stoichiometric C: N and C: P ratios. *Ecol. Lett.*, 24, 208-218.
- 497 Chamberlain, S. (2021). rcol: Catalogue of Life Client. <https://github.com/ropensci/rcol>.
- 498 Chan, J.Y., Bonser, S.P., Powell, J.R. & Cornwell, W.K. (2019). When to cut your losses:  
499 Dispersal allocation in an asexual filamentous fungus in response to competition. *Ecol Evol*,  
500 9, 4129-4137.
- 501 Chan, J.Y., Bonser, S.P., Powell, J.R. & Cornwell, W.K. (2020). Environmental cues for  
502 dispersal in a filamentous fungus in simulated islands. *Oikos*, 129, 1084-1092.
- 503 Chuckran, P.F., Flagg, C., Propster, J., Rutherford, W.A., Sieradzki, E., Blazewicz, S.J. *et al.*  
504 (2021). Edaphic controls on genome size and GC content of bacteria in soil microbial  
505 communities. *bioRxiv*, doi: <https://doi.org/10.1101/2021.1111.1117.469016>.
- 506 Chuckran, P.F., Hungate, B.A., Schwartz, E. & Dijkstra, P. (2022). Variation in genomic  
507 traits of microbial communities among ecosystems. *FEMS Microbes*, 2, 1-13.
- 508 Faizullah, L., Morton, J.A., Hersch-Green, E.I., Walczyk, A.M., Leitch, A.R. & Leitch, I.J.  
509 (2021). Exploring environmental selection on genome size in angiosperms. *Trends Plant Sci.*,  
510 26, 1039-1049.
- 511 Fierer, N. (2017). Embracing the unknown: disentangling the complexities of the soil  
512 microbiome. *Nat. Rev. Microbiol.*, 15, 579-590.

- 1  
2  
3 513 Foerstner, K.U., Von Mering, C., Hooper, S.D. & Bork, P. (2005). Environments shape the  
4 514 nucleotide composition of genomes. *EMBO Rep.*, 6, 1208-1213.
- 5 515 Fox, J., Weisberg, S., Adler, D., Bates, D., Baud-Bovy, G., Ellison, S. *et al.* (2012). Package  
6 516 'car'. *Vienna: R Foundation for Statistical Computing*.
- 7 517 Giovannoni, S.J., Cameron Thrash, J. & Temperton, B. (2014). Implications of streamlining  
8 518 theory for microbial ecology. *ISME J.*, 8, 1553-1565.
- 9 519 Giovannoni, S.J., Tripp, H.J., Givan, S., Podar, M., Vergin, K.L., Baptista, D. *et al.* (2005).  
10 520 Genome streamlining in a cosmopolitan oceanic bacterium. *Science*, 309, 1242-1245.
- 11 521 Grigoriev, I.V., Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otiillar, R. *et al.* (2014).  
12 522 MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Res.*, 42, D699-D704.
- 13 523 Gudelj, I., Weitz, J.S., Ferenci, T., Claire Horner-Devine, M., Marx, C.J., Meyer, J.R. *et al.*  
14 524 (2010). An integrative approach to understanding microbial diversity: from intracellular  
15 525 mechanisms to community structure. *Ecol. Lett.*, 13, 1073-1084.
- 16 526 Guignard, M.S., Nichols, R.A., Knell, R.J., Macdonald, A., Romila, C.A., Trimmer, M. *et al.*  
17 527 (2016). Genome size and ploidy influence angiosperm species' biomass under nitrogen and  
18 528 phosphorus limitation. *New Phytol.*, 210, 1195-1206.
- 19 529 Hartig, F. (2022). DHARMA: Residual diagnostics for hierarchical (multi-level/mixed)  
20 530 regression models. <https://CRAN.R-project.org/package=DHARMA>.
- 21 531 Ho, A., Di Lonardo, D.P. & Bodelier, P.L.E. (2017). Revisiting life strategy concepts in  
22 532 environmental microbial ecology. *FEMS Microbiol. Ecol.*, 93.
- 23 533 Ives, A.R. & Li, D. (2018). rr2: An R package to calculate R<sup>2</sup>s for regression models. *J. Open*  
24 534 *Source Softw.*, 3, 1028.
- 25 535 Jin, Y. & Qian, H. (2019). V. PhyloMaker: an R package that can generate very large  
26 536 phylogenies for vascular plants. *Ecography*, 42, 1353-1359.
- 27 537 Johnson, N.C., Angelard, C., Sanders, I.R. & Kiers, E.T. (2013). Predicting community and  
28 538 ecosystem outcomes of mycorrhizal responses to global change. *Ecol. Lett.*, 16, 140-153.
- 29 539 Kelly, S. (2018). The amount of nitrogen used for photosynthesis modulates molecular  
30 540 evolution in plants. *Mol. Biol. Evol.*, 35, 1616-1625.
- 31 541 Knight, C.A., Molinari, N.A. & Petrov, D.A. (2005). The large genome constraint hypothesis:  
32 542 evolution, ecology and phenotype. *Ann. Bot.*, 95, 177-190.
- 33 543 Konstantinidis, K.T. & Tiedje, J.M. (2004). Trends between gene content and genome size in  
34 544 prokaryotic species with larger genomes. *Proc. Natl. Acad. Sci. U.S.A.*, 101, 3160-3165.
- 35 545 Kraft, N.J., Valencia, R. & Ackerly, D.D. (2008). Functional traits and niche-based tree  
36 546 community assembly in an Amazonian forest. *Science*, 322, 580-582.
- 37 547 Kullman, B., Tamm, H. & Kullman, K. (2005). Fungal genome size database.
- 38 548 Lassalle, F., Périan, S., Bataillon, T., Nesme, X., Duret, L. & Daubin, V. (2015). GC-content  
39 549 evolution in bacterial genomes: the biased gene conversion hypothesis expands. *PLoS Genet.*,  
40 550 11, e1004941.
- 41 551 Leitch, I.J. & Leitch, A.R. (2013). Genome size diversity and evolution in land plants. In:  
42 552 *Plant Genome Diversity, Volume 2, Physical Structure, Behaviour and Evolution of Plant*  
43 553 *Genomes* (eds. Leitch, IJ, Greilhuber, J, Doležal, J & Wendel, JF). Springer, pp. 307-322.
- 44 554 Lekberg, Y., Arnillas, C.A., Borer, E.T., Bullington, L.S., Fierer, N., Kennedy, P.G. *et al.*  
45 555 (2021). Nitrogen and phosphorus fertilization consistently favor pathogenic over mutualistic  
46 556 fungi in grassland soils. *Nat. Commun.*, 12, 1-8.
- 47 557 Li, Y., Steenwyk, J.L., Chang, Y., Wang, Y., James, T.Y., Stajich, J.E. *et al.* (2021). A  
48 558 genome-scale phylogeny of the kingdom Fungi. *Curr. Biol.*, 31, 1653-1665.
- 49 559 Liu, H., Zhang, H., Powell, J., Delgado-Baquerizo, M., Wang, J. & Singh, B. (2023). Warmer  
50 560 and drier ecosystems select for smaller bacterial genomes in global soils. *iMeta*, Doi:  
51 561 10.1002/imt2.70.
- 52  
53  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 562 Lo Presti, L., Lanver, D., Schweizer, G., Tanaka, S., Liang, L., Tollot, M. *et al.* (2015).  
4 563 Fungal effectors and plant susceptibility. *Annu. Rev. Plant Biol.*, 66, 513-545.  
5 564 Lorrain, C., Gonçalves dos Santos, K.C., Germain, H., Hecker, A. & Duplessis, S. (2019).  
6 565 Advances in understanding obligate biotrophy in rust fungi. *New Phytol.*, 222, 1190-1206.  
7 566 Maaroufi, N.I., Nordin, A., Palmqvist, K., Hasselquist, N.J., Forsmark, B., Rosenstock, N.P.  
8 567 *et al.* (2019). Anthropogenic nitrogen enrichment enhances soil carbon accumulation by  
9 568 impacting saprotrophs rather than ectomycorrhizal fungal activity. *Glob Chang Biol*, 25,  
10 569 2900-2914.  
11 570 Makarova, K., Slesarev, A., Wolf, Y., Sorokin, A., Mirkin, B., Koonin, E. *et al.* (2006).  
12 571 Comparative genomics of the lactic acid bacteria. *Proc. Natl. Acad. Sci. U.S.A.*, 103, 15611-  
13 572 15616.  
14 573 Martiny, J.B., Jones, S.E., Lennon, J.T. & Martiny, A.C. (2015). Microbiomes in light of  
15 574 traits: a phylogenetic perspective. *Science*, 350, aac9323.  
16 575 Mitchell, C.E., Reich, P.B., Tilman, D. & Groth, J.V. (2003). Effects of elevated CO<sub>2</sub>,  
17 576 nitrogen deposition, and decreased species diversity on foliar fungal plant disease. *Glob*  
18 577 *Chang Biol*, 9, 438-451.  
19 578 Miyauchi, S., Kiss, E., Kuo, A., Drula, E., Kohler, A., Sánchez-García, M. *et al.* (2020).  
20 579 Large-scale genome sequencing of mycorrhizal fungi provides insights into the early  
21 580 evolution of symbiotic traits. *Nat. Commun.*, 11, 1-17.  
22 581 Naranjo-Ortiz, M.A. & Gabaldón, T. (2019). Fungal evolution: Diversity, taxonomy and  
23 582 phylogeny of the Fungi. *Biological Reviews*, 94, 2101-2137.  
24 583 Nguyen, N.H., Song, Z., Bates, S.T., Branco, S., Tedersoo, L., Menke, J. *et al.* (2016).  
25 584 FUNGuild: an open annotation tool for parsing fungal community datasets by ecological  
26 585 guild. *Fungal Ecol*, 20, 241-248.  
27 586 Nilsson, R.H., Tedersoo, L., Ryberg, M., Kristiansson, E., Hartmann, M., Unterseher, M. *et*  
28 587 *al.* (2015). A comprehensive, automatically updated fungal ITS sequence dataset for  
29 588 reference-based chimera control in environmental sequencing efforts. *Microbes Environ.*,  
30 589 ME14121.  
31 590 Paradis, E. & Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and  
32 591 evolutionary analyses in R. *Bioinformatics*, 35, 526-528.  
33 592 Pellicer, J., Hidalgo, O., Dodsworth, S. & Leitch, I.J. (2018). Genome size diversity and its  
34 593 impact on the evolution of land plants. *Genes*, 9, 88.  
35 594 Pölme, S., Abarenkov, K., Henrik Nilsson, R., Lindahl, B.D., Clemmensen, K.E., Kauserud,  
36 595 H. *et al.* (2020). FungalTraits: a user-friendly traits database of fungi and fungus-like  
37 596 stramenopiles. *Fungal Divers.*, 105, 1-16.  
38 597 R Core Team (2020). R: A language and environment for statistical computing. *R Foundation*  
39 598 *for Statistical Computing, Vienna, Austria.*  
40 599 Rayburn, A.L., Dudley, J. & Biradar, D. (1994). Selection for early flowering results in  
41 600 simultaneous selection for reduced nuclear DNA content in maize. *Plant Breed.*, 112, 318-  
42 601 322.  
43 602 Reich, P.B., Wright, I.J., Cavender-Bares, J., Craine, J., Oleksyn, J., Westoby, M. *et al.*  
44 603 (2003). The evolution of plant functional variation: traits, spectra, and strategies. *Int. J. Plant*  
45 604 *Sci.*, 164, S143-S164.  
46 605 Sabath, N., Ferrada, E., Barve, A. & Wagner, A. (2013). Growth temperature and genome  
47 606 size in bacteria are negatively correlated, suggesting genomic streamlining during thermal  
48 607 adaptation. *Genome Biol. Evol.*, 5, 966-977.  
49 608 Shenhav, L. & Zeevi, D. (2020). Resource conservation manifests in the genetic code.  
50 609 *Science*, 370, 683-687.  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

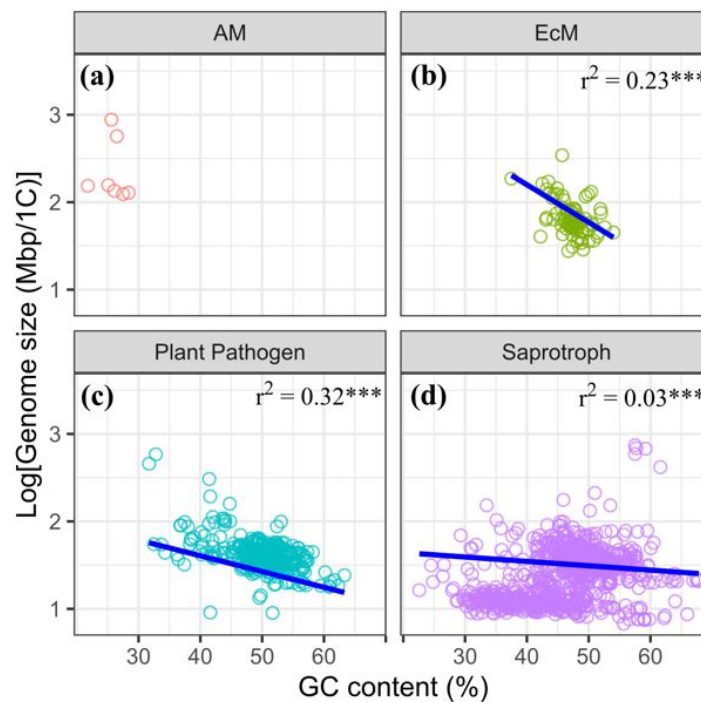
- 1  
2  
3 610 Sinsabaugh, R.L., Manzoni, S., Moorhead, D.L. & Richter, A. (2013). Carbon use efficiency  
4 611 of microbial communities: stoichiometry, methodology and modelling. *Ecol. Lett.*, 16, 930-  
5 612 939.
- 6 613 Šmarda, P., Hejman, M., Březinová, A., Horová, L., Steigerová, H., Zedek, F. *et al.* (2013).  
7 614 Effect of phosphorus availability on the selection of species with different ploidy levels and  
8 615 genome sizes in a long-term grassland fertilization experiment. *New Phytol.*, 200, 911-921.
- 9 616 Spanu, P. & Kämper, J. (2010). Genomics of biotrophy in fungi and oomycetes—emerging  
10 617 patterns. *Curr. Opin. Plant Biol.*, 13, 409-414.
- 11 618 Spanu, P.D. (2012). The genomics of obligate (and nonobligate) biotrophs. *Annu. Rev.*  
12 619 *Phytopathol.*, 50, 91-109.
- 13 620 Sterner, R.W. & Elser, J.J. (2002). *Ecological stoichiometry: the biology of elements from*  
14 621 *molecules to the biosphere*. Princeton university press, Princeton.
- 15 622 Talbot, J.M., Martin, F., Kohler, A., Henrissat, B. & Peay, K.G. (2015). Functional guild  
16 623 classification predicts the enzymatic role of fungi in litter and soil biogeochemistry. *Soil Biol.*  
17 624 *Biochem.*, 88, 441-456.
- 18 625 Tedersoo, L., Bahram, M., Pöhlme, S., Kõljalg, U., Yorou, N.S., Wijesundera, R. *et al.* (2014).  
19 626 Global diversity and geography of soil fungi. *Science*, 346, 1256688.
- 20 627 Tung Ho, L.s. & Ané, C. (2014). A linear-time algorithm for Gaussian and non-Gaussian trait  
21 628 evolution models. *Syst. Biol.*, 63, 397-408.
- 22 629 Webster, M.T. & Hurst, L.D. (2012). Direct and indirect consequences of meiotic  
23 630 recombination: implications for genome evolution. *Trends Genet.*, 28, 101-109.
- 24 631 Westoby, M., Falster, D.S., Moles, A.T., Vesk, P.A. & Wright, I.J. (2002). Plant ecological  
25 632 strategies: some leading dimensions of variation between species. *Annu. Rev. Ecol. Syst.*, 33,  
26 633 125-159.
- 27 634 Westoby, M., Nielsen, D.A., Gillings, M.R., Litchman, E., Madin, J.S., Paulsen, I.T. *et al.*  
28 635 (2021). Cell size, genome size, and maximum growth rate are near-independent dimensions  
29 636 of ecological variation across bacteria and archaea. *Ecol. Evol.*, 11, 3956-3976.
- 30 637 Widder, S., Allen, R.J., Pfeiffer, T., Curtis, T.P., Wiuf, C., Sloan, W.T. *et al.* (2016).  
31 638 Challenges in microbial ecology: building predictive understanding of community function  
32 639 and dynamics. *ISME J.*, 10, 2557-2568.
- 33 640 Zanne, A.E., Abarenkov, K., Afkhami, M.E., Aguilar-Trigueros, C.A., Bates, S., Bhatnagar,  
34 641 J.M. *et al.* (2020). Fungal functional ecology: bringing a trait-based approach to plant-  
35 642 associated fungi. *Biol. Rev.*, 95, 409-433.
- 36 643 Zhang, J. & Elser, J.J. (2017). Carbon: nitrogen: phosphorus stoichiometry in fungi: a meta-  
37 644 analysis. *Front. Microbiol.*, 8, 1281.
- 38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 646 **Table 1** Guild-level sample sizes (number of species) for the overlap between genome size  
4 647 (GS) or genome GC content and other fungal traits, including fungal N and P concentration.  
5 648 Specific fungal guilds are arbuscular mycorrhizal (AM), ectomycorrhizal (EcM), plant  
6 649 pathogenic and saprotrophic fungi.  
7  
8 650

<b>Model</b>	<b>AM</b>	<b>EcM</b>	<b>Pathogen</b>	<b>Saprotroph</b>
GS ~ fungal N	0	34	11	25
GS ~ fungal P	1	39	12	33
GC ~ fungal N	0	10	9	13
GC ~ fungal P	0	14	10	23
Total numbers	1	97	42	94

9  
10  
11  
12  
13  
14  
15  
16 651  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



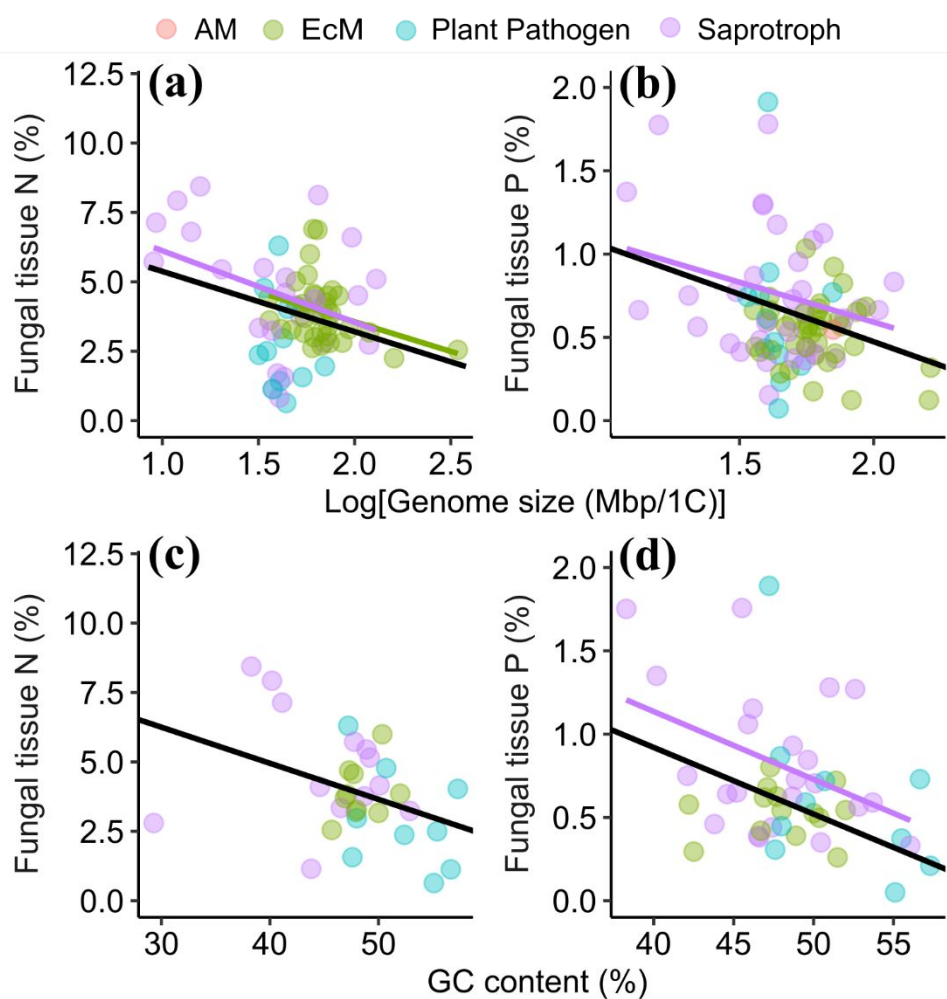
652 **Fig. 1**

653

654

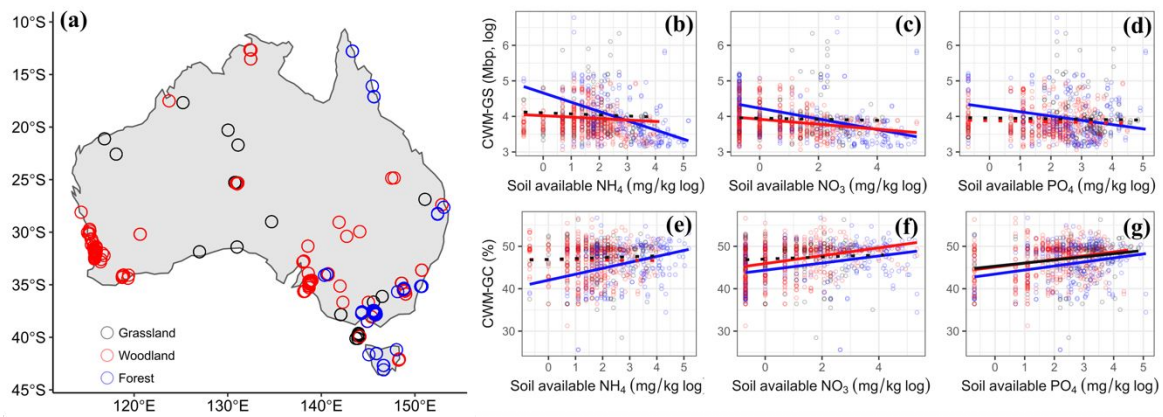
655 **Fig. 1** Correlations between genome size (log-transformed) and GC content (%) for species  
 656 within different fungal guilds. All trait estimates were averaged at the species level. Blue  
 657 lines indicate correlations being significant (\*\*\*)  $P < 0.001$  and were fitted considering  
 658 phylogenetic correlations among fungal species. Points with different colours indicate  
 659 different fungal guilds: arbuscular mycorrhizal (AM, red,  $n = 7$ ), ectomycorrhizal (EcM,  
 660 green,  $n = 78$ ), plant pathogen (blue,  $n = 317$ ), and saprotrophic (purple,  $n = 874$ ) fungi. The  
 661 data in panel (d) represent both yeasts and filamentous fungi, which differ in both genome  
 662 size and GC content, on average (Fig. S7).

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

663 **Fig. 2**

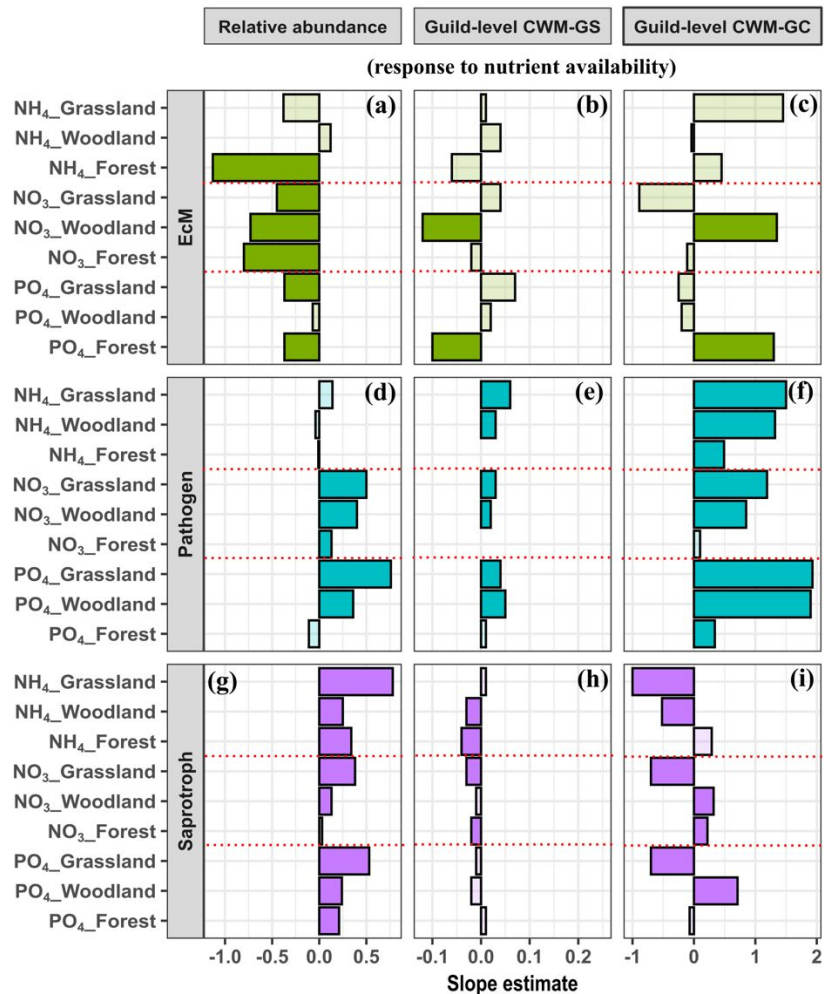
664

665 **Fig. 2** Correlations of genome size (log-transformed) or genome GC content (%) with fungal  
 666 nutrient concentrations, including N (%), (a, c) or P (%), (b, d) concentration in fungal tissues.  
 667 All trait estimates were averaged at the species level. Lines were added when correlations  
 668 were significant across all species (black lines) and within specific fungal guilds (coloured  
 669 lines). Points with different colours indicate different fungal guilds: arbuscular mycorrhizal  
 670 (AM, red), ectomycorrhizal (EcM, green), plant pathogen (blue), and saprotrophic (purple)  
 671 fungi. Points are darker with more overlapping. Models were fitted considering the  
 672 phylogenetic correlations among fungi species, with both genome size and GC content as  
 673 predictors in explaining the variation of fungal tissue N and P. Statistical results and  
 674 explained variation for each predictor were provided in Table S3.

675 **Fig. 3**

676

677 **Fig. 3** Correlations between community weighted mean genome size (CWM-GS, Mbp/1C,  
 678 log-transformed) or GC content (CWM-GC, %) and soil available nutrients (NH<sub>4</sub><sup>+</sup>, NO<sub>3</sub><sup>-</sup>,  
 679 PO<sub>4</sub><sup>-</sup>) for soils from grassland (black), woodland (red), and forest (blue) samples across  
 680 Australia; 463 soil samples were collected from 409 locations. Correlations were fitted with a  
 681 linear mixed model that included soil available nutrients as fixed effects and with DNA  
 682 region sequenced (ITS1 or ITS2) as random effects. Points are darker with more overlapping.  
 683 Solid lines indicate the correlations were significant ( $P < 0.05$ ) while dashed lines indicate  
 684 non-significant (model fitting results were supplied in Table S4).

685 **Fig. 4**

686  
 687 **Fig. 4** Relative abundance of fungal guilds, guild-level community weighted mean genome  
 688 size (CWM-GS) and GC (CWM-GC) content in relation to increasing soil nutrient  
 689 availability for ectomycorrhizal (EcM, green), plant pathogen (blue), and saprotrophic  
 690 (purple) fungi across Australia. Bars display the effect size of these responses, i.e., the  
 691 estimated slopes from models that were fitted with one of the soil nutrients (NH<sub>4</sub><sup>+</sup>, NO<sub>3</sub><sup>-</sup>,  
 692 PO<sub>4</sub><sup>-</sup>) as the predictor for grassland, woodland and forest samples, respectively. Parameters  
 693 were estimated with a linear mixed model that included the DNA region sequenced (ITS1 or  
 694 ITS2) as random effects (detailed results were supplied in Fig. S4-6 and Table S4-6). In each  
 695 panel, positive slopes indicate response variables increased with increasing soil nutrient  
 696 availability while negative slopes indicate the opposite pattern. Solid bars indicate the  
 697 responses were significant ( $P < 0.05$ ) while faded bars indicate responses were not  
 698 significant.