이학석사 학위논문

Metagenomic analysis of clinical fungi associated with urban pigeon feces: implications for public health

메타지놈 분석을 통한 도심 비둘기 분변 내 병원성 진균에 대한 연구: 보건학적 함의를 중심으로

2015년 8월

서울대학교 대학원

생명과학부

이 원 동

Metagenomic analysis of clinical fungi associated with urban pigeon feces: implications for public health

Won Dong Lee

Advisor: Professor Young Woon Lim, Ph.D.

A thesis submitted in partial satisfaction of the requirements for the degree Master of Science in Biological Sciences

August 2015

Graduate School of Biological Sciences

Seoul National University

Metagenomic analysis of clinical fungi associated with urban pigeon feces: implications for public health

Won Dong lee

Graduate School of Biological Sciences

Seoul National University

Abstract

Human infectious fungi are a growing health concern worldwide, with cities posing a higher risk of infection. The dramatic upsurge of pigeon populations in cities has been implicated in the increased incidence of human fungal infection. In the current study, I used a culture-independent, high-throughput sequencing approach to elucidate the diversity of clinical fungi associated with pigeon feces. I mapped the absolute abundance of clinical fungi across Seoul, Korea, using quantitative PCR. In addition, I tested whether certain geographical, sociological, and meteorological factors were significant predictors of either the diversity or the absolute abundance of clinical fungi, or the presence/absence of specific clinical fungi species. Finally, I compared clinical fungi from fresh and old pigeon feces to elucidate the source of the

fungi and the role of pigeons in their dispersal; I inferred fungi in fresh feces to have passed through the pigeon gastrointestinal tract while other species present in old feces colonized after excretion. Our results demonstrated that both the composition and absolute abundance of clinical fungi are unevenly distributed throughout Seoul. The green area ratio and the number of multiplex houses were positively correlated with species diversity, whereas wind speed was negatively correlated. Three significant predictors (distance to city center, humidity, and wind speed) were negatively correlated with the absolute abundance of clinical fungi. Because many clinical fungi were absent in fresh feces, I concluded that most species cannot survive the gastrointestinal tract of pigeons; instead, many clinical fungi are transmitted through soil or air and use pigeon feces as a substrate for proliferation.

Keywords: human infectious fungi • *Columba livia* • next-generation sequencing (NGS) • quantitative PCR (qPCR) • culture-independent • fungal ecology

Student Number: 2013-22961

Contents

List of Figures

List of Tables

1. Introduction

1.1. Emerging fungal threats

Clinical fungi that cause human disease (de Hoog *et al*., 2014) are a growing health concern. Diseases caused by clinical fungi (e.g., candidiasis, aspergillosis, and mucormycosis) often have mortality rates exceeding 30%, and major fungal diseases cause at least as many deaths as tuberculosis or malaria (Brown *et al*., 2012a). The incidence of fungal disease was previously low but is now rising due to the growing population of immunocompromised individuals (Underhill and Iliev, 2014) and the expanding geographic distribution of clinical fungi as a result of global warming (Garcia-Solache and Casadevall, 2010). Although recent genomic approaches have improved antifungal therapies (Reichard *et al*., 2014), the lack of efficient diagnostics (Pfaller and Diekema, 2010), the limited success of drugs in reducing mortality rates (Brown *et al*., 2012b), and the absence of approved antifungal vaccines (Nanjappa and Klein, 2014) make clinical fungi a serious public health hazard.

Cities pose higher risk of fungal infection not only because the

population at risk is higher but for the ecology of clinical fungi itself. In urban areas, land cover is dominated by pavement and buildings, and the natural niche of most fungi (plant and soil) is altered by human activity (Alberti, 2005), the urban heat island effect (Oke, 1995), and pollution (Pickett *et al*., 2001). As a result, fungi are found on substrates such as furniture, cloth, food waste, and pets (Be *et al*., 2014). Most clinical fungi are opportunistic pathogens with a dual life cycle, and an environmental and animal-associated niche (de Hoog *et al*., 2014). Due to limited suitable substrates and a higher chance of contacting susceptible human hosts, opportunistic fungi are more likely to become human pathogens in urban areas (Groll *et al*., 1996; Steenbergen and Casadevall, 2003).

1.2. Pigeons: reservoirs and carriers of clinical fungi?

Pigeon populations have been implicated in the increased incidence of human fungal infection in cities (Jerolmack, 2008). Feral pigeons were domesticated from the rock dove (*Columba livia*) during the Upper Pleistocene (Driscoll *et al*., 2009; Stringham *et al*., 2012; Tchernov, 1984), and populations have dramatically increased with urbanization. Along with fouling and causing decay of infrastructure (Gomez-Heras *et al*., 2004), pigeons are a human

health concern, particularly because *Cryptococcus neoformans* was found in pigeon feces (Emmons, 1955) and shown to directly cause human cryptococcosis (Littman, 1959). Findings from studies worldwide that isolated *Cr. neoformans* and other infectious fungi from urban pigeon feces confirmed the reputation of pigeons as vectors of infectious disease (Haag-Wackernagel and Moch, 2004).

1.3. Impediments of studying the diversity of clinical fungi

Culture-dependent approach

To date, approximately 48 clinical fungi species from 28 genera have been isolated from pigeons around the world (Chae *et al*., 2012; Haag-Wackernagel and Moch, 2004; Jang *et al*., 2011; Khosravi, 1997; Ramirez *et al*., 1976; Soltani *et al*., 2013; Tokarzewski *et al*., 2007). However, this number is likely to be an underestimate, because most of these reports were based on culture-dependent methods, and the majority of fungi are not readily cultured (Pace, 1997; Rappé and Giovannoni, 2003).

Difficulties in species-level identification

Clinical fungi have traditionally been identified based on their morphologies, microscopic features, and physiological profiles. However some clinical fungi have no unique characters, which makes them difficult to be distinguished to the species level using morphology or physiology alone.

Many of these problems can be overcome by using DNA barcoding (Hebert *et al.* 2004). For its high sequence variation, the internal transcribed spacer (ITS) region has been proposed as the primary fungal barcoding gene for the kingdom Fungi (Schoch *et al.* 2012). Recent metagenomic approach, a culture-independent method targeting ITS region and utilizing next-generation sequencing (NGS), has been reported to detect a more complete clinical fungi diversity (Findley *et al*., 2013).

However, there are impediments to highly accurate ITS identification of fungi at the species level: due to tying top taxonomic hits from similarity-based methods such as BLAST and MEGAN, similar species can have different names because of the nomenclature of fungal species (Taylor, 2011), and incorrectly classified fungal species are frequently present in many fungal ITS databases (Nilsson *et al*., 2006). Phylogeny-based method such as Statistical Assignment Program (SAP; an automated phylogeny-based taxon assignment method) provide higher resolution assignment yet require queries

400 bp or longer for the best performance which is not suitable for most high throughput outcome.

1.4. Prerequisites for risk factor identification and effective intervention

Quantifying the risk

NGS, although powerful in detecting species diversity, only provides data on relative abundance (Dannemiller *et al*., 2014a). The prevalence of clinical fungi influences clinical risk; thus, quantification of the absolute abundance is necessary (Downs *et al*., 2001; Rao *et al*., 1996). The combination of quantitative PCR (qPCR) and NGS provides reliable information with which clinical fungi abundance and the resulting clinical risk can be quantified (Dannemiller *et al*., 2014a).

Identifying factors related to clinical fungi

Identifying environmental factors promoting clinical fungi diversity and abundance is a critical step for the prevention and intervention of human mycoses. Fungal diversity in air and soil are known to be influenced by the season, weather, human activity, urbanization, and habitat type (Fierer *et al*.,

2008; Newbound *et al*., 2010; Park, 2013; Pei-Chih *et al*., 2000; Shelton *et al*., 2002; Summerbell *et al*., 1992). The prevalence of clinical fungi (particularly *Cr. neoformans*) has been linked to weather (humidity, temperature, and sunshine), vegetation, microbes, and the chemical properties of pigeon feces (pH, uric acid, and nitrogen) (Granados and Castañeda, 2006; Hubalek, 1975; Ishaq *et al*., 1968; Levitz, 1991; Ruiz *et al*., 1982; Walter and Yee, 1968).

1.5. Objective of this study

I conducted a cross-sectional study to evaluate the potential risk factors (i.e., clinical fungi) associated with pigeon feces across Seoul, Korea. Seoul is an ideal place for such a study, because it is a large, densely populated city where pigeon populations have dramatically increased with urbanization, and previous culture-based studies identified several clinical fungi in pigeon feces (Chae *et al*., 2012; Chee and Lee, 2005; Jang *et al*., 2011). For this study, I used a culture-independent, high-throughput sequencing method to study the diversity and geographic distribution of clinical fungi. I combined NGS and qPCR to evaluate the absolute abundance of clinical fungi and, therefore, public health risk. In addition, I identified environmental factors affecting clinical fungi diversity and abundance using linear modeling. Finally, I tested

whether urban pigeons are vectors of clinical fungi by comparing clinical fungi species found in fresh and old pigeon feces.

2. Materials and Methods

2.1. Study location and sample collection

I sampled pigeon feces from 38 sites across Seoul, Korea in January 2014. Two different types of samples were collected: old (completely dried, decolorized) and fresh (wet, colored). I collected old samples from each locality, whereas fresh samples were collected from 18 of these sites.

Samples were collected using sterilized forceps and placed in sterile 15-ml Falcon tubes (BD Bioscience, Canaan, CT, USA). To minimize soil contamination, all samples were collected from rooftops, window ledges, or sidewalks. Fresh feces were collected immediately after excretion. All samples were transported on ice and stored at -80°C until subsequent analysis.

2.2. Metadata acquisition

Metadata were collected for variables potentially influencing clinical fungi diversity and abundance, including geography (latitude, longitude, and altitude), weather (precipitation, wind speed, maximum/mean/minimum

temperature, and humidity), site conditions (pigeon population size, green area ratio, impermeability, and habitat type), and sociological variables (distance to city center, number of households, number of workers, number of apartments, number of multiplex houses, and number of restaurants). The geographic data were recorded while sampling using the WGS84 coordinate system with an eXplorist 210 GPS (Magellan, San Dimas, CA, USA). All weather data were collected from the Korea Meteorological Administration database from 29 meteorological stations across Seoul in December 2013 and January 2014 (http://sts.kma.go.kr). Pigeon population size was estimated by observing the number of pigeons present during sampling. Data on the green area ratio (e.g., park, lawn, or garden), impermeability (e.g., pavement or building), and habitat type (biotope) were provided by Seoul Metropolitan Government (http://urban.seoul.go.kr). Sociological data were provided by BIZGIS (http://www.biz-gis.com/XsDB).

2.3. DNA extraction

Samples were homogenized (via grinding or mixing), and total DNA was isolated from 0.20 g (dry weight) or 0.30 g (wet weight) of old and fresh samples, respectively. DNA extraction was performed with the PowerSoil®

DNA Isolation Kit (MO BIO Laboratory, Carlsbad, CA, USA) according to the manufacturer's protocol with two modifications: for old pigeon feces, an additional 200 µl of PowerSoil® DNA Kit Bead Solution (MO BIO Laboratory, Carlsbad, CA, USA) was added to each tube before vortexing (according to the recommendation of the manufacturer for treating extremely dry samples), and all samples (both fresh and old) were incubated at 65°C in a dry oven for 15 min before further processing.

2.4. PCR amplification

The internal transcribed spacer (ITS) region was amplified using the primers ITS1F and ITS4 (Gardes and Bruns, 1993; White *et al*., 1990) modified for pyrosequencing; an 'A' pyrosequencing adaptor with a 10-bp multiplex tag was added to ITS4, and a 'B' pyrosequencing adaptor was added to ITS1F (Table 1). Each PCR reaction contained 2 µl of template DNA, AccuPower® PCR PreMix (Bioneer, Daejeon, Korea) (containing 1 unit of Top DNA polymerase; 250 µM dNTPs; 10 mM Tris-HCL; 30 mM KCl; and 1.5 mM MgCl2), and 10 pmol of each primer in a final volume of 20 µl. PCR was performed on a C1000™ thermal cycler (Bio-Rad, CA, USA) under the following conditions: initial denature at 95°C for 5 min, followed by 30 cycles

of 95°C for 40 s, 55°C for 40 s, and 72°C for 1 min, and a final step at 72°C for 10 min. PCR products were electrophoresed through a 1% agarose gel with LoadingSTAR (Dyne Bio, Seoul, Korea). To minimize PCR bias, three PCR reactions were pooled for each sample. PCR products were purified using the Expin™ PCR Purification Kit (GeneAll Biotechnology Co., Ltd., Seoul, Korea), and DNA concentrations were quantified using a NanoDrop ND-1000 spectrophotometer (Thermo, DE, USA). Two PCR libraries were constructed with 28 samples pooled in equimolar concentrations and analyzed using an Agilent 2100 Bioanalyzer (Agilent Technologies, CA, USA). Quality control analyses were performed with a Caliper LabChip® GX using the DNA High Sensitivity Assay (PerkinElmer, MA, USA).

Table 1. Primers used for ITS rRNA gene pyrosequencing. The forward primer was composed of the 454 primer B sequence, the key sequence (TCAG), and the ITS1F sequence. The reverse primer was composed of the 454 primer A sequence, the key sequence, a multiplex identifier (MID), and the ITS4 sequence.

Primer	Sequence
PY-ITS1F	cctatcccctgtgtgccttggcagtctcagCTTGGTCATTTAGAGGAAGTAA
PY-ITS4-1	ccatctcatccctgcgtgtctccgactcagACGAGTGCGTTCCTCCGCTTATTGATATGC
PY-ITS4-2	ccatctcatccctgcgtgtctccgactcagACGCTCGACATCCTCCGCTTATTGATATGC
PY-ITS4-3	ccatctcatccctgcgtgtctccgactcagAGACGCACTCTCCTCCGCTTATTGATATGC
PY-ITS4-4	ccatctcatccctgcgtgtctccgactcagAGCACTGTAGTCCTCCGCTTATTGATATGC
PY-ITS4-5	ccatctcatccctgcgtgtctccgactcagATCAGACACGTCCTCCGCTTATTGATATGC
PY-ITS4-6	ccatctcatccctgcgtgtctccgactcagATATCGCGAGTCCTCCGCTTATTGATATGC
PY-ITS4-7	ccatctcatccctgcgtgtctccgactcagCGTGTCTCTATCCTCCGCTTATTGATATGC
PY-ITS4-8	ccatctcatccctgcgtgtctccgactcagCTCGCGTGTCTCCTCCGCTTATTGATATGC
PY-ITS4-9	ccatctcatccctgcgtgtctccgactcagTCTCTATGCGTCCTCCGCTTATTGATATGC
PY-ITS4-10	ccatctcatccctgcgtgtctccgactcagTGATACGTCTTCCTCCGCTTATTGATATGC
PY-ITS4-11	ccatctcatccctgcgtgtctccgactcagCATAGTAGTGTCCTCCGCTTATTGATATGC
PY-ITS4-12	ccatctcatccctgcgtgtctccgactcagCGAGAGATACTCCTCCGCTTATTGATATGC
PY-ITS4-13	ccatctcatccctgcgtgtctccgactcagATACGACGTATCCTCCGCTTATTGATATGC
PY-ITS4-14	ccatctcatccctgcgtgtctccgactcagTCACGTACTATCCTCCGCTTATTGATATGC
PY-ITS4-15	ccatctcatccctgcgtgtctccgactcagCGTCTAGTACTCCTCCGCTTATTGATATGC
PY-ITS4-16	ccatctcatccctgcgtgtctccgactcagTCTACGTAGCTCCTCCGCTTATTGATATGC
PY-ITS4-17	ccatctcatccctgcgtgtctccgactcagTGTACTACTCTCCTCCGCTTATTGATATGC
PY-ITS4-18	ccatctcatccctgcgtgtctccgactcagACGACTACAGTCCTCCGCTTATTGATATGC
PY-ITS4-19	ccatctcatccctgcgtgtctccgactcagCGTAGACTAGTCCTCCGCTTATTGATATGC
PY-ITS4-20	ccatctcatccctgcgtgtctccgactcagTACGAGTATGTCCTCCGCTTATTGATATGC
PY-ITS4-21	ccatctcatccctgcgtgtctccgactcagTACTCTCGTGTCCTCCGCTTATTGATATGC
PY-ITS4-22	ccatctcatccctgcgtgtctccgactcagTAGAGACGAGTCCTCCGCTTATTGATATGC
PY-ITS4-23	ccatctcatccctgcgtgtctccgactcagTCGTCGCTCGTCCTCCGCTTATTGATATGC
PY-ITS4-24	ccatctcatccctgcgtgtctccgactcagACATACGCGTTCCTCCGCTTATTGATATGC
PY-ITS4-25	ccatctcatccctgcgtgtctccgactcagACGCGAGTATTCCTCCGCTTATTGATATGC

Primer	Sequence
PY-ITS4-26	ccatctcatccctgcgtgtctccgactcagACTACTATGTTCCTCCGCTTATTGATATGC
PY-ITS4-27	ccatctcatccctgcgtgtctccgactcagACTGTACAGTTCCTCCGCTTATTGATATGC
PY-ITS4-28	ccatctcatccctgcgtgtctccgactcagAGACTATACTTCCTCCGCTTATTGATATGC

Lower case letters indicate the portion required for pyrosequencing. Under lined sequences are the MIDs to tag the sequences to the respective sampl e origins (Parameswaran *et al.*, 2007). The sequence specific part of the p rimer is given in upper case letters.

2.5. Pyrosequencing

Pyrosequencing reactions were performed with flow pattern B on ¼ plate of a 454 GS FLX+ System (Roche 454 Life Sciences, CT, USA) and sequenced in the reverse direction. All steps after library construction were performed at Macrogen (Seoul, Korea). Raw sequence data have been deposited in the Sequence Read Archives (Leinonen *et al*., 2011) of NCBI under the accession number SRR2045733.

2.6. Sequence processing

Sequence reads from the two libraries were trimmed, filtered, and merged with the following adjustments using QIIME v1.8.0 (Caporaso *et al*., 2010): minimum length of 200 bp, maximum length of 1000 bp, minimum quality score of 25, no ambiguous bases, no mismatches in the primer sequence, and maximum homopolymer length of 6. I applied these strict filtering requirements to minimize the effect of amplicon noise. Chimera detection and clustering were performed using USEARCH v5.2.236 (Edgar, 2010) implemented in QIIME. Operational taxonomic units (OTUs) were constructed with clustering algorithm ‘nearest’ at 97% sequence similarity,

and representative sequences were selected based on the sequence abundance. Singletons (OTUs represented by a single sequence) were removed from analyses.

2.7. Taxonomic assignment

Accurate identification is important in any molecular study but more so when studying clinical species, because the results have implications for public health. I used a conservative, multi-step identification method that only classifies an OTU as clinical if multiple analyses identify it as such (Figure 1). First, BLAST 2.2.22 (Altschul *et al*., 1990) was used to assign representative sequences to two curated taxonomic databases (EMERENCIA [Ryberg *et al*., 2009] and ISHAM ITS Database for Human and Animal Pathogenic Fungi [Irinyi *et al*., 2015]) at a 98% minimum identity level. The preliminary identification of OTUs was performed using FHiTINGS 1-2 (Dannemiller *et al*., 2014b). Next, neighbor-joining phylogenetic analyses (with 500 bootstrap replicates) were run with type strains and verified strains from the literature for each genus. Identifications were only considered if the representative sequence was monophyletic with sequences of a single species. The final identification of each OTU was based on the combined results of FHiTINGS

and the phylogenetic analyses. To be conservative, I only categorized OTUs as clinical when the species-level identification was strongly supported by both methods. Low-resolution identification often overestimates risk factors by lumping clinical and non-clinical species into a genus-level identification. I followed the fungal taxonomy in Index Fungorum (http://www.indexfungorum.org) and the classification of clinical fungi in the Atlas of Clinical Fungi v. 4.0 (de Hoog *et al.*, 2014).

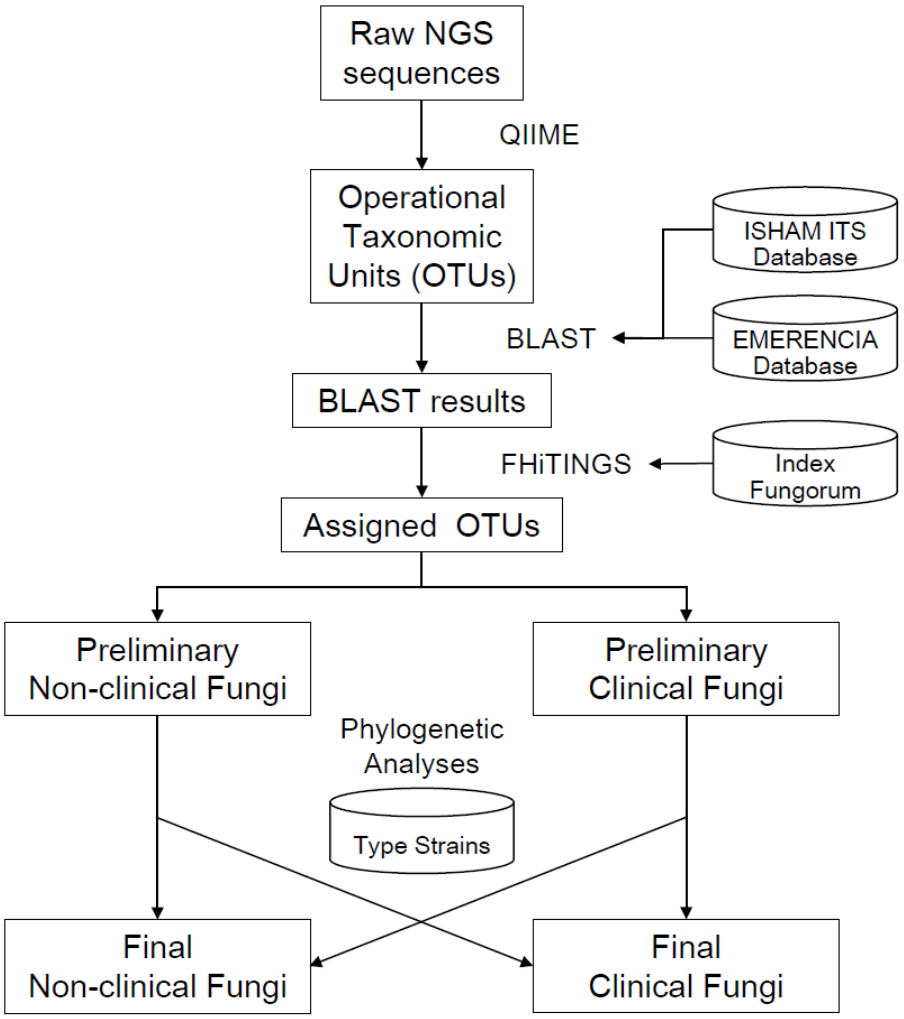


Figure 1. Schematic diagram of assigning and classifying OTUs as clinica l fungi.

2.8. Diversity analyses

I subsampled to 640 sequences per sample, which reduced the number of samples from 56 to 55. A sampling depth of 640 sequences struck a balance between retaining the greatest number of samples while eliminating the samples with low coverage. α-Diversity analyses were performed on the subsampled dataset: OTU diversity was estimated using Chao 1 (Chao, 1984; Chao, 1987), calculated in EstimateS (Colwell and Elsensohn, 2014) based on 100 randomizations, and Good's coverage (Good, 1953), calculated in QIIME.

2.9. Quantitative PCR

qPCR analysis for total fungi was conducted on the same DNA extracts used for pyrosequencing (Yamamoto *et al*., 2011). Three qPCR reactions were performed for each sample using FF2 and FR1 (Zhou *et al*., 2000) universal fungal primers, and the median value was chosen for further analysis. Standard curves were generated using DNA from Candida bovina KCTC 7253. The relative abundance of clinical taxa (from NGS results) was converted to absolute abundance by multiplying by total fungal spore equivalents (SE,

measured by qPCR) for each sample (Dannemiller *et al*., 2014a).

2.10. Statistical analyses

To test for spatial autocorrelation in clinical fungi species diversity and absolute abundance, I used Moran's I correlograms with 500 Monte Carlo simulations in SAM v. 4.0 (Rangel *et al*., 2010).

I preprocessed metadata before linear regression modeling. As most pigeons in cities have a foraging area within 0.3 km of their home lofts (Rose *et al*., 2006), all metadata were averaged to one value for a circle with a radius of 0.3 km using the Spatial Analyst tool in ArcGIS 10.2.2 (ArcGIS Desktop, ESRI, CA, USA). Highly correlated variables (> 0.7) were combined using linear regression to avoid overfitting and bias (Tabachnick and Fidell, 2013; Figure 2), and the remaining variables were standardized (Marquardt, 1980) using the caret package (Kuhn, 2008) in R v. 3.1.0 (R Development Core Team, 2014)

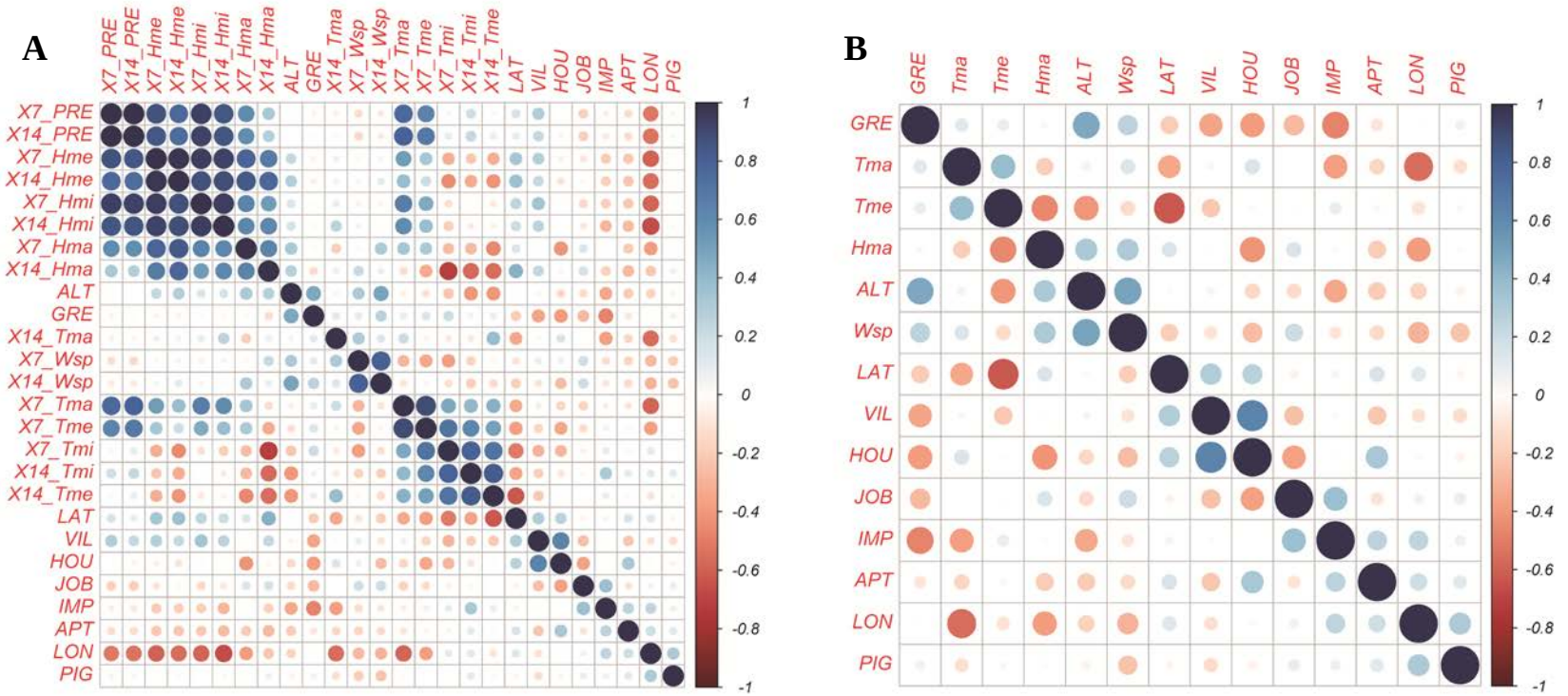


Figure 2. Correlations between environmental variables. (A) Pearson correlation heatmap of 26 variables without linear dependencies. (B) Pearson correlation heatmap of 14 remaining variables after filtering highly correlated (>.7) variables.

The multiple linear regression model was built with a permutation test to identify the significant predictors of the two dependent variables: clinical fungi diversity and absolute abundance of clinical fungi (natural log transformed). The best model was selected according to the corrected Akaike information criterion. All multivariate analyses were run in R. The coefficient for each of the significant variables in the linear model are reported as the mean ± 95% confidence interval. Variables were considered significant if $p < 0.05$.

2.11. Species distribution modeling of the commonly found clinical fungi species

The matrix with presence data was analyzed with the program Maxent version 3.3.3e (http://www.cs.princeton.edu/~schapire/maxent/). Niche models for the species present in at least 10 localities (13 species) were calculated with environmental variables with a 30 arc-seconds resolution. To prevent over-fitting, the variables that were considered to contribute less to the model were removed after observing the estimates of their relative contributions, and the jacknife tests of variable importance implemented in the Maxent software. Finally, the models were run with selected variables only, and with 10 fold

cross validation. Models were evaluated based on receiver operating characteristic analysis that generates the AUC (area under the curve) score.

3. Results

A total of 302,629 raw reads were obtained from the 56 samples; after filtering, 209,780 reads remained for community analysis. This corresponded to 303-7313 fungal ITS sequences per sample, with an average sequence length of 620 bp (ranging from 200 bp to 866 bp). Chao 1 estimators and Good's coverage confirmed that our sampling and sequencing was depth sufficient; Chao 1 estimators approached asymptotes (Figure 3), and Good's coverage indicated that our sequencing detected between 94.4% and 99.8% of the total diversity (Table 2).

Sequence clustering produced 431 fungal OTUs across all samples. These OTUs were assigned to 63 fungal genera where 55 (87%) were potentially clinical (Figure 4). Using our multi-step identification method, 42 OTUs were assigned to 35 clinical fungi species.

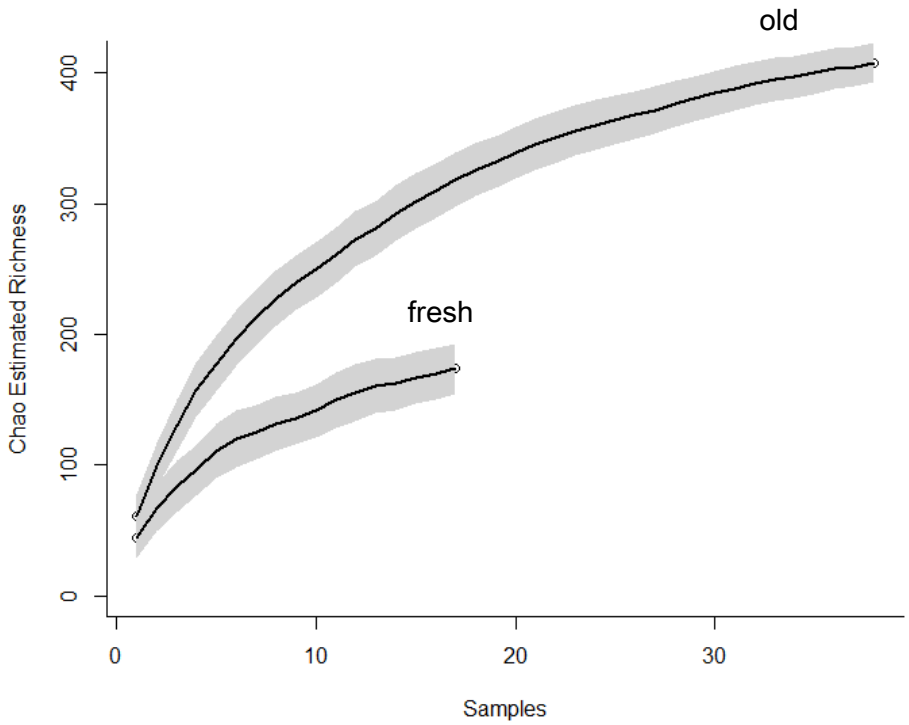


Figure 3. Collector's curves of estimated Chao 1 diversity of fungi in old (n = 38) and fresh (n = 18) pigeon feces. Chao 1 diversity was calculated from the subsampled dataset: 640 sequences per sample.

Table 2. Read number and Good's coverage value for each old and fresh pigeon feces sample calculated from fungal ITS gene pyrosequencing. Good's coverage values were calculated from the subsampled dataset: 640 sequences per sample (sample PD19F excluded).

Sample ID	Old Samples		Sample ID	Fresh Samples	
	Read	Good's coverage		Read	Good's coverage
PD01	3426	0.984			
PD02	2501	0.952			
PD03	4665	0.986			
PD04	3518	0.970			
PD05	2710	0.978	PD05F	879	0.984
PD06	5078	0.983			
PD07	6412	0.980	PD07F	7313	0.978
PD08	4782	0.944			
PD09	1485	0.972	PD09F	3346	0.994
PD10	4837	0.983			
PD11	5317	0.988			
PD12	5858	0.975	PD12F	5801	0.984
PD13	4821	0.992	PD13F	2250	0.975
PD14	4132	0.948	PD14F	4677	0.991
PD15	686	0.964	PD15F	3549	0.984
PD16	3694	0.980			
PD17	3261	0.997	PD17F	7115	0.992
PD18	5203	0.973	PD18F	2561	0.953
PD19	5032	0.992	PD19F	303	
PD20	3314	0.986	PD20F	4102	0.989
PD21	4602	0.997	PD21F	5798	0.984
PD22	6091	0.980			
PD23	3342	0.983			
PD24	2703	0.966			
PD25	3160	0.970			
PD26	3144	0.991			
PD27	3792	0.984			
PD28	2505	0.992			
PD29	4258	0.977	PD29F	663	0.981
PD30	684	0.961	PD30F	1588	0.967
PD31	4129	0.984	PD31F	1445	0.989

PD32	2538	0.995	PD32F	6458	0.981
PD33	3716	0.973	PD33F	2137	0.975
PD34	4619	0.998	PD34F	4813	0.984
PD35	5363	0.958			
PD36	3503	0.978			
PD37	4685	0.967			
PD38	1416	0.972			

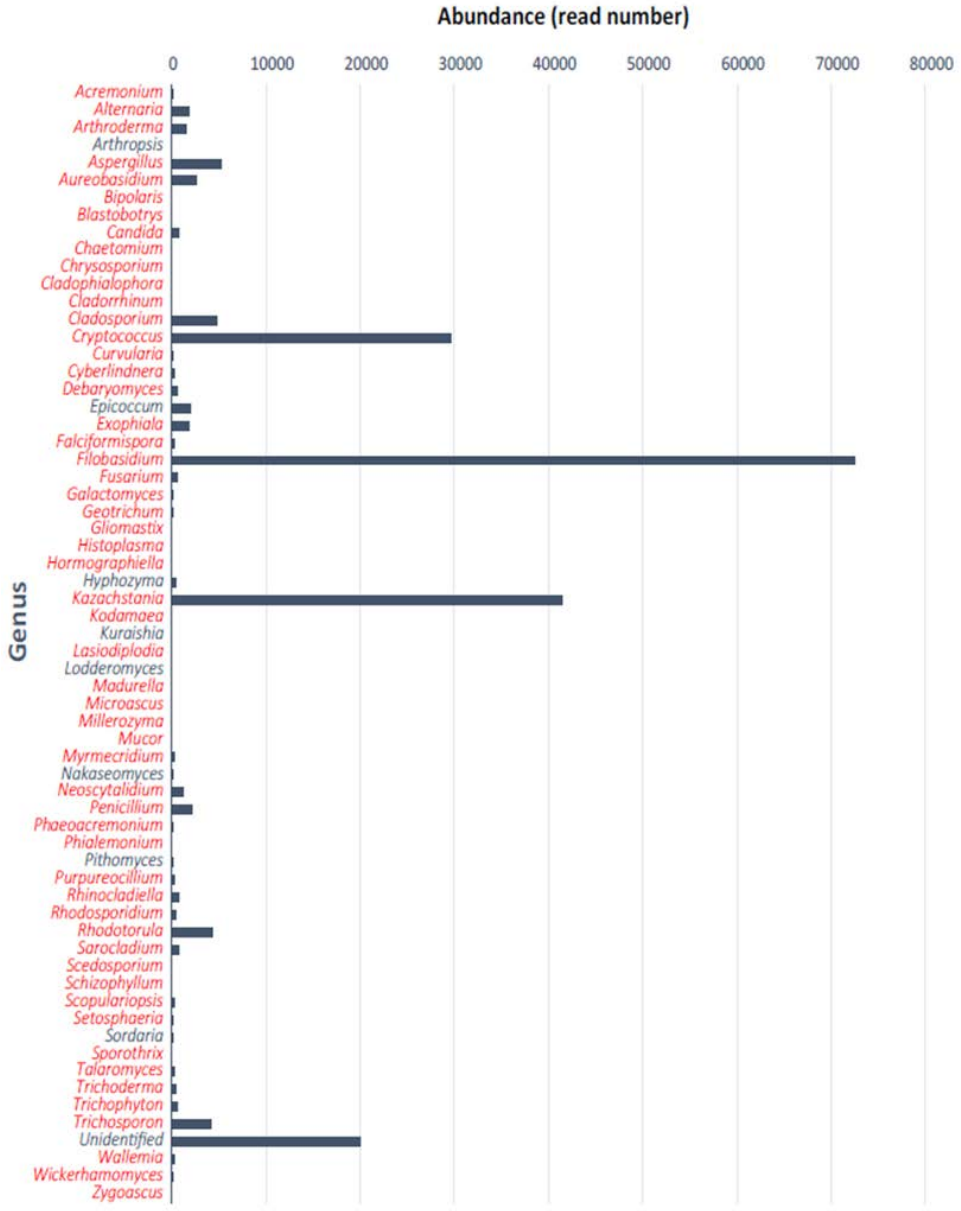


Figure 4. The relative abundance of fungal genera assigned with BLAST and FHiTINGS. Red names indicate fungal genera that include known clin ical species based on the Atlas of Clinical Fungi (de Hoog *et al.*, 2014).

3.1. The diversity and distribution of clinical fungi

I did not find spatial autocorrelation among properties of sampling sites (data not shown); thus, this was excluded from further analyses. The species diversity of clinical fungi associated with pigeon feces was heterogeneous between sites (2–18 species per site; Figure 5). Candida bovina was the most widely distributed clinical fungus species found in old pigeon feces (32/38 sites). Alternaria alternata (31 sites), Cr. magnus (28 sites), and Rhodotorula mucilaginosa (26 sites) were also widely distributed (Table 3).

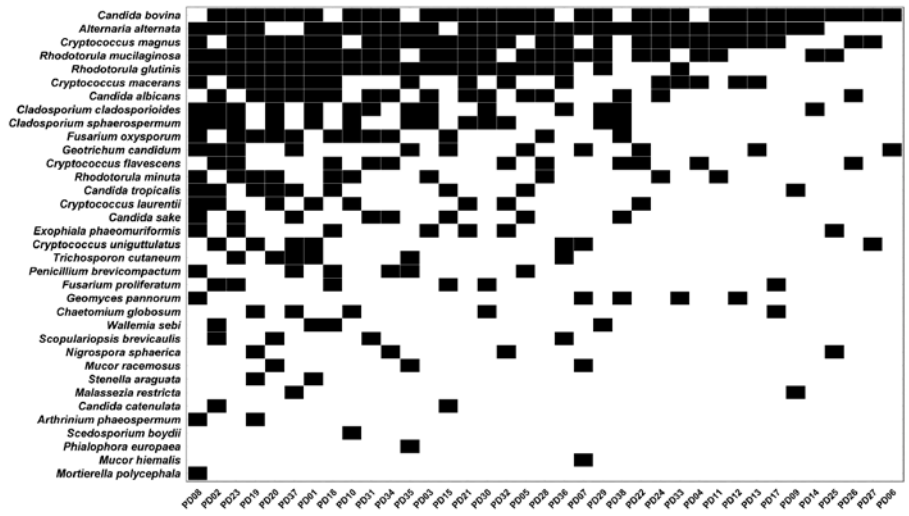


Figure 5. The distribution of clinical fungi across sites. Cells in black represent the presence of a species at a site. Rows toward the top represent species with wider distribution, while columns toward the left represent sites with higher diversity. Site locations refer to Figure 6.

Table 3. List of clinical fungi species identified from pigeon feces and their frequency in fresh and old feces. For each operational taxonomic unit, I list the information (accession and strain number) of the closest related species (type strain when available) from GenBank. Species in bold font are new global records associated with pigeon feces.

Clinical fungal species	Accession no.	Strain no.†	Fresh	Old*	BSL††
Alternaria alternata	KF465761	CBS 916.96 T	8	16 (31)	1
Arthrinium phaeospermum	KF144908	CBS 142.55 T	0	1 (2)	1
Candida albicans	JN944029	CBS 562 T	1	8 (16)	2
Candida bovina[a]	AJ223020	CBS 2760 T	14	17 (32)	1
Candida catenulata	GU246267	CBS 565 T	0	1 (2)	1
Candida sake	AJ549822	CBS 159 T	4	4 (8)	1
Candida tropicalis[a]	AY939810	ATCC 750 T	0	6 (9)	2
Chaetomium globosum	JN209920	CBS 164.62 T	0	3 (5)	1
Cladosporium cladosporioides	AJ300335	CBS 169.54 T	1	5 (14)	1
Cladosporium	DQ780343	CBS 193.54 T	1	5 (13)	1
Cryptococcus flavescens	AB035046	CBS 942 T	0	4 (11)	1
Cryptococcus laurentii	AF410468	CBS 139 T	0	3 (8)	1
Cryptococcus macerans	EU082230	CBS 6532 T	4	8 (16)	1
Cryptococcus magnus	AF190008	CBS 140 T	11	14 (28)	1
Cryptococcus uniguttulatus	AF335938	CBS 1730 T	0	2 (7)	1
Exophiala phaeomuriformis	EF025401	UTHSC88-471	0	3 (7)	2
Fusarium oxysporum	KM030313	NRRL 62960	0	6 (12)	2
Fusarium proliferatum	AB587006	CBS 216.76 T	0	4 (6)	1
Geomyces pannorum	AJ509868	S9A4	2	3 (5)	1
Geotrichum candidum[a]	AY788300	CBS 182.33 T	4	4 (11)	1
Malassezia restricta	AF522062	CBS 7877 T	1	1 (2)	1
Mortierella polycephala	HQ630335	CBS 456.66	2	0 (1)	1
Mucor hiemalis	EU484277	CBS 201.65 T	1	1 (1)	1
Mucor racemosus	FN650642	CBS 260.68 T	1	2 (3)	1
Niarospora sphaerica	FJ478134	xsd08093	0	3 (4)	1
Penicillium brevicompactum	AY484927	NRRL 868 T	0	3 (6)	1
Phialophora europaea	JQ766440	CBS 129.96	0	0 (1)	2
Rhodotorula glutinis	AF444539	CBS 20 T	2	12 (22)	1
Rhodotorula minuta	AB026016	JCM 3777 T	1	3 (10)	1
Rhodotorula mucilaginosa	AF444541	CBS 316 T	6	12 (26)	1

Scedosporium boydii	AY877350	CBS 101.22 T	0	0 (1)	2
Scopulariopsis brevicaulis	KP132734	FMR 12235	0	2 (4)	2
Stenella araguata	EU019250	ATCC 24788 T	0	1 (2)	1
Trichosporon cutaneum [a]	AF444437	CBS 6864 T	1	1 (6)	2
Wallemia sebi	AY302526	CBS 196.56	0	2 (4)	1

[a] Species with new records associated with pigeon feces in Korea but not worldwide

† Strains denoted with a "T" represent type strains

*Comparisons are made between the sites from which both fresh and old feces were collected (n = 18). Additionally, for the "Old" column, the number in parentheses represents the total number of sites in which the species was present (n = 38).

†† Biosafety Level

The best linear model selected explained 29.3% of the total variance and included the following factors: the green area ratio, wind speed, the number of multiplex houses, and habitat type (Table 4). While the green area ratio and the number of multiplex houses were positively correlated with clinical fungi diversity, wind speed was negatively correlated. An analysis of habitat type (categorical variable) revealed that the “commercial and business” and “industrial and urban infrastructure” biotopes displayed significantly higher clinical fungi diversity compared with the other six biotopes.

Table 4. Factors affecting the clinical fungi diversity among pigeon feces based on the highest scoring linear model: GRE + Wsp + MUL + HAB (Adjusted R^2 = 0.293)

Variable	Beta ± SE*	Lower 95% CI	Upper 95% CI	*P*-value
GRE	3.765 ± 1.427	0.968	6.562	0.012
Wsp	-2.576 ± 0.648	-3.846	-1.306	<0.001
MUL	2.311 ± 0.802	0.739	3.883	0.009
HAB				
Commercial and business biotope	3.927 ± 1.848	0.305	4.289	0.035
Industrial and urban infrastructure biotope	3.720 ± 1.794	0.204	4.072	0.038
Landscape biotope	-2.431 ± 2.828	-7.974	-1.877	0.306
Residential biotope	-0.328 ± 1.845	-3.944	0.034	0.922
Rivers and wetlands biotope	0.387 ± 3.252	-5.987	1.024	0.882
Transportation biotope	1.711 ± 3.822	-5.780	2.460	0.615
Woodland biotope	–	–	–	–

SE = standard error, CI = confidence interval,

GRE = green area ratio, Wsp = wind speed, MUL = number of multiplex houses, HAB = habitat type (categorical)

* coefficient and standard error for each variable

3.2. The absolute abundance of clinical fungi

The absolute abundance of clinical fungi among pigeon feces differed greatly across Seoul (ranging from 4 × 104 SE/ml to 9 × 108 SE/ml; Figure 6). The relative abundance (Figure 7) and the total fungal spore equivalents (Figure 8) of clinical fungi varied from site to site. Sites closer to the city center tended to have a higher abundance of clinical fungi compared with peripheral parts of the city. The absolute abundance of clinical fungi and the number of clinical species identified in each sampling site were not related to each other.

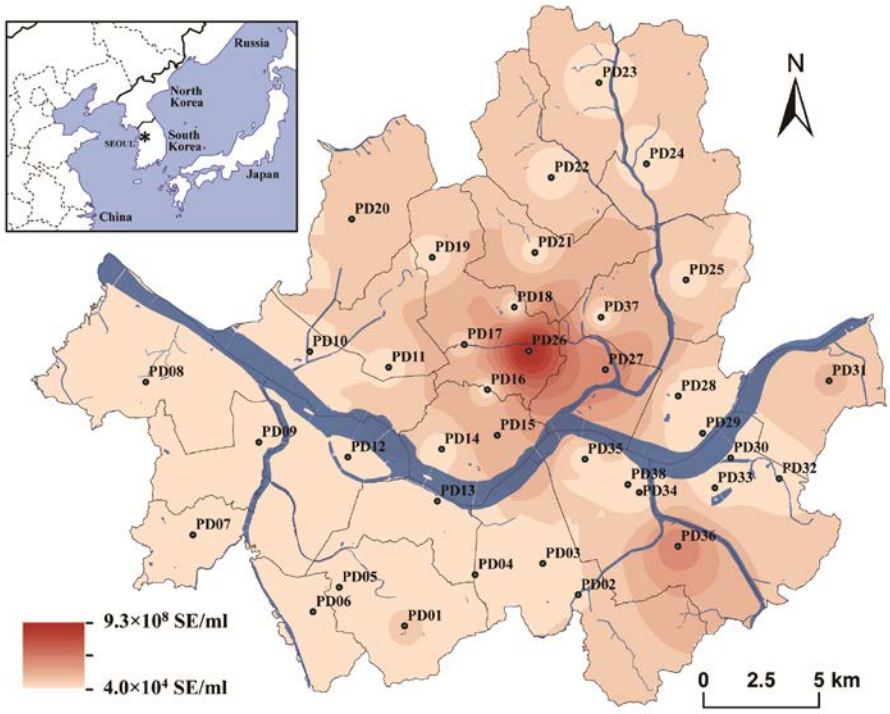


Figure 6. Heatmap of the abundance of clinical fungi (spore equivalent [SE] per ml) associated with pigeon feces across Seoul, Korea. The heatmap was produced using an inverse distance interpolation method in ArcGIS.

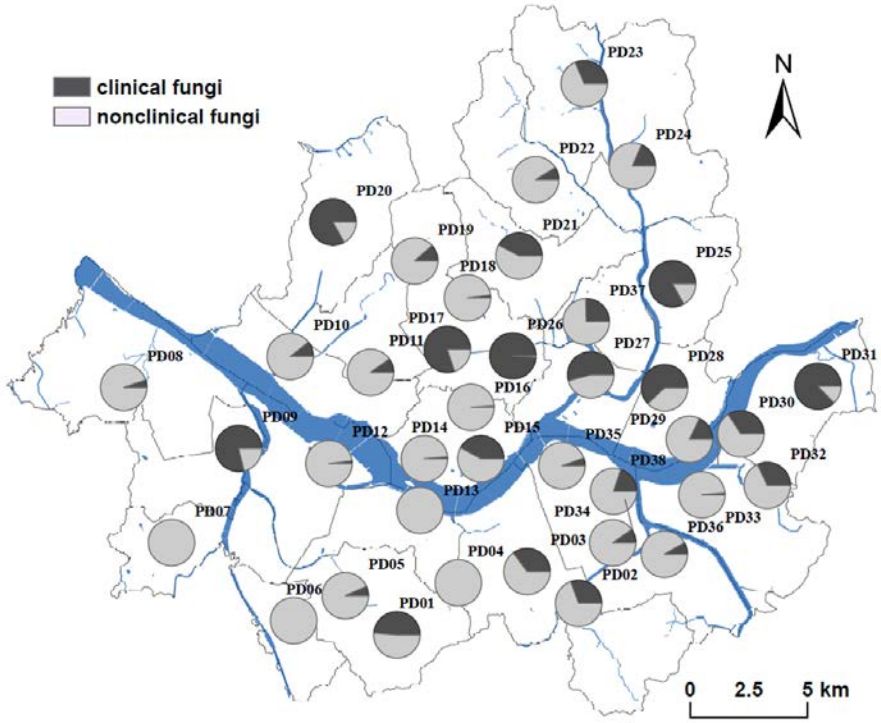


Figure 7. The distribution pattern of the clinical fungi associated with pigeon feces across Seoul, Korea. Pie charts represent the proportion of sequences assigned clinical to nonclinical fungi.

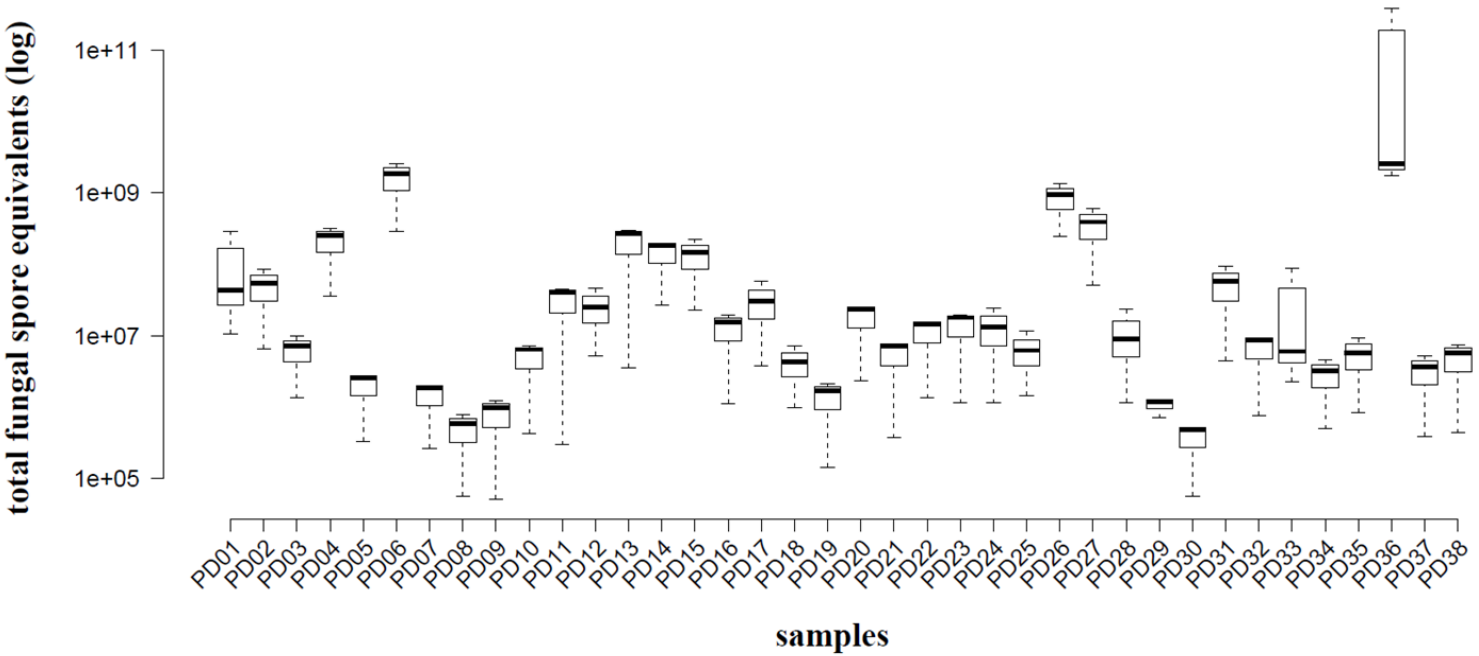


Figure 8. The absolute abundance of fungal biomass in each old pigeon feces sample. The total fungal spore equivalent was measured three times for each sample.

Four variables (humidity, distance to city center, wind speed, and habitat type) were statistically significant predictors of the absolute abundance of clinical fungi (natural log transformed), explaining 21.6% of the total variance (Table 5). All three continuous variables (humidity, distance to city center, and wind speed) were negatively correlated with the absolute abundance of clinical fungi. For habitat type (categorical variable) the “woodland” biotope displayed a significantly higher absolute abundance of clinical fungi compared with the six other biotopes.

Table 5. Factors affecting the absolute abundance of clinical fungi among pigeon feces based on the highest scoring linear model: CEN + HUM + Wsp + HAB (Adjusted R^2=0.216)

Variable	Beta ± SE*	Lower 95% CI	Upper 95% CI	*P*-value
CEN	-1.291 ± 0.465	-2.202	-0.380	0.011
HUM	-1.184 ± 0.447	-2.060	-0.308	0.015
Wsp	-0.837 ± 0.410	-1.641	-0.033	0.035
HAB				
Commercial and business biotope	-1.131 ± 0.767	-2.634	0.829	0.028
Industrial and urban infrastructure biotope	1.562 ± 0.993	-0.384	3.522	0.154
Landscape biotope	1.311 ± 1.276	-1.190	3.271	0.667
Residential biotope	1.896 ± 0.913	0.107	3.856	0.058
Rivers and wetlands biotope	-2.640 ± 2.068	-6.693	-0.680	0.183
Transportation biotope	-0.600 ± 2.063	-4.643	1.360	1.000
Woodland biotope	–	–	–	–

SE = standard error, CI = confidence interval,

CEN = distance to city center, HUM = humidity, Wsp = wind speed, HAB = habitat type (categorical)

* coefficient and standard error for each variable

3.3. Niche models for the common clinical fungi species

Variables selected for each species model and their relative contributions are shown in Table 6. All models had high AUC scores (>0.9) for training data. AUC scores for test data varied ranging from 0.595 to 0.965. Environmental variables that most frequently had a high percentage contribution to the models were biotope type (HAB), number of apartments (APT), and maximum temperature (Tma). According to Maxent predictive ecological models, *Rhodotorula minute*, *Fusarium oxysporum*, *Geotrichum candidum* are the most ubiquitous fungal species in Seoul. While other fungal species have their own prevalence, *Cladosporium* species seem to share similar ecological niche (Figure 9).

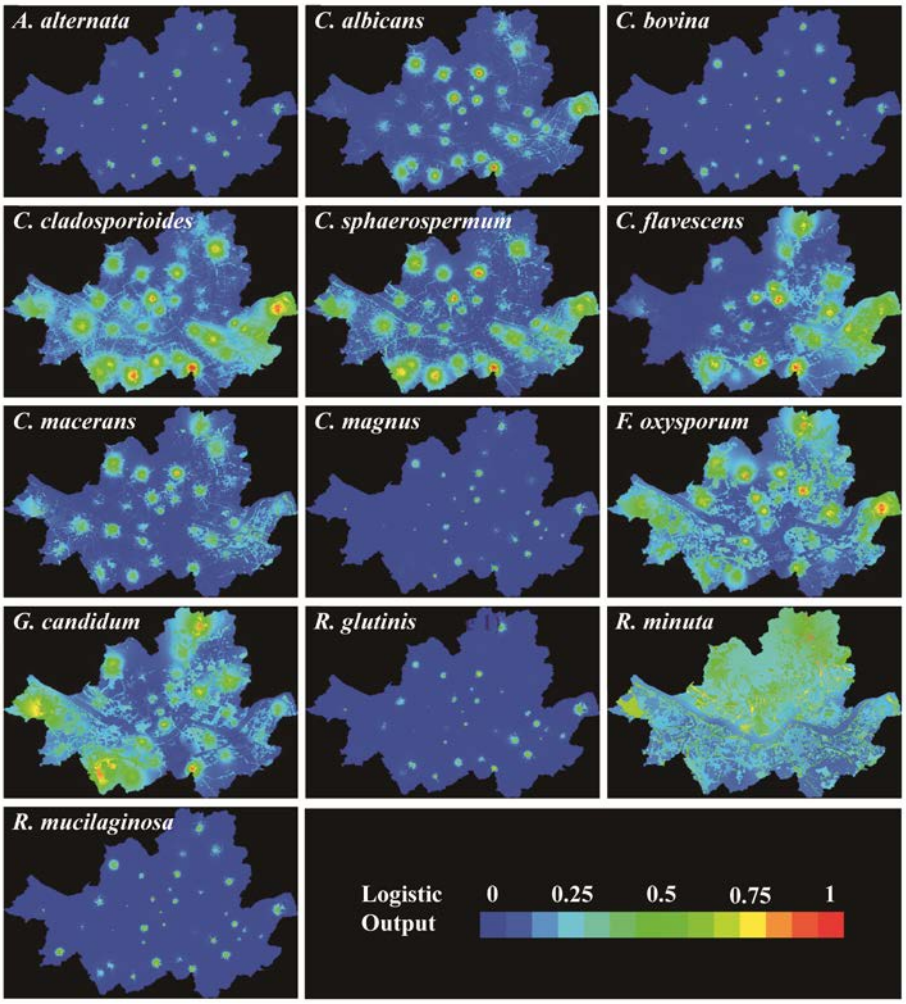
Table 6. Results of the Maxent environmental niche models.

Species	Loc.	Variables															AUC	AUC
		PIG	LAT	LON	ALT	HAB	GRE	IMP	HOU	APT	MUL	JOB	Tma	TEM	HUM	Wsp	train	test
Alternaria alternata	31 (28, 3)	1	1	4	3	8	6	3	0	18	31	7	12	0	4	1	0.996	0.965
Candida albicans	16 (14, 2)	6	0	4	3	11	13	1	6	5	2	12	24	0	8	4	0.981	0.869
Candida bovina	32 (29, 3)	4	1	3	3	7	11	0	5	32	13	4	4	2	7	3	0.994	0.924
Cladosporium cladosporioides	14 (13, 1)	2	0	0	1	17	2	19	6	5	2	1	44	0	1	0	0.940	0.616
Cladosporium sphaerospermum	13 (12, 1)	2	0	0	1	27	0	12	11	3	10	2	29	1	2	0	0.945	0.691
Cryptococcus flavescens	11 (10, 1)	0	1	4	9	18	12	0	0	8	1	5	23	3	15		0.985	0.904
Cryptococcus macerans	16 (14, 2)	1	2	1	3	21	7	4	3	9	13	3	8	7	14	4	0.977	0.694
Cryptococcus magnus	28 (25, 3)	1	0	2	5	7	6	3	0	47	7	4	8		7	2	0.987	0.928
Fusarium oxysporum	12 (11, 1)	2	2	0	3	28	3	0	0	1	0	16	4	0	10	32	0.958	0.650
Geotrichum candidum	11 (10, 1)	19	4	0	5	27	10	0	6	7	1	19	1	0	0	2	0.937	0.669
Rhodotorula glutinis	22 (20, 2)	8	1	3	3	5	4	6	0	22	13	8	11	0	12	4	0.997	0.908
Rhodotorula minuta	10 (9, 1)	0	12	0		80		0			0	1	0	7	0	0	0.907	0.595
Rhodotorula mucilaginosa	26 (23, 3)	3	0	2	5	8	13	5	1	35	4	2	12	0	5	4	0.996	0.912

Abbreviations: AUC test, area under the ROC curve for test data; AUC training, area under the ROC curve for training data; ALT, altitude; APT, number of apartments; GRE, green area ratio; HAB, habitat type; HOU, number of household; HUM, humidity; IMP, impermeable area; LAT, latitude; Loc., total number of localities (training, test); LON, longitude; MUL, number of multiplexhouses; PIG, pigeon population estimated; PRE, precipitation; TEM, temperature; Tma, max temperature; Wsp, wind speed.

The table gives the estimated percentage of relative contribution of selected variables for each species' final model, and the area under the receiver operating characteristic (ROC) curve for training and test data. The variables with higher contributions then 20 percent for each species are highlighted in grey.

Figure 9. Maxent predictive ecological models.



3.4. Comparison between fresh and old feces

The proportion and number of clinical fungi species were significantly higher in old (30%, 33 species) compared with fresh feces (8%, 18 species; Figure 10). All clinical fungi species detected in fresh samples, with the exception of *Mortierella polycephala*, were also present in old samples.

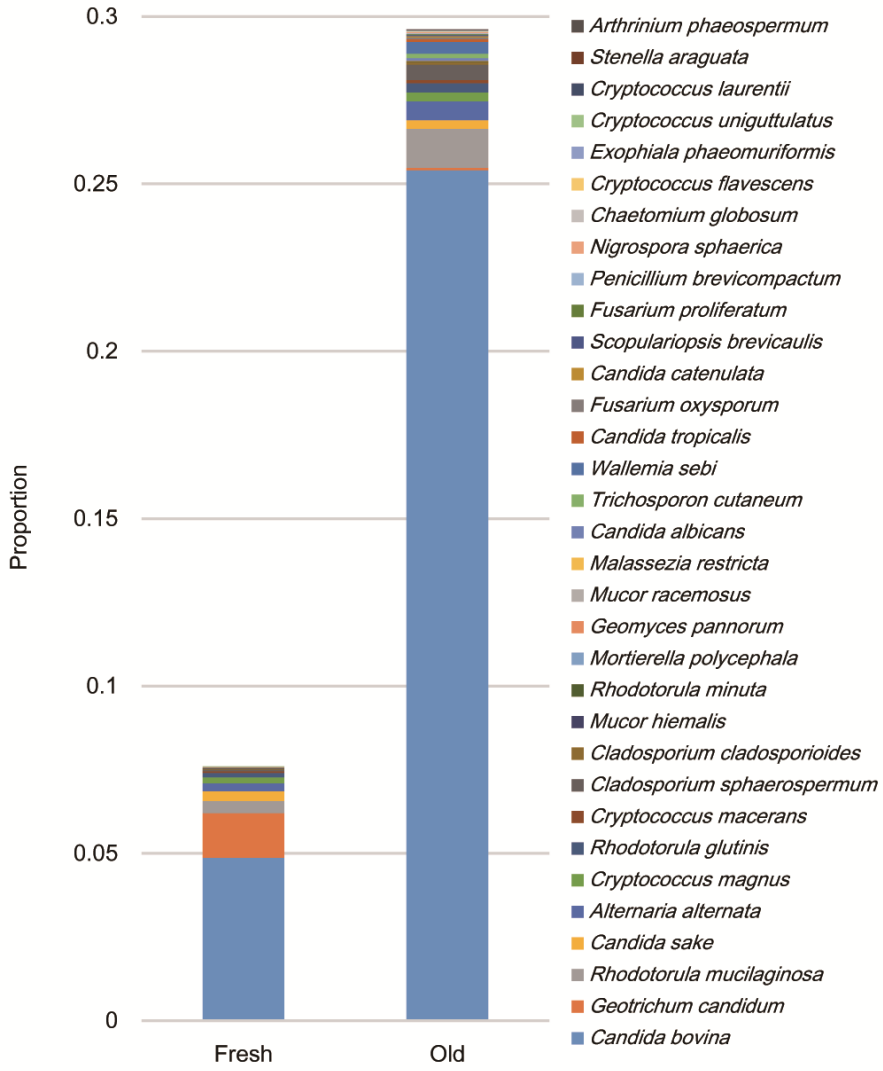


Figure 10. A comparison of clinical fungi diversity relative to total fungal diversity between fresh and old pigeon feces.

4. Discussion

I combined NGS with qPCR and a rigorous species identification method to elucidate the diversity and distribution of clinical fungi in pigeon feces. First, I identified 35 clinical fungi species, 24 of which are new records associated with pigeon feces (Table 3). Statistically significant predictors of clinical fungi diversity were the green area ratio, wind speed, the number of multiplex houses, and habitat type. In addition, I observed that the absolute abundance of clinical fungi was higher toward the city center, and identified four variables (distance from city center, wind speed, humidity, and habitat type) to be statistically significant predictors of the geographic patterns of abundance. Lastly, because a majority of the clinical fungi were found in old and not fresh feces, I inferred that the source of these fungi was environmental, rather than the pigeon itself.

4.1. The diversity and distribution of clinical fungi

I observed high fungal diversity among pigeon feces, with 431 OTUs assigned to 63 fungal genera (Table 2). In comparison to studies performed worldwide,

this accounted for over 50% of the previously reported clinical genera. Notably, I identified nine clinical fungi genera that were previously not known from pigeon feces. Of the 42 OTUs representing 35 clinical fungi species identified in our study, 11 species were found in previous studies whereas 24 were new records associated with pigeon feces (Table 3). In comparison with studies of pigeon feces performed in Seoul using culture-dependent methods (Chae *et al*., 2012; Jang *et al*., 2011), 15 species were unique to previous studies, 28 were unique to our study, and 7 were identified in both previous studies as well as in our study. In addition, 16 genera were newly identified to be associated with pigeon feces in Seoul. Our linear model identified four statistically significant predictors (green area ratio, wind speed, number of multiplex houses, and habitat type) of clinical fungi species diversity (29.3% of variance explained; Table 4).

The green area ratio and the number of multiplex houses were positively correlated with clinical fungi diversity. Based on the coefficient, the green area ratio had the largest effect on the diversity of clinical fungi. This implies that the diversity of clinical fungi isolated from pigeon feces is influenced by the local environment. Residential areas dominated by multiplex houses may influence species diversity because of historical factors (these areas are often in older, less developed parts of the city) (Jun 2011) or

environmental factors (densely packed houses may reduce exposure to sunshine) (Kim *et al*., 2006).Wind speed was negatively correlated with species diversity; high wind speed has been demonstrated to decrease the abundance (Guinea *et al*., 2006; Hasnain, 1993) and influence species composition (Yamamoto *et al*., 2012) of aerial fungi, reducing the probability and species diversity of clinical species settling on pigeon feces. Lastly, an analysis of habitat type (a categorical variable) revealed that more urbanized areas ("commercial and business" and "industrial and urban infrastructure" biotopes) had a significantly higher diversity of clinical fungi compared with that in other areas ("residential" and "woodland" biotopes) in Seoul.

4.2. The absolute abundance of clinical fungi

The health risk from clinical fungi is dosage-dependent; thus, quantification of the absolute abundance is necessary (Downs *et al*., 2001; Rao *et al*., 1996). I used qPCR to measure the absolute abundance of clinical fungi across Seoul and found the range to be between 4 × 104 SE/ml and 9 × 108 SE/ml. These values are at least 105 times higher than typical airborne fungi concentrations (Lee *et al*., 2010; Yamamoto *et al*., 2015). The absolute abundance of clinical fungi was unevenly distributed throughout Seoul (Figure 6), with the highest

abundance found near the city center.

The linear model identified four statistically significant predictors explaining clinical fungi density: distance to city center, wind speed, humidity, and habitat type (29% of the variance explained; Table 5). The three continuous variables (distance to city center, wind speed, and humidity) were negatively correlated with absolute abundance.

The variable with the highest coefficient was the distance to the city center. A similar pattern was reported with pigeon populations in Europe (Hetmanski and Jarosiewicz, 2008; Sacchi *et al*., 2002), with Przybylska *et al*. (2012) suggesting that a higher density of old, tall buildings in the city center promotes a higher pigeon density. The density of clinical fungi may decrease with increasing wind speeds, because high wind speed can decrease the level of airborne fungi and, thus, decrease the number of spores settling on feces (Guinea *et al*., 2006; Hasnain, 1993). The negative correlation of clinical fungi and humidity is likely due to a seasonal effect; in the winter, when sampling was conducted, many fungi are in spore form, which is vulnerable to high humidity (Ishaq *et al*., 1968). Habitat type had the opposite effect to what was shown in the previous linear model analysis, where high levels of fungi were found in the “residential” biotope and low levels were found in the “commercial and business” biotope.

4.3. Niche models for the common clinical fungi species

Prevention strategies for human mycoses could target at reducing the number of clinical fungi species or its absolute abundance. For example, environmental factors I identified earlier, which were significantly related to the number of clinical fungi species or its absolute abundance, are potential candidates for such intervention studies. However, if I are to target and control incidences caused by specific clinical fungi species, I need to identify factors (or niches) closely related to that certain species.

Maximum Entropy Modeling revealed variables selected for each species model and their relative contributions (Table 6). Biotope type (HAB), number of apartments (APT), and maximum temperature (Tma) were environmental variables that most frequently had a high percentage contribution to the models. According to our result, these factors could be primary targets for controlling the most common clinical fungi species.

Meanwhile, Maxent predictive ecological models suggest potential distributions of each clinical species across Seoul (Figure 9). Based on the presence/absence map, government agencies may develop effective strategies to improve public health and prevent human fungal infections caused by certain fungal agents.

4.4. Comparison between fresh and old feces

Fungal communities in pigeon feces are dependent on a combination of what pigeons eat and what is present in the environment (Kwon-Chung, 1976; Partridge and Winner, 1965). By comparing the presence/absence of species in fresh and old feces, I inferred the source of clinical fungi (Table 3). I interpreted species present on at least two fresh feces samples as true presence (one sample could be due to contamination) and an indication that it originated from the pigeon (regardless of its presence/absence on old feces). These species originate either from the pigeon itself or from ingested food after passing through the pigeon gastrointestinal (GI) tract. The identification of species on old and not fresh feces (or fewer than 2 samples) implies that these species originated from the environment (e.g., wind or soil) and colonized the feces after excretion.

Based on our criteria, I observed that the majority of clinical fungi species originated from the environment (60%) as opposed to originating from the pigeon (29%; 11% ambiguous; Table 3). The community of clinical species from old pigeon feces in our study was similar to that observed in studies of airborne fungi (59% of genera shared) (Fang *et al*., 2005; Oh *et al*.,

2014; Shelton *et al*., 2002; Solomon, 1976; Yamamoto *et al*., 2012), supporting the hypothesis that these fungi came from the air. Based on our results, I suggest that airborne fungi, rather than those originating from the pigeons themselves, are the primary source of clinical fungi in cities.

Although routine isolation of *Cr. neoformans* from pigeon feces worldwide gave pigeons a reputation for being a vector of disease (Haag-Wackernagel and Moch, 2004), I did not detect this species in our study. Interestingly, no previous studies have provided evidence that *Cr. neoformans* can pass through the GI tract of pigeons in nature. Several laboratory studies identified *Cr. neoformans* spores in fresh feces (Abou-Gabal and Atia, 1978; Sethi and Randhawa, 1968; Swinne-Desgain, 1974), but these studies were inconclusive because they used an unnaturally high level of spores that wild pigeons would never encounter in nature. Additionally, a recent survey of *Cr. neoformans* in Seoul had an extremely low recovery rate (4/306 samples, 1.3%; Jang *et al*., 2011).

The presence of four species (*Ca. bovina*, *Cr. magnus*, *Ca. sake*, and *Geotrichum candidum*) in fresh feces demonstrates their potential to survive in a pigeon's GI tract. Previous studies have suggested that three of these species (*Ca. bovina*, *Cr. magnus*, and *Ca. sake*) are disseminated by pigeons (Chae *et al*., 2012; Hasenclever and Kocan, 1975). *Candida bovina* displays

qualities of being able to survive harsh environments, as it is thermophilic, colonizes the GI tract of various warm-blooded animals, including humans (Kurtzman *et al*., 2005), and lives in the GI tract of swine (Uden, 1960). Interestingly, most clinical fungi found in fresh pigeon feces are not dangerous clinical fungi (biosafety level 1), as they tend to cause only mild symptoms and are easily cured by commercial treatments (de Hoog *et al*., 2014). The notoriously dangerous clinical species (e.g., *Ca. albicans* and *Ca. tropicalis*; biosafety level 2) were found in old feces (Table 3). The distinction is nuanced, but it appears that pigeons themselves are not dangerous; rather, their feces are a substrate for environmentally transported (e.g., wind) clinical fungi.

5. Conclusion

In this study, I used a culture-independent, high-throughput sequencing approach with qPCR to conduct a cross-sectional study to describe the distribution of clinical fungi among pigeon feces. I identified 35 clinical fungi species, 24 of which are newly associated with pigeon feces. I estimated clinical fungi diversity and abundance based on currently known clinical fungi; however, as many fungal species have not yet been tested for their pathogenicity, our result may be an underestimate. Human activity, geographic location, habitat type, wind speed, and humidity were significant variables that contributed to the diversity and abundance of clinical fungi. Our results suggest that aerial fungi, not pigeons, play a critical role in disseminating clinical fungi; rather, the role of pigeons appears to be to provide a substrate (feces) for clinical fungi to proliferate. I believe these results can help government agencies develop better strategies to improve public health and prevent human fungal infections.

6. References

Abou-Gabal M, Atia M. (1978). Study of the role of pigeons in the dissemination of *Cryptococcus neoformans* in nature. *Sabouraudia* **16:** 63—68.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. (1990). Basic local alignment search tool. *J Mol Biol* **215:** 403—410.

Brown GD, Denning DW, Gow NaR, Levitz SM, Netea MG, White TC. (2012a). Hidden killers: human fungal infections. *Sci Transl Med* **4:** 165rv113.

Brown GD, Denning DW, Levitz SM. (2012b). Tackling human fungal infections. *Science* **336:** 647—647.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK *et al.* (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7:** 335—336.

Chae HS, Jang GE, Kim NH, Son HR, Lee JH, Kim SH *et al.* (2012). Classification of *Cryptococcus neoformans* and yeast-like fungus isolates from pigeon droppings by colony phenotyping and its genotyping and their seasonal variations in Korea. *Avian Dis* **56:** 58—64.

Chao A. (1984). Nonparametric estimation of the number of classes in a population. *Scand J Stat* **11:** 265—270.

Chao A. (1987). Estimating the population size for capture-recapture data with unequal catchability. *Biometrics* **43:** 783—791.

Chee HY, Lee KB (2005). Isolation of *Cryptococcus neoformans* var. *grubii* (serotype A) from pigeon droppings in Seoul, Korea. *J Microbiol* **43:** 469—472.

Colwell RK, Elsensohn JE. (2014), EstimateS turns 20: statistical estimation

of species richness and shared species from samples, with non-parametric extrapolation. *Ecography* **37**: 609—613.

Dannemiller KC, Lang-Yona N, Yamamoto N, Rudich Y, Peccia J. (2014a). Combining real-time PCR and next-generation DNA sequencing to provide quantitative comparisons of fungal aerosol populations. *Atmos Environ* **84:** 113—121.

Dannemiller KC, Reeves D, Bibby K, Yamamoto N, Peccia J. (2014b). Fungal high-throughput taxonomic identification tool for use with next-generation sequencing (FHiTINGS). *J Basic Microbiol* **54:** 315—321.

de Hoog GS, Guarro J, Gené J, Figueras MJ. (2014). Atlas of clinical fungi. Ver 4.0. Centraalbureau voor Schimmelcultures.

Downs SH, Mitakakis TZ, Marks GB, Car NG, Belousova EG, Leuppi JD *et al.* (2001). Clinical importance of *Alternaria* exposure in children. *Am J Respir Crit Care Med* **164:** 455—459.

Driscoll CA, Macdonald DW, O'brien SJ. (2009). From wild animals to domestic pets, an evolutionary view of domestication. *Proc Natl Acad Sci USA* **106:** 9971—9978.

Edgar RC. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26:** 2460—2461.

Emmons CW. (1955). Saprophytic sources of *Cryptococcus neoformans* associated with the pigeon (*Columba livia*). *Am J Hyg* **62:** 227—232.

Fang ZG, Ouyang ZY, Hu LF, Wang XK, Zheng H, Lin XQ. (2005). Culturable airborne fungi in outdoor environments in Beijing, China. *Sci Total Environ* **350:** 47—58.

Fierer N, Liu ZZ, Rodriguez-Hernandez M, Knight R, Henn M, Hernandez MT. (2008). Short-term temporal variability in airborne bacterial and fungal populations. *Appl Environ Microbiol* **74:** 200—207.

Findley K, Oh J, Yang J, Conlan S, Deming C, Meyer JA *et al.* (2013).

Topographic diversity of fungal and bacterial communities in human skin. *Nature* **498:** 367—370.

Garcia-Solache MA, Casadevall A. (2010). Global warming will bring new fungal diseases for mammals. *Mbio* **1**.

Gardes M, Bruns TD. (1993). ITS primers with enhanced specificity for basidiomycetes — application to the identification of mycorrhizae and rusts. *Mol Ecol* **2:** 113—118.

Gomez-Heras M, Benavente D, de Buergo MA, Fort R. (2004). Soluble salt minerals from pigeon droppings as potential contributors to the decay of stone based cultural heritage. *Eur J Mineral* **16:** 505—509.

Good IJ. (1953). The population frequencies of species and the estimation of population parameters. *Biometrika* **40:** 237—264.

Granados DP, Castañeda E. (2006). Influence of climatic conditions on the isolation of members of the *Cryptococcus neoformans* species complex from trees in Colombia from 1992–2004. *FEMS Yeast Res* **6:** 636—644.

Guinea J, Pelaez T, Alcala L, Bouza E. (2006). Outdoor environmental levels of *Aspergillus* spp. conidia over a wide geographical area. *Med Mycol* **44:** 349—356.

Haag-Wackernagel D, Moch H. (2004). Health hazards posed by feral pigeons. *J Infect* **48:** 307—313.

Hasenclever HF, Kocan RM. (1975). *Candida albicans* associated with the gastrointestinal tract of the common pigeon (*Columba livia*). *Med Mycol* **13:** 116—120.

Hasnain SM. (1993). Influence of meteorological factors on the air spora. *Grana* **32:** 184—188.

Hetmanski T, Jarosiewicz A. (2008). Plumage polymorphism and breeding parameters of various feral pigeon (*Columba livia* Gm.) morphs in urban area (Gdansk, North Poland). *Pol J Ecol* **56:** 683—691.

Hubalek Z. (1975). Distribution of *Cryptococcus neoformans* in pigeon habitat. *Folia Parasit* **22:** 73—79.

Irinyi L, Serena C, Garcia-Hermoso D, Arabatzis M, Desnos-Ollivier M, Vu D *et al.* (2015). International society of human and animal mycology (ISHAM)—ITS reference DNA barcoding database—the quality controlled standard tool for routine identification of human and animal pathogenic fungi. *Med Mycol* **53:** 313—337.

Ishaq CM, Bulmer GS, Felton FG. (1968). An evaluation of various environmental factors affecting the propagation of *Cryptococcus neoformans*. *Mycopathologia* **35:** 81—90.

Jang YH, Lee SJ, Lee JH, Chae HS, Kim SH, Choe NH. (2011). Prevalence of yeast-like fungi and evaluation of several virulence factors from feral pigeons in Seoul, Korea. *Lett Appl Microbiol* **52:** 367—371.

Jerolmack C. (2008). How pigeons became rats: the cultural-spatial logic of problem animals. *Soc Probl* **55:** 72—94.

Jun NI. (2011). Social-historical changes of ‘minimal houses’ in Seoul and its spatial characteristics. *J Archit Inst Korea Plan Des* **27:** 191—202.

Khosravi AR. (1997). Isolation of *Cryptococcus neoformans* from pigeon (*Columba livia*) droppings in northern Iran. *Mycopathologia* **139:** 93—95.

Kim HJ, Lee SY, Yang WH. (2006). A study on efficient development size and design guideline proposal for the small multi-family houses in urban residential blocks. *Journal of The Urban Design Institute of Korea* **7:** 25—46.

Kuhn M. (2008). Building predictive models in R using the caret package. *J Stat Softw* **28:** 1—26.

Kurtzman CP, Robnett CJ, Ward JM, Brayton C, Gorelick P, Walsh TJ. (2005). Multigene phylogenetic analysis of pathogenic *Candida* species in the *Kazachstania* (*Arxiozyma*) *telluris* complex and description of

their ascosporic states as *Kazachstania bovina* sp. nov., *K. heterogenica* sp. nov., *K. pintolopesii* sp. nov., and *K. slooffiae* sp. nov. *J Clin Microbiol* **43:** 101—111.

Kwon-Chung KJ. (1976). Morphogenesis of *Filobasidiella neoformans*, the sexual state of *Cryptococcus neoformans*. *Mycologia* **68:** 821—833.

Lee SH, Lee HJ, Kim SJ, Lee HM, Kang H, Kim YP. (2010). Identification of airborne bacterial and fungal community structures in an urban area by T-RFLP analysis and quantitative real-time PCR. *Sci Total Environ* **408:** 1349—1357.

Leinonen R, Sugawara H, Shumway M, C INSD. (2011). The sequence read archive. *Nucleic Acids Res* **39:** D19—D21.

Levitz SM. (1991). The ecology of *Cryptococcus neoformans* and the epidemiology of cryptococcosis. *Rev Infect Dis* **13:** 1163—1169.

Littman ML. (1959). Cryptococcosis (Torulosis): current concepts and therapy. *Am J Med* **27:** 976—998.

Marquardt DW. (1980). You should standardize the predictor variables in your regression models. *J Am Stat Assoc* **75:** 87—91.

Nanjappa SG, Klein BS. (2014). Vaccine immunity against fungal infections. *Curr Opin Immunol* **28:** 27—33.

Newbound M, Mccarthy MA, Lebel T. (2010). Fungi and the urban environment: a review. *Landscape Urban Plann* **96:** 138—145.

Oh SY, Fong JJ, Park MS, Chang L, Lim YW. (2014). Identifying airborne fungi in Seoul, Korea using metagenomics. *J Microbiol* **52:** 465—472.

Pace NR. (1997). A molecular view of microbial diversity and the biosphere. *Science* **276:** 734—740.

Parameswaran P, Jalili R, Tao L, Shokralla S, Gharizadeh B, Ronaghi M *et al.* (2007). A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing. *Nucleic Acids Res*

35: e130.

Park D. (2013). The ecology of terrestrial fungi. In: Ainsworth, GC and Sussman, AS (ed). The Fungal Population: An Advanced Treatise**.** Academic Press: New York, **pp** 5—39.

Partridge BM, Winner HI. (1965). *Cryptococcus neoformans* in bird droppings in London. *The Lancet* **285:** 1060—1062.

Pei-Chih W, Huey-Jen S, Chia-Yin L. (2000). Characteristics of indoor and outdoor airborne fungi at suburban and urban homes in two seasons. *Sci Total Environ* **253:** 111—118.

Pfaller MA, Diekema DJ. (2010). Epidemiology of invasive mycoses in North America. *Crit Rev Microbiol* **36:** 1—53.

Przybylska K, Haidt A, Myczko L, Ekner-Grzyb A, Rosin ZM, Kwiecinski Z *et al.* (2012). Local and landscape-level factors affecting the density and distribution of the feral pigeon *Columba livia* var. *domestica* in an urban environment. *Acta Ornithologica* **47:** 37—45.

Ramirez R, Robertstad GW, Hutchinson LR, Chavez J. (1976). Mycotic flora in the lower digestive tract of feral pigeons (*Columba livia*) in the El Paso, Texas area. *J Wildl Dis* **12:** 83—85.

Rangel TF, Diniz-Filho JAF, Bini LM. (2010). SAM: a comprehensive application for spatial analysis in macroecology. *Ecography* **33:** 46—50.

Rao CY, Burge HA, Chang JCS. (1996). Review of quantitative standards and guidelines for fungi in indoor air. *J Air Waste Manage Assoc* **46:** 899—908.

Rappé MS, Giovannoni SJ. (2003). The uncultured microbial majority. *Annu Rev Microbiol* **57:** 369—394.

Reichard U, Herrmann S, Asif A. (2014). Vaccination approaches against opportunistic fungal infections caused by *Aspergillus fumigatus*. *Curr Protein Peptide Sci* **15**: 424—429.

Rose E, Nagel P, Haag-Wackernagel D. (2006). Spatio-temporal use of the urban habitat by feral pigeons (*Columba livia*). *Behav Ecol Sociobiol* **60:** 242—254.

Ruiz A, Neilson JB, Bulmer GS. (1982). Control of *Cryptococcus neoformans* in nature by biotic factors. *Sabouraudia* **20:** 21—29.

Ryberg M, Kristiansson E, Sjökvist E, Nilsson RH. (2009). An outlook on the fungal internal transcribed spacer sequences in genbank and the introduction of a web-based tool for the exploration of fungal diversity. *New Phytol* **181:** 471—477.

Sacchi R, Gentilli A, Razzetti E, Barbieri F. (2002). Effects of building features on density and flock distribution of feral pigeons *Columba livia* var. *domestica* in an urban environment. *Can J Zool* **80:** 48—54.

Sethi KK, Randhawa HS. (1968). Survival of *Cryptococcus neoformans* in gastrointestinal tract of pigeons following ingestion of organism. *J Infect Dis* **118:** 135—138.

Shelton BG, Kirkland KH, Flanders WD, Morris GK. (2002). Profiles of airborne fungi in buildings and outdoor environments in the United States. *Appl Environ Microbiol* **68:** 1743—1753.

Solomon WR. (1976). A volumetric study of winter fungus prevalence in the air of midwestern homes. *J Allergy Clin Immunol* **57:** 46—55.

Soltani M, Bayat M, Hashemi SJ, Zia M, Pestechian N. (2013). Isolation of *Cryptococcus neoformans* and other opportunistic fungi from pigeon droppings. *J Res Med Sci* **18:** 56—60.

Stringham SA, Mulroy EE, Xing J, Record D, Guernsey MW, Aldenhoven JT *et al.* (2012). Divergence, convergence, and the ancestry of feral populations in the domestic rock pigeon. *Curr Biol* **22:** 302—308.

Summerbell RC, Staib F, Dales R, Nolard N, Kane J, Zwanenburg H *et al.* (1992). Ecology of fungi in human dwellings. *Med Mycol* **30:** 279—

285.

Swinne-Desgain D. (1974). The pigeon as reservoir for *Cryptococcus neoformans*. *Lancet* **304:** 842—843.

Tabachnick BG, Fidell LS. (2013). Using multivariate statistics, 6th ed. Pearson: Boston.

Tchernov E. (1984). Commensal animals and human sedentism in the Middle East. In: Clutton-Brock J and Grigson C (ed). Animals and Archaeology: 3. Early herders and their flocks. British Archaeological Reports:

Oxford, **pp** 91—115.

R Development Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation For Statistical Computing: Vienna, Austria (http://www.R-project.org).

Tokarzewski S, Ziolkowska G, Lopuszynski W, Nozdryn-Plotnicki Z. (2007). *Aspergillus fumigatus* infection in a pigeon flock. *Bull Vet Inst Pulawy* **51:** 563—567.

Uden NV. (1960). The occurrence of *Candida* and other yeasts in the intestinal tracts of animals. *Ann N Y Acad Sci* **89:** 59—68.

Underhill DM, Iliev ID. (2014). The mycobiota: interactions between commensal fungi and the host immune system. *Nature Rev Immunol* **14:** 405—416.

Walter JE, Yee RB. (1968). Factors that determine the growth of *Cryptococcus neoformans* in avian excreta. *Am J Epidemiol* **88:** 445—450.

White TJ, Bruns T, Lee S, Taylor J. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innins MA, Gelfand DH, Sninsky JJ, White TJ (eds). PCR Protocols. Academic Press: San Diego, **pp** 315—322.

Yamamoto N, Shendell DG, Peccia J. (2011). Assessing allergenic fungi in house dust by floor wipe sampling and quantitative PCR. *Indoor Air*

21: 521—530.

Yamamoto N, Bibby K, Qian J, Hospodsky D, Rismani-Yazdi H, Nazaroff WW *et al.* (2012). Particle-size distributions and seasonal diversity of allergenic and pathogenic fungi in outdoor air. *ISME J* **6:** 1801—1811.

Yamamoto N, Hospodsky D, Dannemiller KC, Nazaroff WW, Peccia J. (2015). Indoor emissions as a primary source of airborne allergenic fungal particles in classrooms. *Environ Sci Technol* **49**: 5098–5106.

Zhou G, Whong WZ, Ong T, Chen B. (2000). Development of a fungus-specific PCR assay for detecting low-level fungi in an indoor environment. *Mol Cell Probes* **14:** 339—348.

7. Abstract in Korean

메타지놈 분석을 통한 도심 비둘기 분변 내 병원성 진균에 대한 연구: 보건학적 함의를 중심으로

이원동

생명과학부 대학원

서울대학교

초 록

도시 지역에서의 감염 위험이 증가함에 따라 인체 감염성 진균으로 인한 건강 우려가 전세계적으로 확산되고 있다. 이 같은 우려는 도심 비둘기 개체수의 급격한 증가와 맞물려, 도심 비둘기가 병원성 진균의 전파와 밀접한 관련이 있는 것으로 해석되기에 이르렀다. 현재까지 전세계적으로 약 48종의 병원성 진균이 비둘기 분변에 존재하는 것으로 보고되어 왔으나 이는 배양이 되지 않는 종들을 배제한 것으로 실제로는 더 많은 종이 비둘기 분변에 존재할 것으로 생각된다.

본 연구의 목적은, 첫째, 기존의 배양 기반이 아닌 차세대 염기서열 분석방법을 이용하여 비둘기 분변에 존대하는 병원성

진균의 다양성을 조사하고, 둘째, 실시간 유전자정량증폭 기술을 이용하여 병원성 진균의 절대량과 서울시내 지역적 분포를 확인하며, 셋째, 병원성 진균의 다양성 및 절대량과 관련 있는 지리, 사회, 그리고 기후 인자들을 파악하는 것이다. 마지막으로 신선한 비둘기 분변과 오래된 분변 내 존재하는 병원성 진균들을 조사하여 병원성 진균 전파에 실제로 비둘기가 관여하는지 여부를 확인해보고자 하였다. 신선한 비둘기 분변에서 발견되는 병원성 진균은 비둘기의 장을 통과하는 것으로서 비둘기에 의해 전파되는 종들로 간주하였고, 오래된 분변에서만 발견되는 종들은 비둘기가 아닌 주변의 환경에서 유래한 것으로 해석하였다.

연구 결과, 비둘기 분변 내 병원성 진균 다양성과 절대량은 서울시내에 불균등하게 분포되어 있음을 알 수 있었다. 병원성 진균의 다양성과 높은 양의 상관관계를 가지는 인자들로는 녹지율과 다세대주택의 수가 있었고, 바람의 세기는 음의 상관관계를 나타내었다. 한편 병원성 진균의 절대량은 도시 중심으로부터의 거리, 습도, 바람의 세기와 모두 음의 상관관계를 보였다. 오래된 비둘기 분변에서 발견되는 병원성 진균의 상당수가 신선한 분변에서는 발견되지 않은 것으로 미루어 우리는 대부분의 병원성 진균이 비둘기를 통해서가 아닌 주변 환경으로부터 유래한 것이라고 판단하였다. 즉, 비둘기가 아닌 비둘기의 분변이 병원성 진균의 전파에 더 중요한 역할을 하는 것으로 확인되었다.

이 같은 연구결과는 향후 보건당국이 보건 정책을 수립하고

병원성 진균으로 인한 피해를 예방하는데 있어 유용한 정보를 제공할 것으로 기대된다.

주요어: 인체 감염성 진균 • 비둘기 • 차세대 염기서열 분석 • 실시간 유전자정량증폭 • 비배양법 • 미생물 생태학

학 번: 2013-22961