Association of *Streptomyces* community composition determined by PCR-denaturing gradient gel electrophoresis with indoor mold status

Elisabet Johansson[1], **Tiina Reponen**[1], **Jarek Meller**[1], **Stephen Vesper**[2], and **Jagjit Yadav**[1]

[1]Department of Environmental Health, University of Cincinnati College of Medicine, Box 670056, Cincinnati, OH 45267-0056, USA

[2]US Environmental Protection Agency, 26 W. M. L. King Drive, Cincinnati, OH 45268, USA

Abstract

Both *Streptomyces* species and mold species have previously been isolated from moisture-damaged building materials; however, an association between these two groups of microorganisms in indoor environments is not clear. In this study we used a culture-independent method, PCR denaturing gradient gel electrophoresis (PCR-DGGE) to investigate the composition of the *Streptomyces* community in house dust. Twenty-three dust samples each from two sets of homes categorized as high-mold and low-mold based on mold specific quantitative PCR-analysis were used in the study. Taxonomic identification of prominent bands was performed by cloning and sequencing. Associations between DGGE amplicon band intensities and home mold status were assessed using univariate analyses, as well as multivariate recursive partitioning (decision trees) to test the predictive value of combinations of bands intensities. In the final classification tree, a combination of two bands was significantly associated with mold status of the home (p = 0.001). The sequence corresponding to one of the bands in the final decision tree matched a group of *Streptomyces* species that included S. *coelicolor* and *S. sampsonii*, both of which have been isolated from moisture-damaged buildings previously. The closest match for the majority of sequences corresponding to a second band consisted of a group of *Streptomyces* species that included *S. hygroscopicus*, an important producer of antibiotics and immunosuppressors. Taken together, the study showed that DGGE can be a useful tool for identifying bacterial species that may be more prevalent in mold-damaged buildings.

Keywords

Streptomyces; denaturing gradient gel electrophoresis; house dust; mold; decision trees

1. Introduction

Moisture in buildings is associated both with material damage and health outcomes, notably upper respiratory tract symptoms and asthma. There is strong epidemiological evidence that

---

*Corresponding author: Elisabet Johansson, University of Cincinnati, Department of Environmental Health, P.O. Box 670056 (3223 Eden Avenue), Cincinnati, OH, 45267-0056, USA, Tel. 513-558-4614; Fax 513-558-2263; Hanna.Johansson@cchmc.org.

mold growth in buildings is associated with respiratory health effects, even though a causal link between indoor microbial exposure and respiratory disease has not been conclusively established (Jones et al. 2011; Mendell et al. 2011; Reponen et al. 2011; Sahakian et al. 2008).

Moisture problems and mold growth are usually accompanied by bacterial growth as well, and species from both Gram-negative and Gram-positive bacterial taxa have been isolated from moisture-damaged building materials (Rintala et al. 2002; Suihko et al. 2009; Torvinen et al. 2006). Although species of other bacterial genera are often found in moisture damaged buildings, streptomycetes have attracted particular attention, and are considered indicators of moisture damage (Nevalainen et al. 1991). S*treptomyces* is a large genus of actinomycetes which are Gram-positive, spore-forming soil bacteria that can thrive on building materials under high moisture conditions. Streptomycetes are very versatile in their nutrient requirements and thrive on a wide variety of substrates, including many man-made materials used in building construction, such as concrete, ceramics, paint, and plasterboard.Several *Streptomyces* species have been isolated and identified directly from moisture-damaged building materials using both culturing and DNA-based techniques such as sequencing and ribotyping, with *Streptomyces griseus* and *Streptomyces coelicolor* being among the most commonly found species (Suihko et al. 2009; Torvinen et al. 2006). A further reason for the study of streptomycetes in the indoor environment is their production of secondary metabolites, with biological activities including antimicrobial, antitumor, immunosuppressive, antinflammatory, and cytotoxic properties, among others. *In vitro* and *in vivo* studies have demonstrated the toxic and inflammatory potential of some *Streptomyces* species, which makes airborne streptomycetes relevant to human health (Andersson et al. 1998; Hirvonen et al. 1997; Jussila et al. 1999; Jussila et al. 2003; Kirst et al. 1996). Secondary metabolites produced by streptomycetes have been shown to frequently co-occur with mycotoxins in moisture-damaged buildings (Täubel et al. 2011).

Microbial growth in association with moisture damage can lead to the release of inhalable spores and microbial fragments in indoor air. Both air and dust sampling have been used to approximate airborne microbial exposure. While air sampling may give a more accurate estimate of short-term exposure to aerosolized microbial components, dust samples represent integrated sampling over longer periods of time. Total levels of streptomycetes in household dust have been investigated in several studies. In a Finnish study using conventional PCR, Rintala et al (2004) observed a borderline significant association between indoor moisture damage and amplification of dust-borne streptomycetes. A later report based on quantitative PCR (qPCR), however, did not show a significant association between dust-borne levels of streptomycetes and moisture damage (Lignell et al., 2008). We recently investigated the sources of indoor dust-borne streptomycetes using qPCR, and did not detect a significant association between moisture damage and levels of streptomycetes (Johansson et al. 2011).

Streptomycetes in indoor environments can originate from both indoor and outdoor sources. While certain strains are known to thrive on moisture damaged building materials, much of what is found by indoor air and dust sampling is likely transported in from the outside environment by ground traffic or through the air (Johansson et al. 2011). This may be the

reason for the lack of associations between moisture damage and total levels of streptomycetes in earlier studies. A more detailed characterization of the indoor *Streptomyces* community may make it possible to identify species that are characteristic of moisture damage. Denaturing gradient gelelectrophoresis (DGGE), a culture-independent genetic fingerprinting technique, is one such approach that has the potential to address this knowledge gap. This technique was originally developed for the detection of single base mutations in DNA sequences, and has later found numerous applications in the field of microbial ecology (Muyzer and Smalla 1998). DGGE is rapid and reproducible, and can often resolve PCR-amplified ribosomal DNA fragments that differ by as little as a single nucleotide in fragments up to 600-700 base pairs. Sequencing of the ribosomal DNA-based PCR products (amplicons) can further allow identification of the source organism at the genus and/or species level. The study presented here was designed as a pilot study to explore the potential of PCR-DGGE in conjunction with amplicon DNA sequencing to resolve the *Streptomyces* community into individual species or groups of species, and to identify species with levels that differ between mold-damaged buildings and undamaged buildings.

2. Matherials and methods

2.1. Dust samples

The dust was collected from homes of children that were part of a birth cohort, the Cincinnati Childhood Asthma and Air Pollution Study (CCAAPS) (LeMasters et al. 2006). One hundred and seventy-eight homes in the Greater Cincinnati/Northern Kentucky area had previously been selected from the CCAAPS cohort for dust sampling and home inspection at child's age seven. The selection of the 178 homes was based on home inspections at child's age one, with 50% of the selected homes previously classified as having low observed mold and 50% classified as having high observed mold (Reponen et al. 2010). Homes for the present study were selected based on the mold status determined using Environmental Relative Moldiness Index (ERMI; described below). Among the 46 study homes, 23 homes had the highest mold-burden (ERMI > 10) and 23 homes had the lowest mold-burden (ERMI < −5) at the time of dust sampling at child's age seven. The study was approved by the University of Cincinnati Institutional Review Board, and informed consent was obtained by a parent at the time of each home visit.

The dust sampling has been described previously (Johansson et al. 2011; Cho et al. 2006). Briefly, dust was collected from carpeted floors by vacuuming a 2-m^2 area at a rate of 2 min/m^2. For hardwood floors, an entire room was vacuumed at a rate of 1 min/m^2. Collected dust was sieved through a 355-μm sieve, and the resulting fine dust was stored at −20°C.

2.2. Measurement of dust-borne indoor molds and streptomycetes

ERMI is based on the quantification of 26 mold strains characteristic of moisture-damaged buildings (group 1 molds) and 10 strains of outdoor origin that are characteristic of undamaged buildings (group 2 molds). The measurement of mold strains in dust by quantitative PCR and the calculation of the ERMI values has been described in detail elsewhere (Haugland et al. 2004; Vesper et al., 2007). Briefly, the sum of the log-

transformed values for levels of group 2 molds is subtracted from the sum of the log-transformed values for group 1 molds to obtain the value for ERMI.

Total concentration of streptomycetes was measured by quantitative PCR (qPCR) as described earlier (Johansson et al. 2011).

2.3. DNA extraction and PCR amplification for DGGE

DNA was extracted from dust using a bead beating protocol. Ten mg dust from each sample was added to 2-ml sterile tubes containing 0.3 g acid-washed glass beads (#G1277; Sigma-Aldrich, St Louis, MO, USA). After addition of 0.3 ml Lysis buffer (GeneRite, North Brunswick, NJ, USA) the tubes were shaken in a Mini Bead-Beater (Biospec Products, Bartlesville, OH, USA) at maximum speed for 1 minute. DNA was isolated using the DNA-EZ kit from GeneRite (North Brunswick, NJ, USA) according to manufacturer's instructions. Sixty µl of elution buffer were used to elute the DNA.

Nested amplification of a portion of the 16S rRNA gene was used to obtain amplified DNA for the gel electrophoresis. For the outer reaction primers 235F and 1392R (Inbar et al. 2005) were used. Each reaction contained one µl DNA, 0.25 µM of each primer, 200 µM of each deoxynucleotide, and 0.5 U TaKaRa Ex Taq polymerase in a total volume of 20 µl. To ensure similar total intensities in all lanes, qPCR results for total concentrations of streptomycetes in the dust samples were used to dilute extracted DNA to correspond to 167 *Streptomyces* cell equivalents/µl template. Amplifications were performed in a Perkin-Elmer PE 9700 Thermal Cycler, using the following protocol: initial denaturation at 95°C for 4 min, followed by 20 cycles of denaturation at 94°C for 30 s, annealing at 62°C for 30 s, and extension at 72°C for 60 s, with an additional extension step at 72°C for 7 minutes. Products from the outer amplification were diluted 1:50, and an aliquot (1 µl) was used in nested amplifications. The primers used were S661-F and F/S1218-R; the latter contained a 40-bp GC-clamp as described (Inbar et al. 2005). The reaction conditions for the nested PCR were the same as those for outer reactions, except that 0.5 µM of each primer was used, extensions carried out for 45 s, and amplifications proceeded for a total of 24 cycles.

2.4. Denaturing gradient gel electrophoresis (DGGE)

Amplified DNA was resolved by DGGE using the DCode Universal Mutation Detection System from Bio-Rad Laboratories (Hercules, California, USA) according to manufacturer's instructions. Six µl nested PCR product was loaded on a 6% polyacrylamide gel with a 30-60% denaturant gradient in 1× Tris-acetate-EDTA (TAE). Electrophoresis was performed at 150 V for 7 hours, and the gels were stained with $1 \times$ SYBR® Gold (Life Technologies, Grand Island, NY, USA) in $1 \times$ TAE for 10 min and photographed using a Kodak digital camera (Eastman Kodak Co., New Haven, Connecticut, USA). Gel images were processed and bands were identified and quantified using GelCompar II (Applied Maths, Inc., Austin, Texas, USA).

2.5. Cloning and sequencing of individual bands

For the identification of phylotypes corresponding to individual bands, DNA from two dust samples was directly cloned into the pCR®2.1-TOPO® vector (Life Technologies, Grand

Island, NY, USA) according to manufacturer's instructions using One Shot® Mach1™ Chemically Competent *E. coli* included in the cloning kit for transformation. Individual colonies were used for colony PCR (15-μl PCR reactions) as well as for preparing a subculture by inoculating a 2-ml LB broth containing 50 μg/ml ampicillin. PCR primers and reaction conditions were the same as for nested PCR described above. The PCR products were analyzed by DGGE alongside PCR products from whole dust DNA. The clones that produced amplicons that co-migrated with bands of interest from whole dust DNA were selected for further analysis. Plasmid DNA was isolated from the 2-ml cultures using the QIAprep Spin Miniprep Kit (QIAGEN Inc., Valencia, CA, USA), and sequenced (Genewiz, South Plainfield, NJ,USA) using the T3 universal primer (5'-ATTAACCCTCACTAAAGGGA-3'). At least four clones for each band of interest were sequenced. DECIPHER (http://DECIPHER.cee.wisc.edu) was used to check for chimeric sequences. Edited sequences, with primer sequences excluded, were compared to GenBank entries by BLAST analysis (Altschul et al. 1990) at the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov), and to 16S rRNA sequences of type strains in the database of the Ribosomal Database Project, release 10 (Cole et al. 2009). Sequences representing relevant DGGE band classes determined in this study have been deposited in the GenBank database under accession numbers KC542345-KC542348. For further analysis, 16S rRNA sequences from 33 *Streptomyces* known species, including sequences from species previously associated with moisture damage and mold growth, were downloaded from GenBank and edited to correspond to the nucleotide positions of the amplified DGGE fragments. These sequences, together with sequences representing relevant bands in our study, were used as input for the construction of a phylogenetic tree using the PhyML online tool at Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier (http://www.phylogeny.fr/version2_cgi/phylogeny.cgi) with Approximate Likelihood-Ratio Test for branch support (Anisimova and Gascuel 2006).

2.6. Statistical and machine learning analyses

Bands were normalized in the GelCompar analyses as percentage of total lane intensity. Results from the qPCR analyses, which were expressed as the total number of streptomycetes per mg house dust, were used to recalculate the amount of streptomycetes in each DGGE band as number of cells per mg dust.

Histograms and Mann-Whitney rank sum tests were used to assess associations of individual bands, expressed as cells/mg, with the mold status. To create binary variables, histograms, Chi-square tests, and Fisher's exact tests were used to establish cut-off values that best predicted highversus low-mold homes for each band. Bands that differed between low- and high-mold homes at $p < 0.2$ in Mann-Whitney rank sum tests were explored further in the construction of decision trees and the resulting classification rules.

Here, we specifically used the standard CART decision tree (Breiman et al. 1984), as implemented in R (www.r-project.org). To evaluate the accuracy and stability of the resulting trees, we used the leave-one-out (LOO) form of cross-validation. LOO involves repeated training on all but one data point, which is put aside and subsequently used for testing. The procedure is repeated for each data point at a time, and it allows estimation of

accuracy on data not used for the training. For 46 samples (data points) in our case, LOO involved generating 46 different decision trees using a different subset of 45 data points each time. Importantly, features (data bands) that are most informative (provide best discrimination) were selected in each run using a subset of the data as well, thus providing an assessment of the stability of the tree (i.e., consistency of features selected as most discriminatory). In addition, we also generated decision trees using the entire data set, as discussed in the results section.

3. Results

3.1. Band identification

A total of 17 amplicon band classes (designated B1 through B17) were identified by gel image analysis using GelCompar. A majority of these band classes were of low intensity and/or occurred in less than a third of the samples. All band classes were analyzed for associations with mold status by Mann-Whitney rank sum tests, and identification was focused on those bands that were associated with mold at $p < 0.2$. One of these bands, B7, was faint, occurred in few of the samples, and was of somewhat poor reproducibility. Therefore, it was excluded from further analysis. Cloning and sequencing identified consensus sequences for bands B1, B4, B11, and B15 (Figure 1).

Out of eight clones (amplicons) analyzed that co-migrated as band B8, only two had over 98% sequence similarity, which suggests that this band is highly heterogeneous and made up of several different genotypes that were poorly resolved on the gel. The consensus sequence for B1, obtained from identical sequences from eight clones, was a 100% match to more than 20 species in the database of the Ribosomal Database Project (RDP), including *S. anulatus*, *S. californicus*, *S. finlayi*, *S. microflavus*, *S. flavogriseus*, and *S. fimicarius*, and 99.5% match to *S. griseus* (type strain KACC 20084). Six identical sequences from clones with DGGE bands that co-migrated with B15 were 100% matches to another group of more than 20 type strains, including *S. sampsonii*, *S. coelicolor*, *S. albidoflavus*, *S. luteogriseus*, and *S. violaceus*. Out of four analyzed clones that co-migrated with B4, two produced sequences that were 100% identical with *S. torulosus*, *S. hygroscopicus*, and *S. ipomoeae* type strains. A third clone had 99.6% similarity with these species. Two clones that co-migrated with B11 produced sequences with 100% similarity to *S. glauciniger* (AY314782), whereas two additional clones were less than 98% similar and did not produce identical matches to any entries in RDP or GenBank.

Figure 2 shows the phylogenetic relationships between consensus sequences for B1, B4, B11, and B15, and 16S rRNA sequences for 35 *Streptomyces* type strains downloaded from GenBank.

3.2. Analysis of association with mold status

After each band, expressed as percentage of total gel lane intensity, was recalculated as number of cells per mg of dust, five bands were identified that differed between low- and high-mold homes at $p < 0.2$ in Mann-Whitney rank sum tests: B1, B4, B8, B11, and B15. As can be seen in Table 1, the discriminatory power of these bands taken individually is limited.

Although two of the bands, namely B1 and B4 provide statistically significant results, their discriminatory power is also very limited. In order to further improve classification of sample type, we combined the individual bands using recursive partitioning techniques in the form of C4.5 decision trees. Both the numerical values for band intensities, as well as their binary projections (using thresholds identified as described above), were used to build and assess decision trees and the resulting logical rules that can be used for the classification of samples as high- vs. low-mold.

An example of a decision tree that combines bands B4 and B15 to provide more accurate prediction of high mold levels is shown in Figure 3A. As can be seen from Figure 3, very similar trees were obtained when using binary variables, defined by applying cut points that best predicted high-versus low-mold homes, and when using numerical variables (importantly, cut-off values optimized independently in the latter case were very close to those used for binary projections). Overall, the logical rules corresponding to these trees can be formulated as follows:

IF ((B4 400) OR ((B4 < 400) AND (B15 295))) THEN HIGH MOLD;

OTHERWISE LOW MOLD.

Thus, high mold can be predicted either by sufficiently high B4 amounts, or, when B4 amount is low, a combination of high B15 and low B4 amounts.

The stability and accuracy of decision rules (and associated logical rules) were assessed using leave-one-out (LOO) cross-validation. The use of LOO, as opposed to alternative strategies to evaluate generalization and accuracy was dictated by the small sample size. The trees obtained when using numerical values showed a relatively higher level of variation, with only 17 of the resulting 46 alternative trees identical to the original tree. The remaining 29 alternative trees represented seven different types involving B4, B15, B1, and B8. The tree for binary variables, on the other hand, was very stable, with 38 of the 46 alternative trees identical to the original tree. Seven of the alternative trees involved B4 only, and one tree involved B4 and B8.

To refine the classification trees, and to further simplify the resulting decision rules, we included in the classification models combinations of variables, which were created using the logical operator "OR". Such obtained combinations of variables were used again in two alternative models, one in conjunction with the continuous band variables, and one with the binary projections. The two resulting trees were identical, however (Figure 1B), and could be described by the corresponding simple logical rule as follows:

IF ((B4 400) OR (B15 295)) THEN HIGH MOLD;

OTHERWISE LOW MOLD.

Cross-validation analyses of both models produced trees that remained the same for 43 of the 46 training subsets. For the remaining three samples, removal resulted in a tree involving B4 and B11. The prediction accuracy in LOO cross-validation was estimated to be 70%, as opposed to 76% using the whole data set for training, or in other words 76% training

accuracy. The drop in prediction accuracy from 76% to 70% indicates that the extent of overfitting is limited despite the small sample size. It should be also noted that a simple baseline classifier that assigns all samples to the bigger class has 50% accuracy in this case, as both high- and low-mold classes consist of 23 samples.

Table 1 shows the p-values for each decision tree obtained with all 46 samples, as well as p-values for prediction accuracy of each individual band. As can be seen, the significance for all three composite decision rules involving more than 1 band was higher than that of any of the individual band variables.

4. Discussion

In this study, we addressed the hypothesis that the composition of bacterial communities from the *Streptomyces* genus can be used as a predictive marker of indoor mold levels. In order to test this hypothesis, we assessed associations between home mold status and *Streptomyces* DGGE band intensities. The latter represent the abundance of distinct *Streptomyces* species and their specific mixtures, as captured by gel migration patterns. The classification tree analysis revealed that two bands (bands 15 and 4) were significantly associated with mold status.

The sequence corresponding to band 15 (B15), which remained in the final decision trees, matched a large group of *Streptomyces* type strains, of which *S. coelicolor* and *S. sampsonii* have been isolated from moisture-damaged buildings previously (Rintala et al. 2002; Suihko et al. 2009). Among the three database species that matched the sequence of B4, the second variable in the final decision trees, was *S. hygroscopicus*, which is an important producer of a large variety of antibiotics and medically significant immunosuppressors (Kirst et al. 1996; Park et al. 2010). Although B1 was not a component of the final decision trees, this band alone as a binary variable was significantly associated with high-mold status (Table 1). The sequence of B1 matched a large group of *Streptomyces* species, of which *S. flavogriseus*, *S. anulatus*, *S. griseus*, and *S. californicus* have been associated with building moisture damage (Jussila et al. 1999; Rintala et al. 2002; Suihko et al. 2009).

When considered independently, individual bands provided only weak discriminatory power for the classification of homes as high- vs. low-mold. In order to test whether combinations of band intensities can be used to more accurately predict high mold levels, we used recursive partitioning in the form of decision trees. Decision trees are a standard machine learning technique for multivariate data analysis and classification. The advantage of decision trees is that they provide the ability to select the most discriminating features (here band classes) and intuitive logical rules that can be used to classify the samples. Decision trees can also be viewed as a recursive partitioning approach, in which data is hierarchically divided into strata (subtypes) that can be classified by simple logical rules. Such rules combine individual variables (here band intensities) into an easy-to-interpret multivariate classifier.

In PCR-DGGE based community composition analysis in microbial ecology studies, band patterns are most often evaluated using whole profile analysis. However, this may not be as

useful a strategy for environmental samples where different components of the microbial community are not always in ecological balance. House dust is considered a long-term reservoir of microbial species of both indoor and outdoor origin, and can be expected to contain a large part of dead or dormant spores and cells.

In this study we used indoor mold as an indirect measure of moisture damage and deduced its association with the *Streptomyces* composition. The mold index ERMI utilized for the purpose is calculated from levels of 26 indicator mold species, such as Aspergillus niger, A. flavus, A. versicolor, Penicillium brevicompactum, *Penicillium spinulosum*, and *Stachybotrys chartarum,* that have been associated with moisture damage in buildings (Thrasher and Crawley 2009; Vesper et al. 2007), and 10 ubiquitous mold species that are found in healthy buildings as well. *In vitro* studies have demonstrated that synergistic interactions occur between indoor molds and *Streptomyces* spores with regard to cytotoxic effects (Huttunen et al. 2004; Penttinen et al. 2005), and that co-cultivation potentiates immunotoxicity (Penttinen et al. 2006). Furthermore, endotoxin has been shown to augment the effects of fungi on respiratory symptoms and the proinflammatory reactions induced by the mycotoxin roridin A (Park et al. 2006; Islam et al. 2007). Taken together, these studies suggest that it may be necessary to assess a range of microbial exposures (such as *Streptomyces* composition as defined in this study) in addition to molds in risk assessments of indoor moisture damage. In contrast with the demonstrated association between *Streptomyces* composition and mold status, we were not able to detect significant associations between the *Streptomyces* DGGE bands and extent of observed moisture damage, which may have been due to the relatively small sample number, and the small number of homes with extensive moisture damage in our study (data not shown).

The achieved resolution in DGGE analysis depends on amplified fragment length, sequence variability in the amplified fragment for the targeted taxon, and the complexity of the microbial community under study. The *Streptomyces* community in dust was complex, and the amplicon of $\approx$ 614 bp was in the upper range of what can be resolved by DGGE. One of the analyzed bands in our study, B8, most likely consisted of a poorly resolved mix of phylotypes with no sequence variant predominating. In addition, the *Streptomyces* genus is a large genus composed of closely related species, and the amplicon produced by the primers S661-F and F/S1218-R may not always have enough sequence variability to provide resolution at the species level.

The dust samples used in this study were collected from homes in the Greater Cincinnati/ Northern Kentucky area, and the results may not be generalizable to other geographic locations. Also, the association with mold is specific to the four groups of *Streptomyces* species identified, and may not be applicable to other Gram-positive genera and species. The study does demonstrate, however, that associations with mold and moisture damage may not always be generalizable to entire microbial genera, and that associations may be unveiled by resolving genera into clades or individual species.

In this study we identified four groups of *Streptomyces* species that were more prevalent in homes with high mold compared with homes with low mold. The results may form a basis for designing more detailed comparative studies on the composition of the *Streptomyces*

community in moisture-damaged and moldy buildings. It may be possible to further resolve species and strain-level compositional differences in indoor *Streptomyces* populations by designing DGGE primers better suited to resolve the groups associated with high mold levels. Availability of such specific information may then allow quantitative PCR-based determination of the predominance of particular species and/or strains in these groups associated with the indoor environment, thereby enabling a more specific risk assessment.

Acknowledgments

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. Journal of Molecular Biology. 1990; 215:403–410. [PubMed: 2231712]

Andersson MA, Mikkola R, Kroppenstedt RM, Rainey FA, Peltola J, Helin J, Sivonen K, Salkinoja-Salonen MS. The mitochondrial toxin produced by *Streptomyces griseus* strains isolated from an indoor environment is valinomycin. Applied and Environmental Microbiology. 1998; 64:4767–4773. [PubMed: 9835560]

Anisimova M, Gascuel O. Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. Systematic Biology. 2006; 55:539–552. [PubMed: 16785212]

Breiman, L.; Friedman, J.; Olshen, RA.; Stone, CJ. Classification and regression trees. Wadsworth; Belmont, CA: 1984.

Cho SH, Reponen T, Bernstein DI, Olds R, Levin L, Liu X, Wilson K, LeMasters G. The effect of home characteristics on dust antigen concentrations and loads in homes. Science of the Total Environment. 2006; 371:31–43. [PubMed: 17049968]

Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ, Kulam-Syed-Mohideen AS, McGarrell DM, Marsh T, Garrity GM, Tiedje JM. The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. Nucleic Acids Research. 2009; 37:D141–D145. [PubMed: 19004872]

Haugland RA, Varma M, Wymer LJ, Vesper SJ. Quantitative PCR analysis of selected *Aspergillus, Penicillium* and *Paecilomyces* species. Systematic and AppliedMicrobiology. 2004; 27:198–210.

Hirvonen MR, Nevalainen A, Makkonen N, Monkkonen J, Savolainen K. Induced production of nitric oxide, tumor necrosis factor, and interleukin-6 in RAW 264.7 macrophages by streptomycetes from indoor air of moldy houses. Archives of Environmental Health. 1997; 52:426–432. [PubMed: 9541363]

Huttunen K, Pelkonen J, Nielsen KF, Nuutinen U, Jussila J, Hirvonen MR. Synergistic interaction in simultaneous exposure to Streptomyces californicus and Stachybotrys chartarum. Environmental Health Perspectives. 2004; 112:659–665. [PubMed: 15121507]

Inbar E, Green SJ, Hadar Y, Minz D. Competing factors of compost concentration and proximity to root affect the distribution of streptomycetes. Microbial Ecology. 2005; 50:73–81. [PubMed: 16052380]

Islam Z, Amuzie CJ, Harkema JR, Pestka JJ. Neurotoxicity and inflammation in the nasal airways of mice exposed to the macrocyclic trichothecene mycotoxin roridin a: kinetics and potentiation by bacterial lipopolysaccharide coexposure. Toxicological Sciences. 2007; 98:526–541. [PubMed: 17483119]

Johansson E, Vesper S, Levin L, LeMasters G, Grinshpun S, Reponen T. Streptomycetes in house dust: associations with housing characteristics and endotoxin. Indoor Air. 2011; 21:300–310. [PubMed: 21204988]

Jones R, Recer GM, Hwang SA, Lin S. Association between indoor mold and asthma among children in Buffalo, New York. Indoor Air. 2011; 21:156–164. [PubMed: 21204984]

Jussila J, Ruotsalainen M, Komulainen H, Savolainen K, Nevalainen A, Hirvonen MR. *Streptomyces anulatus* from indoor air of moldy houses induce NO and IL-6 production in a human alveolar epithelial cell-line. Environmental Toxicology and Pharmacology. 1999; 7:261–266. [PubMed: 21781934]

Jussila J, Pelkonen J, Kosma VM, Maki-Paakkanen J, Komulainen H, Hirvonen MR, Systemic immunoresponses in mice after repeated exposure of lungs to spores of Streptomyces californicus. Clinical and Diagnostic Laboratory Immunology. 2003; 10:30–37. [PubMed: 12522036]

Kirst HA, Mynderse JS, Martin JW, Baker PJ, Paschal JW, Rios Steiner JL, Lobkovsky E, Clardy J. Structure of the spiroketal-macrolide ossamycin. Journal of Antibiotics (Tokyo). 1996; 49:162–167.

LeMasters GK, Wilson K, Levin L, Biagini J, Ryan P, Lockey JE, Stanforth S, Maier S, Yang J, Burkle J, Villareal M, Khurana Hershey GK, Bernstein DI. High prevalence of aeroallergen sensitization among infants of atopic parents. Journal of Pediatrics. 2006; 149:505–511. [PubMed: 17011322]

Lignell U, Meklin T, Rintala H, Hyvarinen A, Vepsalainen A, Pekkanen J, Nevalainen A. Evaluation of quantitative PCR and culture methods for detection of house dust fungi and streptomycetes in relation to moisture damage of the house. Letters in Applied Microbiology. 2008; 47:303–308. [PubMed: 19241524]

Mendell MJ, Mirer AG, Cheung K, Tong M, Douwes J. Respiratory and allergic health effects of dampness, mold, and dampness-related agents: a review of the epidemiologic evidence. Environmental Health Perspectives. 2011; 119:748–756. [PubMed: 21269928]

Muyzer G, Smalla K. Application of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) in microbial ecology. Antonie Van Leeuwenhoek. 1998; 73:127–141. [PubMed: 9602286]

Nevalainen A, Pasanen AL, Niininen M, Reponen T, Kalliokoski P, Jantunen MJ. The indoor air quality in Finnish homes with mold problems. Environment International. 1991; 17:299–302.

Park JH, Cox-Ganser J, Rao C, Kreiss K. Fungal and endotoxin measurements in dust associated with respiratory symptoms in a water-damaged office building. Indoor Air. 2006; 16:192–203. [PubMed: 16683938]

Park SR, Yoo YJ, Ban YH, Yoon YJ. Biosynthesis of rapamycin and its regulation: past achievements and recent progress. Journal of Antibiotics (Tokyo). 2010; 63:434–441.

Penttinen P, Pelkonen J, Huttunen K, Toivola M, Hirvonen MR. Interactions between Streptomyces californicus and Stachybotrys chartarum can induce apoptosis and cell cycle arrest in mouse RAW264.7 macrophages. Toxicology and Applied Pharmacology. 2005; 202:278–288. [PubMed: 15667833]

Penttinen P, Pelkonen J, Huttunen K, Hirvonen MR. Co-cultivation of *Streptomyces californicus* and *Stachybotrys chartarum* stimulates the production of cytostatic compound(s) with immunotoxic properties. Toxicology and Applied Pharmacology. 2006; 217:342–351. [PubMed: 17095029]

Reponen T, Singh U, Schaffer C, Vesper S, Johansson E, Adhikari A, Grinshpun SA, Indugula R, Ryan P, Levin L, Lemasters G. Visually observed mold and moldy odor versus quantitatively measured microbial exposure in homes. Science of the Total Environment. 2010; 408:5565–74. [PubMed: 20810150]

Reponen T, Vesper S, Levin L, Johansson E, Ryan P, Burkle J, Grinshpun SA, Zheng S, Bernstein DI, Lockey J, Villareal M, Khurana Hershey GK, LeMasters G. High environmental relative moldiness index during infancy as a predictor of asthma at 7 years of age. Annals of Allergy, Asthma & Immunology. 2011; 107:120–126.

Rintala H, Nevalainen A, Suutari M. Diversity of streptomycetes in water-damaged building materials based on 16S rDNA sequences. Letters in Applied Microbiology. 2002; 34:439–443. [PubMed: 12028426]

Rintala H, Hyvarinen A, Paulin L, Nevalainen A. Detection of streptomycetes in house dust-- comparison of culture and PCR methods. Indoor Air. 2004; 14:112–119. [PubMed: 15009417]
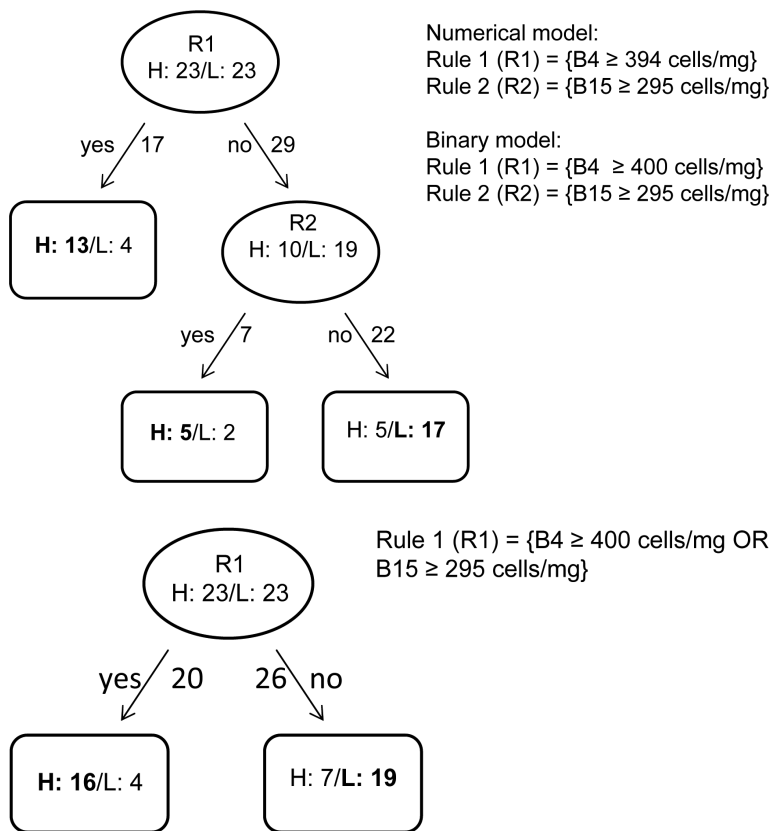
Sahakian NM, Park JH, Cox-Ganser JM. Dampness and mold in the indoor environment: implications for asthma. Immunology and Allergy Clinics of North America. 2008; 28:485–505. vii. [PubMed: 18572103]

Suihko ML, Priha O, Alakomi HL, Thompson P, Malarstig B, Stott R, Richardson M. Detection and molecular characterization of filamentous actinobacteria and thermoactinomycetes present in water-damaged building materials. Indoor Air. 2009; 19:268–277. [PubMed: 19302504]

Täubel M, Sulyok M, Vishwanath V, Bloom E, Turunen M, Järvi K, Kauhanen E, Krska R, Hyvärinen A, Larsson L, Nevalainen A. Co-occurrence of toxic bacterial and fungal secondary metabolites in moisture-damaged indoor environments. Indoor Air. 2011; 21:368–75. [PubMed: 21585551]

Thrasher JD, Crawley S. The biocontaminants and complexity of damp indoor spaces: more than what meets the eyes. Toxicology and Industrial Health. 2009; 25:583–615. [PubMed: 19793773]

Torvinen E, Meklin T, Torkko P, Suomalainen S, Reiman M, Katila ML, Paulin L, Nevalainen A. Mycobacteria and fungi in moisture-damaged building materials. Applied Environmental Microbiology. 2006; 72:6822–6824. [PubMed: 17021236]

Vesper S, McKinstry C, Haugland R, Wymer L, Bradham K, Ashley P, Cox D, Dewalt G, Friedman W. Development of an Environmental Relative Moldiness index for US homes. Journal of Occupational Environmental Medicine. 2007; 49:829–833. [PubMed: 17693779]
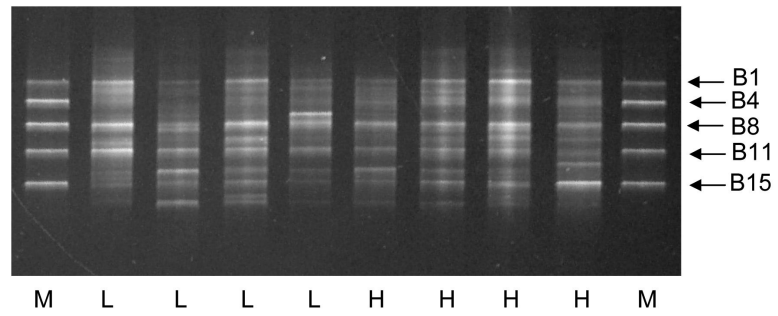
Figure 1.
Representative DGGE gel showing amplified PCR products profile from four low-mold and four high-mold house dust samples. M: marker; L: DNA from low-mold sample; H: DNA from high-mold sample.
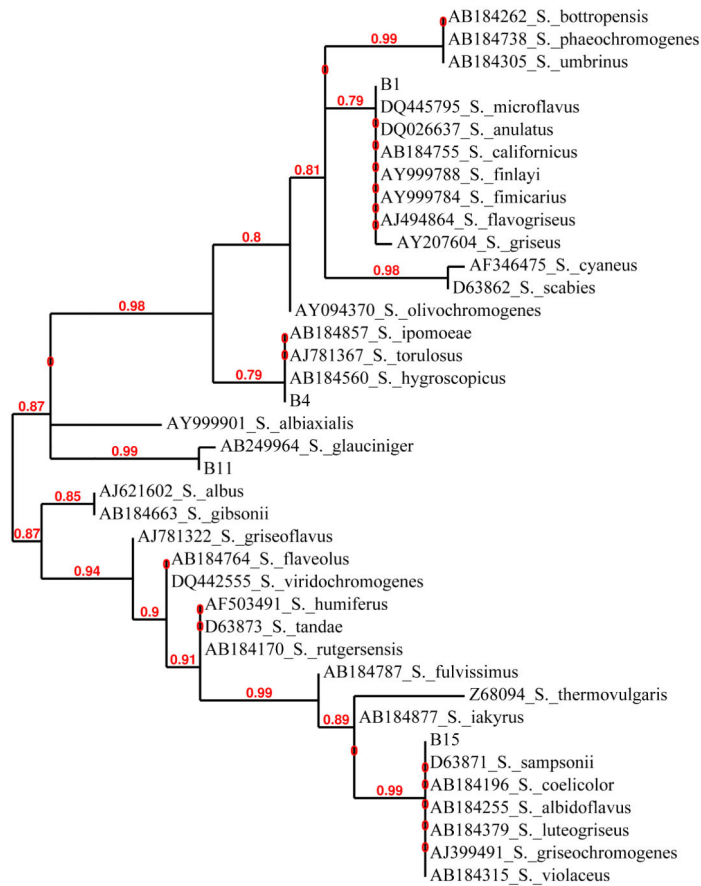
Figure 2.
Phylogenetic tree drawn based on the sequences for Streptomyces DGGE bands B1, B4, B11, and B15, together with the sequences for 33 *Streptomyces* species downloaded from GenBank. The tree was constructed and drawn with the PhyML online tool using Approximate Likelihood-Ratio Test for branch support.

R1
H: 23/L: 23

yes / 17   no \ 29

H: **13**/L: 4

R2
H: 10/L: 19

yes / 7   no \ 22

H: **5**/L: 2   H: 5/**L: 17**

Numerical model:
Rule 1 (R1) = {B4 ≥ 394 cells/mg}
Rule 2 (R2) = {B15 ≥ 295 cells/mg}

Binary model:
Rule 1 (R1) = {B4 ≥ 400 cells/mg}
Rule 2 (R2) = {B15 ≥ 295 cells/mg}

R1
H: 23/L: 23

yes / 20   26 \ no

H: **16**/L: 4   H: 7/**L: 19**

Rule 1 (R1) = {B4 ≥ 400 cells/mg OR
B15 ≥ 295 cells/mg}

Figure 3.
Classification trees for all 46 dust samples using four different classification models. L: Low-mold sample. H: High-mold sample. **3A.** Model based on numerical values for Streptomycetes content in each band. **3B.** Model based on pair-wise combinations of relevant bands using the logical operator "OR". Combination variables were used together with numerical values and binary values for individual bands in two different models, which produced the same tree.

Table 1

Significance of decision rules for prediction of mold levels (high or low). P-values were obtained by Chi-square tests.

Decision rule	Number of low-mold samples[a]	Number of high-mold samples[a]	P value
B1 500			0.049
yes	3	10	
no	20	13	
B4 400			0.015
yes	4	13	
no	19	10	
B8 600			0.24
yes	8	13	
no	15	10	
B11 500			0.17
yes	3	8	
no	20	15	
B15 295			0.11
yes	4	10	
no	19	13	
BC4 400 OR (BC4 < 394 AND BC15 295)			0.001
yes	6	18	
no	17	5	
BC4 400 OR (BC4 < 400 AND BC15 295)			0.001
yes	6	18	
no	17	5	
B4 400 OR B15 295			0.001
yes	4	16	
no	19	7	