

<https://helda.helsinki.fi>

Diversity of indoor fungi as revealed by DNA metabarcoding

Korpelainen, Helena

2017-01

Korpelainen , H & Pietiläinen , M 2017 , ' Diversity of indoor fungi as revealed by DNA metabarcoding ' , Genome , vol. 60 , no. 1 , pp. 55-64 . <https://doi.org/10.1139/gen-2015-0191>

<http://hdl.handle.net/10138/307915>

<https://doi.org/10.1139/gen-2015-0191>

acceptedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17

Diversity of indoor fungi as revealed by DNA metabarcoding

Helena Korpelainen, and Maria Pietiläinen

Department of Agricultural Sciences, P.O. Box 27 (Latokartanonkaari 5), FI-00014 University of Helsinki, Finland

Corresponding author: Helena Korpelainen (e-mail helena.korpelainen@helsinki.fi).

1 **Abstract:** In the present study, we conducted DNA metabarcoding (the nuclear ITS2 region)
2 for indoor fungal samples originating from two nursery schools with a suspected mould problem
3 (sampling before and after renovation), from two university buildings and from an old
4 farmhouse. Good-quality sequences were obtained, and the results showed that DNA
5 metabarcoding provides high resolution in fungal identification. The numbers of fungal classes,
6 orders, families and genera per sample varied greatly among sampling sites (pooled results per
7 building) and times, between 12-21, 15-58, 20-118 and 29-248, respectively. Comparable ranges
8 of Shannon's diversity indices were 0.47-2.12, 0.65-2.91, 0.82-3.30 and 0.87-3.59, respectively.
9 The pooled proportions of filamentous ascomycetes, filamentous basidiomycetes, yeasts and
10 other fungi equalled 62.3%, 8.0%, 28.3% and 1.4%, respectively, and the total number of fungal
11 genera found during the study was 585. When comparing fungal diversities and taxonomic
12 composition between different types of buildings, no obvious pattern was detected. The average
13 pairwise values of Sørensen_{Chao} indices that were used to compare similarities for taxon
14 composition between samples among the samples from the two university buildings, two
15 nurseries and farmhouse equaled 0.693, 0.736, 0.852, 0.928 and 0.981, respectively, while the
16 mean similarity index for all samples was 0.864. We discovered that making explicit conclusions
17 on the relationship between the indoor air quality and mycoflora is complicated by the lack of
18 appropriate indicators for air quality and by the occurrence of wide spatial and temporal changes
19 in diversity and compositions among samples.

20

21 *Key words:* indoor air quality, fungi, metabarcoding, next generation sequencing

1 **Introduction**

2

3 Environmental microbes can have both beneficial and harmful effects on health, and the
4 interactions between environment, microbiota and health may be complicated. For instance,
5 biodiversity of bacteria is believed to be an important factor explaining the lower incidence of
6 allergic diseases in children living in high-biodiversity conditions (rural environments) when
7 compared to children living in urban environments with lower biodiversity (Hanski et al. 2012;
8 Ruokolainen et al. 2015). Such an environmental effect may be mediated via the effect of
9 environmental microbiota on the commensal microbiota influencing immunotolerance. Based on
10 the proposed biodiversity hypothesis, early exposure to an environment with high biodiversity
11 may prevent the development of allergic diseases (Hanski et al. 2012; Ruokolainen et al. 2015).

12 People spend most of their time in indoor environments, which contain a variety of
13 microbes. Serious problems may develop in buildings with long-lasting dampness, where the
14 moisture supports the growth of bacteria and fungi (i.e., mould). Based on epidemiological
15 studies, mould in buildings is positively associated with several allergic and respiratory effects,
16 and certain moulds are toxigenic, meaning that they can produce mycotoxins (Fisk et al. 2007;
17 Mendell et al. 2011; Jacobs et al. 2014). There are estimates that allergic diseases caused by
18 plant, animal and fungal allergens affect more than 30% of the population in industrialized
19 countries (Cramer et al. 2013), and there is increasing awareness and concern over exposure to
20 moulds in indoor environments. The phenomenon has become known as Sick Building
21 Syndrome (SBS), where the occupants describe a complex range of vague and often subjective
22 health complaints (Jones 1999). Since a presumed mould problem may lead to expensive
23 renovations or even to the abandonment of buildings, it is important to be able to evaluate the
24 mould situation, as well as the potential presence of other indoor air pollutants, correctly and
25 precisely.

1 Indoor fungi are traditionally determined by culture-dependent methods (e.g. Ebbelhøj et
2 al. 2002), which have a low taxonomic resolution, underestimate diversity, and bias results
3 towards fungi that grow well on generic growth media and produce characteristic morphological
4 structures allowing identification. In fact, there are many cryptic fungal species that cannot be
5 distinguished morphologically or based on reproductive characteristics (Sato and Murakami
6 2008; Brown et al. 2013). Presently, taxon-specific microbial markers combined with
7 quantitative PCR methods are also used for identifying fungal specimens (e.g. Simoni et al.
8 2011; Jacobs et al. 2014).

9 Recent advances in DNA sequencing provide an effective tool for species detection and
10 biomonitoring using DNA present in the environment. Specifically, DNA metabarcoding
11 through high-throughput sequencing (next generation sequencing) allows the characterization of
12 the species composition of bulk samples, including both intact and degraded DNA extracted
13 from environmental samples (eDNA, i.e., cellular DNA from living cells or organisms and
14 extracellular DNA resulting from cell death and subsequent destruction of cell structure)
15 (Taberlet et al. 2012; Bohmann et al. 2014; Yang et al. 2014; Valentini et al. 2015).
16 Metabarcoding uses universal PCR primers to mass-amplify a taxonomically informative gene
17 from mass collections of organisms or from environmental DNA.

18 In the present study, to increase precision in analyses and to provide useful data and tools
19 for end-users on the environmental quality of indoor spaces, and to discover existing biodiversity
20 in indoor fungal communities, we conducted DNA barcoding (the nuclear ITS2 region) for
21 indoor fungal samples. The internal transcribed spacer region (ITS, comprising spacers ITS1 and
22 ITS) of the nuclear ribosomal DNA is the formal DNA barcoding region for molecular
23 identification of fungi (Schoch et al. 2012). It has been shown that ITS1 and ITS2 yield closely
24 similar results when used as DNA barcodes for fungi (Blaalid et al. 2013). Thus, the use of ITS2
25 in fungal metabarcoding is justified. In this study, we wanted to test, how effective DNA

1 barcoding is when analysing the taxonomic diversity of fungal communities in indoor spaces.
2 The additional novelty was that we collected samples from different types of buildings and
3 included multiple samples from each building at different time points (i.e., a longitudinal study
4 approach). Two buildings were sampled both before and after renovation in order to discover,
5 whether the renovation affected the fungal composition.

6

7

8 **Materials and methods**

9

10 Samples were collected from five buildings, including two university buildings, two
11 nursery schools and an old inhabited farmhouse (field crop production nearby). The farmhouse is
12 located in the county of Porvoo, about 40 km to the east from Helsinki, while other buildings are
13 located in Helsinki (about 60°14' N, 25°01' E). Sampling was conducted four times: January
14 2013, July 2013, January 2014, and July 2014. All buildings were not sampled on every occasion
15 (Table 1). Both nursery schools were renovated during the study due to respiratory symptoms
16 reported by some employees and minor visible water damage and mould growth, and we
17 sampled them both before and after renovation, which primarily included changes in surface
18 materials and improved ventilation. Sampling was performed using a collector with a disposable
19 filter (DUSTREAM Collector, Indoor Biotechnologies Inc., Charlottesville, VA, USA; mesh size
20 40 µm) attached to the tube of a vacuum cleaner with the suction power of 32 L/s. Both a
21 horizontal (tables or shelves) and vertical (walls) sample were collected by vacuuming an area of
22 about 2 m²/sample (i.e., two 2 m² samples per room) from two rooms in each of five buildings
23 (two office rooms in each university building, two playrooms in each nursery school, and two
24 bedrooms in the old farmhouse; nursery schools and the farmhouse had suspected mould
25 problems).

1 After vacuuming, the filter containing the dust was removed from the collector and
2 placed in a plastic bag until processing. In a fume hood in the lab, filters were cut, rinsed with
3 water, and the dust and water were poured into a petri dish, where large non-biological particles
4 were removed. Thereafter, the samples were dipped in liquid nitrogen and ground in a ball mill,
5 and DNA was extracted using the CTAB (cetyl trimethylammonium bromide) method (Doyle
6 and Doyle 1987). The final volume was 100 µl.

7 For the metabarcoding of the fungal samples, genomic ITS2 sequences were amplified
8 and sequenced using two approaches. All sequencing was conducted at the DNA Sequencing and
9 Genomics Laboratory, Institute of Biotechnology, University of Helsinki. The sequencing for the
10 samples from January and July 2013 were conducted using 454 FLX pyrosequencing (Roche
11 Applied Science, Penzberg, Germany), as described in Korpelainen et al. (2015). The following
12 primer systems were used:

13 1) Reverse *ITS3_Ampl_B* (adapter + ITS3 primer) [the same one in all reactions]

14 5'-CTATGCGCCTTGCCAGCCCGCTCAG + GCATCGATGAAGAACGCAGC-3'

15

16 2) Forward *ITS4_Ampl_A+Tg* (adapter + tag marker (6 bp) + ITS4 primer) [different tag marker
17 alternatives], for instance

18 5'-CGTATGCGCCTCCCTCGCGCCATCAG + TCTGTA + TCCTCCGCTTATTGATATGC-3'

19

20 The used tag marker sequences were as follows: TCTGTA, CTA CTG, CAGCTC, ATCATG, AGATAT,
21 CGACGC, CATGCA and TCTATG.

22

23 However, sequencing for the samples from January and July 2014 were performed using
24 Illumina MiSeq sequencing (San Diego, CA, USA), for which ITS2 sequences were first
25 amplified using the following primer system (forward ITS4 mix + reverse ITS3 mix):

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

Forward ITS4 mix including three primers:

- ITS4_F1 5'-ATCTA CACTCTTTCCCTACACGACGCTCTTCCGATCTTCCTCCGCTTATTGATATGC-3'
- ITS4_F2 5'-ATCTA CACTCTTTCCCTACACGACGCTCTTCCGATCT(c/g)TCCTCCGCTTATTGATATGC-3'
- ITS4_F3 5'-ATCTA CACTCTTTCCCTACACGACGCTCTTCCGATCTagt(a/g)(a/g)TCCTCCGCTTATTGATATGC-3'

Reverse ITS3 mix including three primers:

- ITS3_R1 5'-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGCATCGATGAAGAACGCAGC-3'
- ITS3_R2 5'-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT(c/t)GCATCGATGAAGAACGCAGC-3'
- ITS3_R3 5'-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTa(a/t)GCATCGATGAAGAACGCAGC-3'

All 20- μ l PCR reactions contained 2 μ l of template DNA, and the concentration of each primer was 0.25 μ M. All PCR products were gel-purified (Omega Bio-Tek Gel Extraction Kit, Norcross, Georgia, USA). After Illumina sequencing, primer sequences were removed from the raw reads, and quality control, as described by Brown et al. (2013), followed. During this process, low-quality reads (below average PHRED score of 25) and short sequences (< 100 bp) were removed. Then, all other sequence data were subjected to similarity search against GenBank (www.ncbi.nlm.nih.gov/genbank), and assignment of taxonomic identities using TAXAassign (<https://github.com/umerijaz/taxaassign>) was conducted with 60, 70, 80, 95, 95, and 97% thresholds for different taxonomic ranks, which may correspond to phylum, class, order, family, genus and species levels, respectively. However, these thresholds are tentative and should be treated with special caution, except for the 97% threshold, which is, by convention, used as a divergence threshold for operational taxonomic units (OTUs) that serve as a proxy for species (Brown et al. 2015). The sequence data were submitted to the EMBL (European Molecular Biology Laboratory) database under accession number PRJEB8345. Based on the numbers of sequences representing each taxon (i.e., taxon distribution), Shannon's diversity

1 indices (Shannon 1948) were calculated at class, order, family, and genus level for each sample.
2 To compare the taxonomic composition of pooled samples (each including all four samples from
3 a building at the same time point), a principal component analysis (PCA) was conducted for the
4 ITS2 sequence-based generic data (fungal genera and their frequencies) using SAS 9.4 software
5 (SAS Institute Inc., Cary, NC, USA). In addition, EstimateS 9.1.0. (<http://purl.oclc.org/estimates>)
6 was used to calculate similarities for taxon composition between pooled longitudinal samples
7 from the same site and between all pooled samples. The used estimator was the Sørensen_{Chao}
8 abundance-based similarity index (corrected for unseen shared species), which can also handle
9 different sample sizes (Chao et al. 2005).

10

11

12 **Results**

13

14 Relatively small-scale pyrosequencing was conducted for the first two sets of samples
15 (winter 2013 and summer 2013), and the number of good sequences averaged 6569 and 4967
16 sequences/sample (original 2 m² sample), respectively. For the last two sets of Illumina-
17 sequenced samples (winter 2014 and summer 2014), the number of good sequences averaged
18 213 894 and 558 756 sequences/sample (original 2 m² sample), respectively. Of all samples,
19 85.8% were successfully assigned to the genus level and 49.0% to the species level. We present
20 diversity and taxonomic information based on genus level data unless otherwise specified.

21 Fungal taxa per sample varied greatly among sampling sites (pooled results per building)
22 and times. Based on 454 FLX pyrosequencing (January and July 2013 samples), numbers of taxa
23 were as follows; 12-21 classes, 15-58 orders, 20-114 families, and 29-176 genera, while based
24 on Illumina MiSeq sequencing (January and July 2014 samples), the numbers were as follows:
25 15-19 classes, 46-58 orders, 82-118 families, and 144-248 genera (tentative classification; Table

1 1). Comparable ranges of Shannon's diversity indices for 454 FLX pyrosequenced data were
2 0.88-2.12, 1.09-2.91, 1.18-3.30 and 1.18-3.59, respectively, and for Illumina MiSeq data 0.47-
3 1.87, 0.65-2.57, 0.82-3.00, and 0.87-3.52. The total number of fungal genera found during the
4 study was 585. The change of the sequencing method did not result in increased diversity indices
5 (genus level, determined for individual 2 m² samples, mean±standard error; 2.79±1.00 in 2013
6 and 2.56±0.94. However, there was an increase in the numbers of detected taxa, with a mean of
7 89.1±57.7 genera in 2013 and 122.4±57.7 genera in 2014 (t=2.627, df=54, P=0.012). Thus,
8 apparently a more comprehensive coverage of low-frequency taxa was obtained using Illumina
9 sequencing. The numbers of taxa and diversity indices were similar among individual horizontal
10 (100.2±41.2 genera; 2.83±0.65) and vertical (105.7±56.7 genera; 2.55±1.16)and, and among
11 winter (105.2±51.1 genera; 2.58±1.11) and summer samples (102.5±44.6 genera; 2.77±0.88). In
12 Nursery 1, which underwent a small-scale renovation, the diversity index did not change, while
13 in Nursery 2, which was renovated extensively, the diversity index increased from 1.18 to 3.07.

14 When the fungal taxa detected in each sample were divided into four groups, filamentous
15 ascomycetes, filamentous basidiomycetes, yeasts, and other fungi, the results showed great
16 variation in proportions among sampling sites and times (Fig. 1), and no detectable patterns
17 among samples within and between buildings were found. The proportions of sequences
18 corresponding to filamentous ascomycetes, filamentous basidiomycetes, yeasts, and other fungi
19 were 62.3%, 8.0%, 28.3% and 1.4%, respectively.

20 Table 2 lists the five most frequent fungal taxa detected in each building at each
21 sampling time. Besides genus, the species name is given in the case of a species-level
22 identification. The results show that there was a seasonal turnover in the proportions of dominant
23 taxa, except for the farmhouse, in which *Cyberlindnera jadinii* and *Candida* sp. were the two
24 most frequent taxa at both sampling times (summer 2013 and winter 2014). *Aureobasidium*
25 *pullulans* was in the top five taxa in 9 out of 16 building samples, *Cladosporium* sp. in 7

1 samples, and *Cryptococcus* sp. and *Saccharomyces cerevisiae* in 6 samples each (Table 2). In
2 several samples, one specific taxon highly dominated, such as *S. cerevisiae* in University 1 and
3 Nursery 2 in winter 2013 (47.5% and 45.4%, respectively), *C. jadinii* in the farmhouse in
4 summer 2013 (38.7%), and *Preussia* sp. in University 1 in summer 2014 (85.7%). Overall, the
5 most frequent taxa were *A. pullulans* (10.5%) and *S. cerevisiae* (7.8%).

6 The taxonomic content of the 14 pooled fungal samples, based on the ITS2 sequence
7 data, was analyzed using principal component analysis (PCA) (Fig. 2). Two components
8 explained 57.7% of the variability. This projection of taxonomic data also confirmed the
9 presence of a great temporal turnover in the composition of samples, except for the two pooled
10 farmhouse samples (F-S13, summer 2013; F-W14, winter 2014). Otherwise, Fig. 2 does not
11 show any seasonal pattern or any definite pattern in the fungal composition of different types of
12 buildings. Correspondingly, Sørensen_{Chao} indices that were calculated to compare similarities for
13 taxon composition between samples did not show any clear pattern. The average pairwise values
14 for temporal pooled samples among University 1, University 2, Nursery 1, Nursery 2 and
15 farmhouse samples equaled 0.693, 0.736, 0.852, 0.928 and 0.981, respectively, while the mean
16 similarity index for all samples was 0.864. The similarity indices of Nursery 1 and Nursery 2
17 samples for before and after renovation samples equaled 0.739 and 0.928, respectively. For
18 comparison, Shannon's genus-level diversity indices of Nursery 1 equaled 3.42 and 3.41 before
19 and after renovation, and those of Nursery 2 equaled 1.18 and 3.07, respectively.

20

21

22 **Discussion**

23

24 The present study shows that DNA metabarcoding gives a good resolution in fungal
25 identification. The used method is highly effective until the genus level identification of fungi

1 (85.8%) and reasonably effective for species identification (49.0%). The change of sequencing
2 method from 454 FLX pyrosequencing to Illumina MiSeq sequencing and resulting 70-fold
3 increase in sequence numbers are suggested to increase detection of infrequent taxa, with the
4 mean number of genera per building increasing from 89 to 122. However, year-to-year variation
5 may also contribute to changes in taxon numbers. Considering sequencing platform qualities,
6 Kozich et al. (2013) have demonstrated that Illumina MiSeq platform can provide data that are at
7 least as good as that generated by the 454 platform while providing considerably higher
8 sequencing coverage at a lower cost. Previously, Pitkäranta et al. (2011) have shown that
9 molecular profiling may reveal a five to ten times higher diversity at the genus level than culture-
10 based methods. However, we do not know what fungal diversity culture-dependent methods
11 might reveal in the buildings studied here.

12 Fungal diversities in samples collected from different buildings (university offices,
13 nursery schools, farmhouse, supposed with or without a mould problem), during different
14 seasons (summer vs. winter) or using different sampling methods (horizontal vs. vertical surface)
15 showed considerable variation and turnover but no definite pattern. Previously, Adams et al.
16 (2013) surveyed temporal variation in airborne fungal assemblages, both indoors and outdoors,
17 using ITS1 pyrosequencing. They discovered that indoors fungal assemblages were diverse and
18 strongly determined by dispersal from outdoors, and no fungal taxa were found as indicators of
19 indoor air quality. Also, human occupancy has been shown to result in significantly elevated
20 airborne bacterial and fungal concentrations as compared to vacant conditions (Hospodsky et al.
21 2015).

22 The considerable variation in fungal composition found to occur even within the same
23 building emphasizes the importance of multiple sampling. Also, a diverse array of fungi occurred
24 even in a normal indoor environment (recently renovated University 1 and relatively new
25 University 2) considered to have a good indoor air quality. Nursery 1 and Nursery 2, with some

1 moisture damage and employees suffering from possibly mould-related symptoms, possessed
2 highly divergent diversity indices before renovation, 3.41 and 1.18, respectively, but similar
3 values after renovation, 3.41 and 3.07, respectively. The Sørensen_{Chao} similarity index for before
4 and after comparisons of taxon composition was lower for Nursery 1 (0.739) than for Nursery 2
5 (0.928). Cleaning frequency and coverage in different buildings may also contribute to the
6 observed fungal diversities, although all studied buildings are cleaned regularly. When
7 examining fungal diversity by metabarcoding using ITS1 in different apartments in South Korea,
8 An and Yamamoto (2016) observed that Shannon diversity indices were variable but quite low,
9 ranging from 0.14 to 2.29 (mean = 1.11) in indoor spaces considered alike. In addition,
10 renovation may not instantly affect the fungal and bacterial composition, as shown by Emerson
11 et al. (2015), who compared flood-damaged and non-flooded homes. The flooded homes had
12 higher fungal abundances, and the bacterial and fungal communities continued to be affected by
13 flooding, even after relative humidity had returned to baseline levels and remediation had
14 removed any visible evidence of flood damage.

15 The presence or dominance of fungal taxa known to cause allergic and respiratory effects
16 and/or being indicators of moisture damage could show something of the air quality. Of the 78
17 fungal genera listed by Simon-Nobbe et al. (2007) to have been shown to induce allergies in
18 atopic (hypersensitive to allergens) individuals, 51 (65%) were found in this study, although
19 most at very low frequencies. Among them, 11 genera were found at the frequency of more than
20 1% in the whole data set, namely the filamentous ascomycetes *Aspergillus*, *Aureobasidium*,
21 *Chaetomium*, *Cladosporium*, *Epicoccum*, *Leptosphaeria* and *Penicillium*, and the yeasts
22 *Candida*, *Malassezia*, *Rhodotorula* and *Saccharomyces*. Previously, An and Yamamoto (2016)
23 found several allergy-related genera in apartments in South Korea, where the most abundant
24 genera were *Cladosporium*, *Crivellia*, *Rhodotorula* and *Alternaria*. Among university samples,
25 the most common fungal taxa were the filamentous ascomycetes *Aureobasidium pullulans* and

1 *Preussia* sp., and the yeasts *Saccharomyces cerevisiae*, *Cyberlindnera jadinii*, *Rhodotorula* sp.
2 and *Cryptococcus* sp., of which *A. pullulans*, *S. cerevisiae* and *Rhodotorula* are listed among the
3 allergy-inducing fungi (Simon-Nobbe et al. 2007). Among nursery school samples, the most
4 common taxa were filamentous ascomycetes *Cadophora*, *A. pullulans* and *Pestalotiopsis*, and
5 the yeast *S. cerevisiae* and *Cryptococcus* sp., of which *A. pullulans* and *S. cerevisiae* are
6 presumed to induce allergies (Simon-Nobbe et al. 2007). In the farm samples, the yeasts *C.*
7 *jadinii* and *Candida* sp. were the dominant taxa. Among these fungi, *Candida* may cause
8 clinically significant infections (Simon-Nobbe et al. 2007). We also determined the percentages
9 of sequences representing presumably allergy-inducing fungi in each of the 14 pooled samples
10 based on the classification of Simon-Nobbe et al. (2007). The percentages ranged among
11 university samples between 8-81%, among nursery schools before and after renovation between
12 32-70% and 54-56%, respectively, and in the two farmhouse samples the proportions equalled 54
13 and 56%. Thus, no obvious pattern was detected. Yet, the view of the nursery school having a
14 mould problem may be subjective, as there has not been a proper medical examination for its
15 users. Also in previous studies, which have been generally based on a narrower range of
16 information, there have been difficulties in the interpretation of relationships between mycoflora
17 and allergic symptoms (e.g. Simoni et al. 2011; Jacobs et al. 2014). On the other hand,
18 Dannemiller et al. (2014) demonstrated significant associations between low fungal diversity and
19 childhood asthma development in a low-income, Mexican immigrant community in the USA.
20 However, one characteristic of populations of Mexican descent is low asthma prevalence rates
21 compared with other racial/ethnic groups in the USA (Lara et al. 2006).

22 DNA metabarcoding is a very promising approach to biodiversity investigations, and its
23 effectiveness to recover the diversity present in mixed-species samples has been already tested
24 for a range of organisms and environments, and even for poor-quality and low-quantity DNA
25 (e.g. Taberlet et al. 2012; Bohmann et al. 2014; Valentini et al. 2015). However, there are also

1 potential limitations and difficulties that should be considered, such as errors during PCR and
2 sequencing, quantitative assessment of different organisms, sequence coverage of reference
3 databases (although fast improving), and species with incomplete lineage sorting for the barcode
4 markers, which can lead to errors in identification. Yet, the increased and all the time improving
5 precision obtained through DNA metabarcoding provides a highly potential tool for analysing,
6 for example, indoor mycoflora. However, the full interpretation of even very accurate
7 biodiversity results can be challenging. For instance, in the present study, we discovered that
8 making explicit conclusions on the relationship between the indoor air quality and mycoflora is
9 complicated by the occurrence of wide changes in spatial and temporal diversities and
10 compositions among samples. In future DNA metabarcoding studies, a wider range of buildings,
11 both with and without mould problems, should be investigated to allow deeper insights into the
12 air quality issue of indoor space. In addition, the whole problem concerning fungi and other
13 microbes in indoor air may be closely linked with decreases in the overall biodiversity and
14 consequent alterations in the indigenous microbiota, and increased susceptibility to allergies.
15 Evidence for the biodiversity hypothesis proposing a connection between biodiversity and
16 allergic diseases has been provided in several recent investigations (e.g. Hanski et al. 2012;
17 Ruokolainen et al. 2015).

18

19

20 **Acknowledgements**

21 We thank the City of Helsinki representatives for providing sampling sites. This work
22 was supported by the Marjatta and Eino Kolli Foundation through a research project 1055 to H.
23 Korpelainen. The authors declare that they have no conflict of interest concerning this study.

24

25

1 **References**

- 2
- 3 Adams, R.I., Miletto, M., Taylor, J.W., and Bruns, T.D. 2013. Dispersal in microbes: fungi in
4 indoor air are dominated by outdoor air and show dispersal limitation at short distances.
5 The ISME Journal **7**: 1262–1273. doi:10.1038/ismej.2013.28.
- 6 An, C., and Yamamoto, N. 2016. Fungal compositions and diversities on indoor surfaces with
7 visible mold growths in residential buildings in the Seoul Capital Area of South Korea.
8 Indoor Air, in press. doi:10.1111/ina.12261.
- 9 Błaalid, R., Kumar, S., Nilsson, R.H., Abarenkov, K., Kirk, P.M., and Kauserud, H. 2013. ITS1
10 versus ITS2 as DNA metabarcodes for fungi. Mol. Ecol. Res. **13**: 218–224.
11 doi:10.1111/1755-0998.12065.
- 12 Bohmann, K., Evans, A., Gilbert, M.T.P., Carvalho, G.R., Creer, S., Knapp, M., et al. 2014.
13 Environmental DNA for wildlife biology and biodiversity monitoring. Trends Ecol. Evol.
14 **29**: 358-367. doi:10.1016/j.tree.2014.04.003.
- 15 Brown, E.A., Chain, F.J.J., Crease, T.J., MacIsaac, H., and Cristescu, M. 2015. Divergence
16 thresholds and divergent biodiversity estimates: can metabarcoding reliably describe
17 zooplankton communities. Ecol. Evol. **5**: 2234-2251. doi:10.1002/ece3.1485.
- 18 Brown, E.M., McTaggart, L.R., Zhang, S.X., Low, D.E., Stevens, D.A., and Richardson, S.E.
19 2013. Phylogenetic analysis reveals a cryptic species *Blastomyces gilchristii*, sp. nov.
20 within the human pathogenic fungus *Blastomyces dermatitidis*. PLoS ONE **8(3)**: e59237.
21 doi:10.1371/journal.pone.0059237.
- 22 Brown, S.P., Callahan, M.A., Oliver, A.K., and Jumpponen, A. 2013. Deep Ion Torrent
23 sequencing identifies soil fungal community shifts after frequent prescribed fires
24 in a southeastern US forest ecosystem. FEMS Microbiol Ecol. **86**: 557-566. doi:10.1111/
25 1574-6941.12181.

- 1 Chao, A., Chazdon, R.L., Colwell, R.K., and Shen, T.-J. 2005. A new statistical approach for
2 assessing similarity of species composition with incidence and abundance data. *Ecology*
3 *Letters* **8**: 148–159. doi:10.1111/j.1461-0248.2004.00707.x.
- 4 Cramer, R., Garbani, M., Rhyner, C. and Huitema, C. 2014. Fungi: the neglected allergenic
5 sources. *Allergy* **69**: 176–185. doi:10.1111/all.12325.
- 6 Dannemiller, K.C., Mendell, J.M., Macher, J.M., Kumagai, K., Bradman, A., Holland, N. et al.
7 2014. Next-generation DNA sequencing reveals that low fungal diversity in house dust is
8 associated with childhood asthma development. *Indoor Air* **24**: 236–247.
9 doi:10.1111/ina.12072.
- 10 Doyle, J.J., and Doyle, J.L. 1987. A rapid DNA isolation procedure for small quantities of
11 fresh leaf tissue. *Phytochem. Bull.* **19**: 11-15.
- 12 Ebbenhøj, N.E., Hansen, M.Ø., Sigsgaard, T., and Larsen L. 2002. Building-related symptoms and
13 molds: a two-step intervention study. *Indoor Air* **12**: 273–277. doi:10.1034/j.1600-
14 0668.2002.02141.x
- 15 Emerson, J.B., Keady, P.B., Brewer, T.E., Clements, N., Morgan, E.E., Awerbuch, J., et al.
16 2015. Impacts of flood damage on airborne bacteria and fungi in homes after the 2013
17 Colorado Front Range Flood. *Envir. Sci. Technol.* **49**: 2675-2684.
18 doi:10.1021/es503845j.
- 19 Fisk, W.J., Lei-Gomez, Q, and Mendell, M.J. 2007. Meta-analyses of the associations of
20 respiratory health effects with dampness and mold in homes. *Indoor Air* **17**: 284-296.
21 doi:10.1111/j.1600-0668.2007.00475.x.
- 22 Hanski, I., von Hertzen, L., Fyhrquist, N., Koskinen, K., Torppa, K., Laatikainen, T., et al. 2012.
23 Environmental biodiversity, human microbiota, and allergy are interrelated. *Proc. Natl.*
24 *Acad. Sci. USA* **109**: 8334–8339. doi:10.1073/pnas.1205624109.
- 25 Hospodsky, D., Yamamoto, N., Nazaroff, W. W., Miller, D., Gorthala, S. and Peccia, J. 2015.

1 Characterizing airborne fungal and bacterial concentrations and emission rates in six
2 occupied children's classrooms. *Indoor Air* **25**: 641–652. doi:10.1111/ina.12172.

3 Jacobs, J., Borràs-Santos, A., Krop, E., Täubel, M., Leppänen, H., Haverinen-Shaughnessy, U. et
4 al. 2014. Dampness, bacterial and fungal components in dust in primary schools and
5 respiratory health in schoolchildren across Europe. *Occup. Environ. Med.* **71**: 704-712.
6 doi:10.1136/oemed-2014-102246.

7 Jones, A.P. 1999. Indoor air quality and health. *Atmosph. Envir.* **33**: 4535-4564.
8 doi:10.1016/S1352-2310(99)00272-1.

9 Korpelainen, H., Pietiläinen, M., and Huotari, T. 2015. Effective detection of indoor fungi by
10 metabarcoding. *Ann. Microbiol.* **66**: 495-498. doi:10.1007/s13213-015-1118-x.

11 Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K., and Schloss, P.D. 2013.
12 Development of a dual-index sequencing strategy and curation pipeline for analyzing
13 amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Environ.*
14 *Microbiol.* **79**: 5112-5120. doi:10.1128/AEM.01043-13.

15 Lara, M., Akinbami, L., Flores, G. and Morgenstern, H. 2006. Heterogeneity of childhood
16 asthma among Hispanic children: Puerto Rican children bear a disproportionate burden,
17 *Pediatrics* **117**: 43–53. doi:10.1542/peds.2004-1714.

18 Mendel, M.J., Mirer, A.G., Cheung, K., Tong, M., and Douwes, J. 2011. Respiratory and allergic
19 health effects of dampness, mold, and dampness-related agents: a review of the
20 epidemiologic evidence. *J. Environ. Health Perspect.* **119**: 748-756.
21 doi:10.1289/ehp.1002410.

22 Pitkäranta, M., Meklin, T., Hyvärinen, A., Nevalainen, A., Paulin, L., Auvinen, P., Lignell, U,
23 and Rintala, H. 2011. Molecular profiling of fungal communities in moisture damaged
24 buildings before and after remediation – a comparison of culture-dependent and culture-
25 independent methods. *BMC Microbiology* **11**: 235. doi:10.1186/1471-2180-11-235.

- 1 Ruokolainen, L., von Hertzen, L., Fyhrquist, N., Laatikainen, T., Lehtomäki, J., Auvinen, P. et
2 al. 2015. Green areas around homes reduce atopic sensitization in children. *Allergy* **70**:
3 195–202. doi:10.1111/all.12545.
- 4 Sato, H., and Murakami, N. 2008. Reproductive isolation among cryptic species in the
5 ectomycorrhizal genome *Strobilomyces*: population-level caps marker-based genetic
6 analysis. *Mol. Phyl. Evol.* **48**: 326-334. doi:10.1016/j.ympev.2008.01.033.
- 7 Schoch, C.L., Seifert, K., Huhndorf, S., Robert, V., Spougea, J.L., Levesque, C.A., et al. 2012.
8 Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode
9 marker for Fungi. *Proc. Natl. Acad. Sci. USA* 109: 6241–6246.
10 doi:10.1073/pnas.1117018109
- 11 Shannon CE (1948) A mathematical theory of communication. *Bell System Techn J*
12 **27**: 379–423, 623–656
- 13 Simon-Nobbe, B., Denk, U., Pöll, V., Rid, R., Breitenbach, M. 2007. The spectrum of fungal
14 allergy. *Int. Arch. Allergy Immunol.* **145**: 58–86. doi:10.1159/000107578.
- 15 Simoni, M., Cai, G.-H., Norback, D., Annesi-Maesano, I., Lavaud, F., Sigsgaard, T. et al. 2011.
16 Total viable molds and fungal DNA in classrooms and association with respiratory health
17 and pulmonary function of European schoolchildren. *Pediatric Allergy and Immunology*
18 **22**: 843–852. doi:10.1111/j.1399-3038.2011.01208.x.
- 19 Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., and Willerslev, E. 2012. Towards
20 next-generation biodiversity assessment using DNA barcoding. *Mol. Ecol.* **21**: 2045-
21 2050. doi:10.1111/j.1365-294X.2012.05470.x.
- 22 Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P.F. et al. 2015. Next-
23 generation monitoring of aquatic biodiversity using environmental DNA metabarcoding.
24 *Mol. Ecol.*, in press. doi:10.1111/mec.13428.
- 25 Yang, C., Wang, X., Miller, J.A., de Blécourt, M., Ji, Y., Yang, C., Harrison, R.D., and Yu,

1 D.W. 2014. Using metabarcoding to ask if easily collected soil and leaf-litter samples can
2 be used as a general biodiversity indicator. *Ecol. Indicators* **46**: 379-389.
3 doi:10.1016/j.ecolind.2014.06.028.

4
5

1 **Figure captions**

2

3 **Fig. 1.** Percentages (%) of sequences corresponding to filamentous ascomycetes, filamentous
4 basidiomycetes, yeasts and other fungi among indoor fungi in five buildings, including two
5 university office buildings, two nursery schools and a farmhouse, based on ITS2 sequences.
6 Each pooled sample included sampling of both horizontal and vertical surfaces. Sequencing for
7 January and July 2013 samples were conducted using 454 FLX pyrosequencing, while
8 sequencing for January and July 2014 samples were performed using Illumina MiSeq
9 sequencing.

10

11 **Fig. 2.** Principal component analysis (PCA) conducted for the pooled fungal samples, based on
12 the ITS2 sequence data. University, nursery and farmhouse samples are marked with grey, black
13 and white dots, respectively. U1 and U2 university samples from buildings 1-2, N1 and N2
14 nursery samples from buildings 1-2, F farmhouse sample; W13, S13, W14, S14, winter and
15 summer samples from years 2013-14. Sequencing for January and July 2013 samples were
16 conducted using 454 FLX pyrosequencing, while sequencing for January and July 2014 samples
17 were performed using Illumina MiSeq sequencing.

Fig. 1.

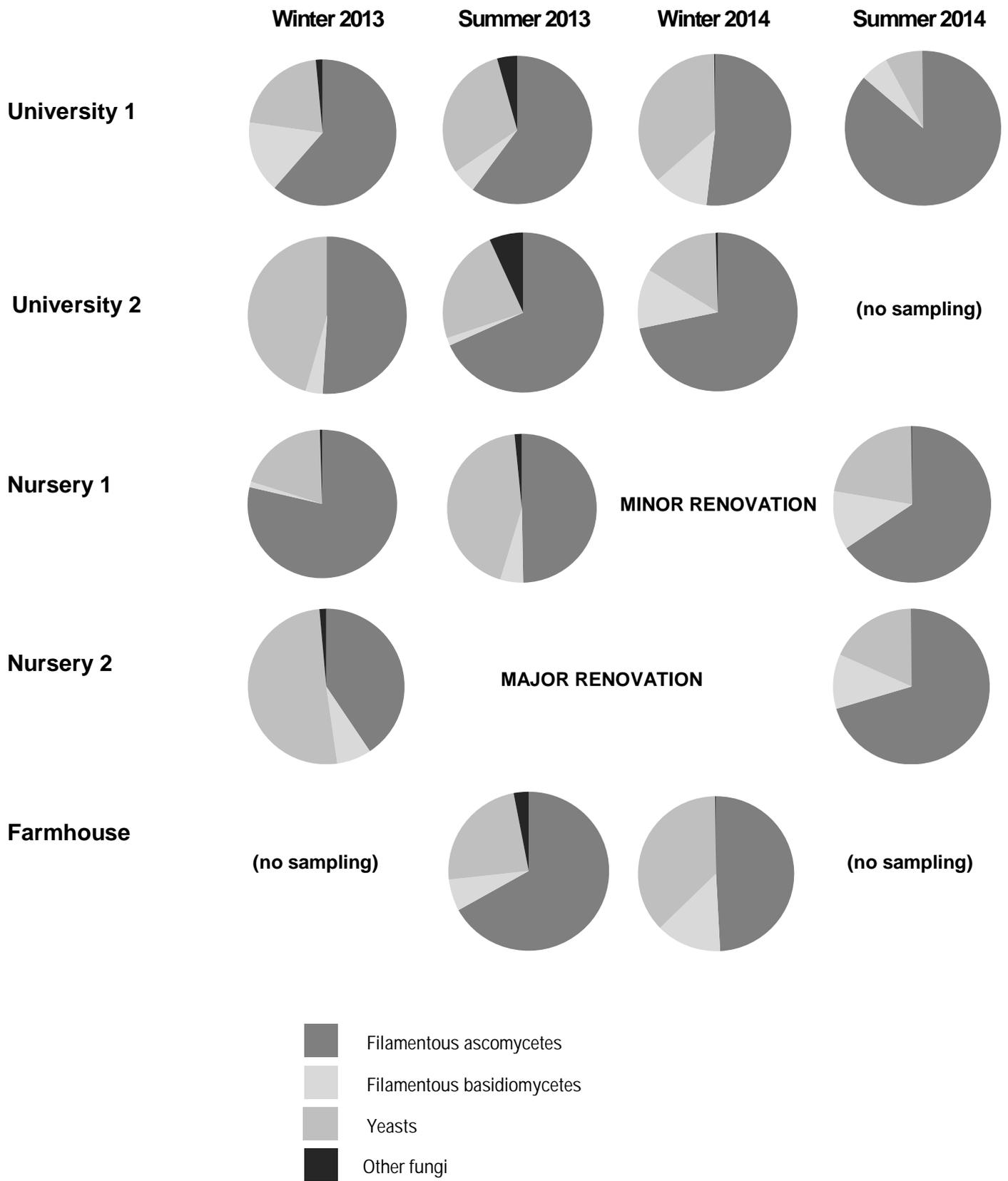


Fig. 2

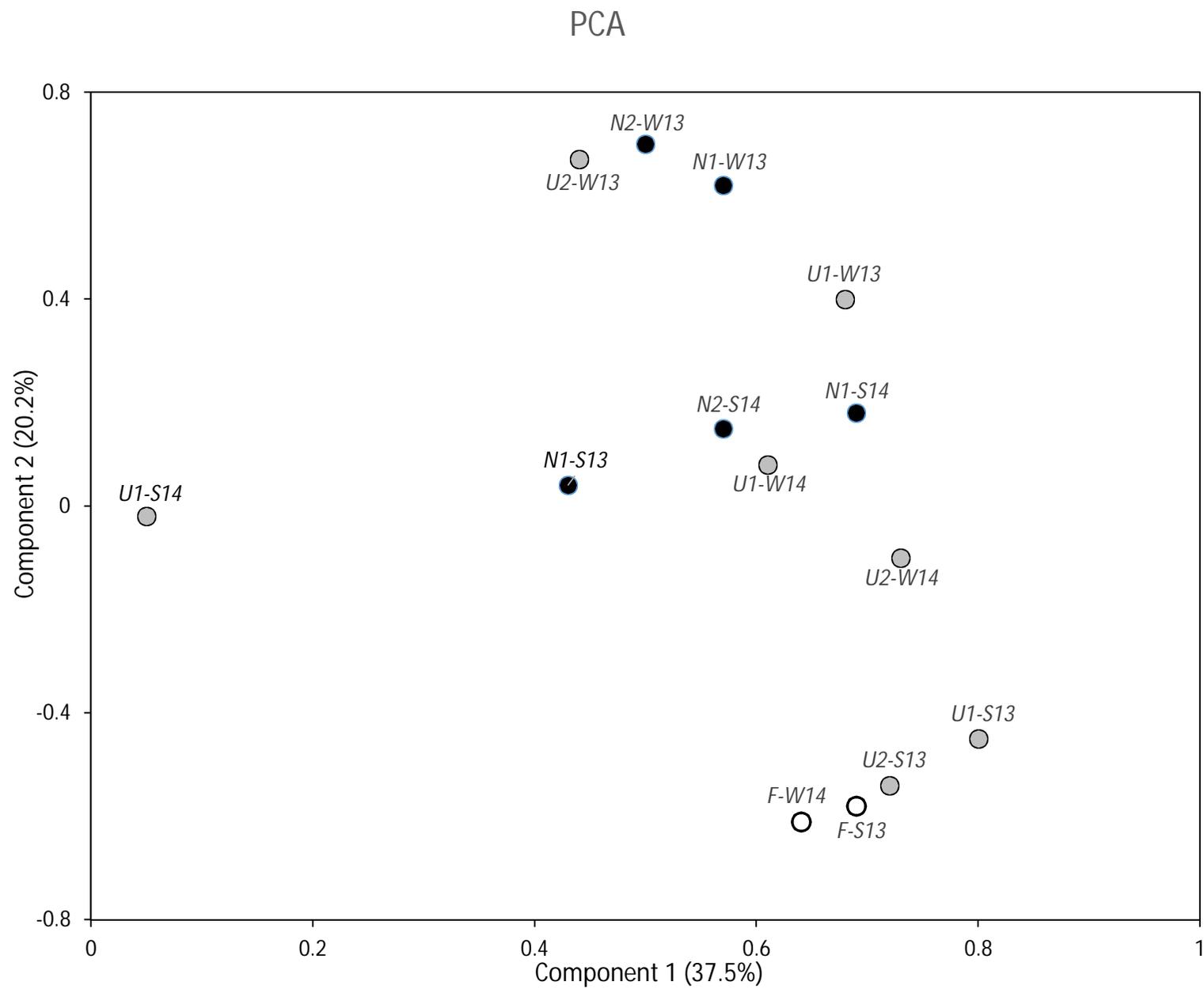


Table 1. Taxonomic diversity of indoor fungi at class, order, family and genus level in five buildings, including two university office buildings, two nursery schools and a farmhouse, based on ITS2 sequences. Each pooled sample included sampling of both horizontal and vertical surfaces. Range of variables among individual samples is given in parentheses. N, number of taxa; H, Shannon's diversity index.

Site		Winter 2013		Summer 2013		Winter 2014		Summer 2014	
		N	H	N	H	N	H	N	H
University 1	Class	19 (11-16)	1.95 (1.08-1.74)	21 (14-17)	2.08 (1.82-2.20)	16 (11-16)	1.62 (0.85-1.84)	15 (12-13)	0.47 (0.17-1.41)
	Order	55 (14-31)	2.66 (1.26-2.44)	57 (29-39)	2.80 (2.39-2.97)	47 (21-41)	2.12 (1.54-2.12)	48 (37-40)	0.65 (0.25-2.16)
	Family	90 (15-46)	2.98 (1.16-2.88)	97 (44-56)	3.30 (2.88-3.56)	82 (28-66)	2.47 (1.28-2.77)	88 (58-73)	0.82 (0.34-2.30)
	Genus	108 (26-61)	3.16 (1.43-3.10)	143 (56-69)	3.59 (2.83-3.49)	146 (40-108)	2.73 (1.63-3.00)	165 (92-122)	0.87 (0.35-2.94)
University 2	Class	12 (8-10)	2.12 (1.70-2.12)	17 (10-17)	2.04 (1.84-2.04)	19 (14-18)	1.87 (1.66-1.87)	(no sampling)	
	Order	15 (10-12)	2.24 (1.87-2.10)	37 (16-33)	2.67 (2.43-2.57)	46 (31-40)	2.57 (2.22-2.52)		
	Family	20 (10-13)	2.45 (1.80-2.16)	55 (16-47)	2.94 (2.49-2.84)	83 (46-71)	3.00 (2.38-2.92)		
	Genus	29 (17-18)	2.86 (2.25-2.54)	73 (27-60)	3.20 (2.96-2.98)	144 (72-111)	3.52 (2.62-3.46)		
Nursery 1	Class	19 (12-15)	1.95 (0.70-1.69)	18 (13-14)	1.92 (1.41-1.75)	(minor renovation, no sampling)		18 (15-17)	1.47 (1.23-1.42)
	Order	58 (22-33)	2.91 (0.80-2.46)	54 (25-34)	2.86 (1.80-2.49)			58 (35-46)	2.47 (1.72-2.39)
	Family	97 (35-44)	3.22 (0.88-2.89)	109 (38-58)	3.25 (1.89-2.80)			118 (60-81)	2.80 (2.01-2.56)
	Genus	133 (50-81)	3.41 (0.78-3.42)	136 (46-80)	3.42 (1.74-3.17)			248 (108-151)	3.41 (2.22-3.05)
Nursery 2	Class	17 (6-17)	0.88 (0.13-2.06)	(major renovation, no sampling)		(major renovation, no sampling)		17 (11-13)	1.43 (1.40-1.41)
	Order	54 (14-45)	1.09 (0.17-2.81)					53 (28-36)	2.47 (1.97-2.42)
	Family	114 (20-83)	1.18 (0.18-3.40)					93 (35-62)	2.64 (2.00-2.58)
	Genus	176 (20-115)	1.18 (0.18-3.65)					170 (54-108)	3.07 (2.11-2.79)
Farmhouse	Class	(no sampling)		17 (12-17)	1.85 (1.36-1.95)	16 (14-16)	1.86 (1.68-1.99)	(no sampling)	
	Order			57 (25-40)	2.48 (1.64-2.67)	50 (28-45)	2.12 (1.93-2.15)		
	Family			104 (43-58)	2.92 (1.72-2.99)	102 (42-90)	2.44 (2.08-2.32)		
	Genus			158 (58-88)	3.01 (1.87-3.32)	198 (65-172)	2.95 (2.54-2.76)		

Table 2. Five most frequent fungal taxa and their proportions (% , in parentheses) in five buildings, including two university office buildings, two nursery schools and a farmhouse, based on ITS2 sequences. Each pooled sample included sampling of both horizontal and vertical surfaces.

Site	Winter 2013	Summer 2013	Winter 2014	Summer 2014
University 1	<i>Aureobasidium pullulans</i> (30.7) <i>Cryptococcus</i> (16.7) <i>Saccharomyces cerevisiae</i> (11.0) <i>Debaromyces hansenii</i> (8.7) <i>Cladosporium</i> (5.8)	<i>Cyberlindnera jadinii</i> (22.4) <i>Candida</i> (11.1) <i>Cladosporium</i> (10.8) <i>Penicillium</i> (6.2) <i>Aureobasidium pullulans</i> (5.0)	<i>Rhodotorula</i> (20.8) <i>Aureobasidium pullulans</i> (15.7) <i>Debaromyces hansenii</i> (12.9) <i>Candida</i> (10.0) <i>Caproventuria hanliniana</i> (6.1)	<i>Preussia</i> (85.7) <i>Penicillium</i> (3.0) <i>Aureobasidium pullulans</i> (1.8) <i>Cryptococcus</i> (1.7) <i>Pyrenophora</i> (0.9)
University 2	<i>Saccharomyces cerevisiae</i> (47.5) <i>Fusarium oxysporum</i> (14.1) <i>Malassezia</i> (7.7) <i>Capnobotryella</i> (3.6) <i>Rhodotorula</i> (3.3)	<i>Cyberlindnera jadinii</i> (23.7) <i>Candida</i> (11.7) <i>Exophiala</i> (6.7) <i>Cladosporium</i> (6.4) <i>Fontanospora</i> (5.3)	<i>Aureobasidium pullulans</i> (13.6) <i>Chaetomium</i> (7.5) <i>Cyberlindnera jadinii</i> (6.6) <i>Penicillium</i> (6.6) <i>Candida</i> (5.6)	(no sampling)
Nursery 1	<i>Cadophora</i> (46.3) <i>Saccharomyces cerevisiae</i> (11.9) <i>Aureobasidium pullulans</i> (6.0) <i>Cladosporium</i> (4.5) <i>Exophiala</i> (2.7)	<i>Pestalotiopsis</i> (20.1) <i>Lasiodiplodia</i> (14.4) <i>Saccharomyces cerevisiae</i> (8.8) <i>Aspergillus</i> (6.6) <i>Cyberlindnera jadinii</i> (6.5)	(minor renovation, no sampling)	<i>Aureobasidium pullulans</i> (20.3) <i>Cryptococcus</i> (13.5) <i>Candida</i> (9.2) <i>Rhodotorula</i> (4.7) <i>Saccharomyces cerevisiae</i> (4.2)
Nursery 2	<i>Saccharomyces cerevisiae</i> (45.4) <i>Cladosporium</i> (7.2) <i>Cryptococcus</i> (4.3) <i>Aureobasidium pullulans</i> (3.7) <i>Rhodotorula</i> (3.1)	(major renovation, no sampling)	(major renovation, no sampling)	<i>Cryptococcus</i> (16.6) <i>Aureobasidium pullulans</i> (13.8) <i>Chaetomium</i> (12.9) <i>Cladosporium</i> (6.7) <i>Leptosphaeria</i> (6.5)
Farmhouse	(no sampling)	<i>Cyberlindnera jadinii</i> (38.7) <i>Candida</i> (12.9) <i>Cladosporium</i> (6.5) <i>Malassezia</i> (5.7) <i>Aureobasidium pullulans</i> (5.1)	<i>Cyberlindnera jadinii</i> (20.9) <i>Candida</i> (15.9) <i>Aspergillus</i> (15.5) <i>Penicillium</i> (9.1) <i>Cryptococcus</i> (7.4)	(no sampling)