

## Strain Diversity within a Microbial Collection

DENNIS DiTULLIO<sup>a,†</sup>, NANCY ROCHELEAU<sup>a,††</sup>, M. KELLY TALBOT<sup>a,†††</sup>, WILLIAM MILLETT<sup>a,†</sup>,  
JAMES MCALPINE<sup>a,\*</sup> and PHIL ROBAKIEWICZ<sup>b</sup>

<sup>a</sup>Phytera, Inc.

377 Plantation Street, Worcester, MA 01605

<sup>b</sup>Worcester Polytechnic Institute

(Received for publication June 11, 2002)

Random subsets of a modest-sized microbial collection have been examined for culture redundancy, initially by morphology, both to the naked eye and microscopically, of cultures grown on a variety of agar-based solid media. Subsequent analysis, by simple TLC, of the extractable metabolites produced by morphologically similar cultures grown in submerged shaken fermentation was carried out. Apparent duplicate cultures were further examined on Biolog SF-P MicroPlates<sup>TM</sup> for differentiation. The results were subjected to a statistical analysis and the contribution of each stage in the process to resolving culture uniqueness was noted. A statistical extrapolation of the results, from the subsets of each culture type, to the total for that type within the entire collection, with 95% confidence limits, is presented. Morphological comparison, on four different agar media, gives a significant underestimation of the metabolic diversity of the collection. The weighted mean from the three types of cultures indicate that the expected content of the collection is approximately 93% unique strains.

For more than 60 years the pharmaceutical industry has sought new drugs, or drug leads, *via* screening of microbial fermentations. Initially directed at antibiotics, the search expanded following the impetus from HAMA O UMEZAWA<sup>1)</sup> in 1965 to include enzyme inhibitors and receptor binding agonists and antagonists, applicable to virtually all areas of therapy. The compounds sought are usually described as secondary metabolites. Although the definition of this term tends to vary with user, most expect that specific secondary metabolites will be produced by a limited number of species and that the goal of chemical diversity is well served by biological diversity. Thus large collections of microbes have been assembled to achieve this. Within a collection, one aims to have each organism represented only once, as metabolite production by a given organism grown under specific conditions tends to be highly reproducible, and redundancy within the collection leads to wasted effort. A difficulty arises, in that assessing the

extent of, and minimizing, duplication of cultures within a collection becomes disproportionately more difficult as the collection increases in size. The gold standard for comparing cultures is 16S ribosomal-RNA analysis, and we have considered this methodology. Although it is definitive and widely accepted by taxonomists in particular, it has two major disadvantages. The most obvious is cost, which is considerable and can be prohibitive for even a medium sized library. A more fundamental concern is the applicability of the results. Whereas RNA analysis is ideal for taxonomic differentiation of strains, configuring a library consisting entirely of unique species may not accomplish the primary goal of optimizing chemical diversity production. The literature abounds in examples of different strains of the same species making very different secondary metabolites. The common streptomycete, *S. hygroscopicus* is listed in the Chapman and Hall database<sup>2)</sup> as the source of 261 different natural products entries. An

\* Corresponding author: mcalpine@ecopiabio.com

Current address: Ecopia Biosciences Inc., 7290 Frederick Banting, Saint Laurent, H4-S 2A1, Canada.

† Current address: Cetek Corporation, 260 Cedar Hill St., Marlborough, MA 01752, USA.

†† Work undertaken as an intern at Phytera, while a graduate student at Worcester State College.

††† Deceased October 26th, 2001.

inspection of these reveals 179 different elucidated structures spread across 25 distinctly different chemical classes. The significance here is that seldom does one strain of a given species produce more than one or two chemical classes of antibiotic. Moreover some structural types tend to be restricted to a single strain isolate. E.g. Rapamycin<sup>3</sup> was, until the report last year, produced only by a single known strain of *S. hygroscopicus* isolated from a soil sample collected on remote Easter Island in the South Pacific. Last year a rapamycin-producing strain, discovered in Fuzhou, China was reported.<sup>4</sup> Hence restricting one's microbial collection to a single isolated strain of each species is counterproductive to the goal of acquiring production capacity for the maximum chemical diversity. Nevertheless, it is still highly desirable to minimize the occurrence of multiple copies of strains, identical in their metabolic capabilities, in such a library.

Initially Phytera, on the rationale that marine-derived microbes have been less explored than their terrestrial counterparts and by virtue of their environment have probably gained different metabolic potentials<sup>5</sup>, has assembled a collection of over 8,000 such microbial isolates. This collection represents 26 expeditions to various environmentally different locations in US territorial waters, including the US Virgin Islands, Alaska and Hawaii. From each expedition the isolated strains were dereplicated with respect to colony morphology, and unique morphotypes from the expedition were deposited in the collection with a sequential culture number. There was no attempt to dereplicate cultures across different collecting expeditions in keeping with the principle of same species metabolite differences as noted above. Nonetheless there is acknowledgement that this practice may have led to multiple copies of strains with identical metabolic potential within the total collection, and an assessment of the extent of this redundancy was the aim of this study.

Phytera's other main program to generate chemical diversity is based on the use of proprietary technology to obtain maximal chemical expression by plant cells in suspension culture. Plant endophytes are just one of many hurdles that have to be overcome in the process of establishing a suspension culture of a given plant species. Recently, rather than discard endophytic fungi we have begun to add them as a distinct and identified subset of the microbial library. Fungi, endophytic to plants, have been used as a source of secondary metabolites, by other groups<sup>6</sup> and this approach has yielded some unusual and complex chemical structures. Endophytes have been deposited in the collection after dereplication, on the basis of morphology, only across isolates from a particular plant

specimen.

A simple examination of the extent of possible culture duplication within the collection has been made and is reported here.

In order to limit this task, a small, but statistically significant, sample of each major taxonomic group within the collection was examined. The three groups chosen were: marine-derived fungi, marine-derived actinomycetes, and fungal plant endophytes, represented within the total collection by 3852, 2470 and 776 strains respectively.

### Materials and Methods

Selection of cultures to be examined was made without bias by capitalizing on the history of assigning a sequential culture number to each culture as it was added to the collection. By selecting cultures at a set interval throughout the collection, it was assured that most expeditions would be represented in the selection and that the collection would be examined as stringently as possible. Each expedition resulted in a series of sequential culture numbers in the collection as acquisition of cultures from one was complete before the next expedition was undertaken. Redundancy was most likely to occur across expeditions as dereplication had already been carried out within each expedition. For marine-derived fungi every 75th culture was selected resulting in a sample of 51 cultures and for the marine-derived actinomycetes, the interval was 50 and the sample size 49 cultures. In the case of the endophytes, there is no equivalent to expeditions and cultures isolated from different plant samples often overlapped in number sequence. Approximately every 17th endophytic culture was selected for a sample size of 50.

In the first phase, all cultures in a given group were grown on a single agar medium and compared on the basis of gross morphology. The media were: Difco potato-dextrose agar made with sterile sea water for marine-derived fungi, Difco potato-dextrose agar made with distilled water for endophytic fungi and ATCC Number 5<sup>7</sup> media made with sterile sea water for marine-derived actinomycetes.

In a second phase, those cultures, which were not clearly singular were regrown on three agar media and again compared morphologically. Actinomycetes were grown on ATCC Number 172<sup>8</sup>, Starch casein nitrate<sup>9</sup> and Difco ISP Medium 4. Fungi were grown on Difco agars of corn meal, yeast malt extract, and Sabouraud maltose.

Phase 3 consisted of an examination of the metabolic profiles of morphologically undifferentiated strains. The

organisms were inoculated into 250 ml wide-mouthed Erlenmeyer flasks containing 40 ml of autoclaved seed media<sup>10)</sup> and then allowed to grow on a shaker for three days whereupon 1 ml was transferred to a similar flask of production media and grown for 6 days (actinomycetes) or 8 days (fungi). At harvest the whole broths were extracted with an equal volume of ethyl acetate, the organic layer was separated and dried and the residues obtained were digested in methanol (1 ml) and spotted on Merck SiO<sub>2</sub> TLC plates. The plates were developed with a mixture of dichloromethane (10 vols), methanol (1 vol) and acetic acid (0.01 vols), dried and visualized under UV light. They were then sprayed with sulfuric acid (1.8N) in ethanol, heated with a hot air gun to dryness and further sprayed with a 1% solution of vanillin in ethanol and again heated with a hot air gun. The colors and R<sub>f</sub>s of spots were noted.

In a fourth phase, cultures which were identical in their extractable metabolite patterns were examined for their comparative growth characteristics under a wide variety of nutritional conditions. These cultures were inoculated into all 96 wells of Biolog SF-P MicroPlates<sup>TM</sup> and incubated at 26°C according to the directions provided by Biolog, Inc.<sup>11)</sup> Each well of these plates contains a different individual nutrient medium and these media are designed to explore the primary metabolic capabilities of the inoculants.

At the completion of each of the first three phases (no further distinctions emerged from the fourth phase), the number of unique and apparently duplicate cultures was recorded. The data from these subsets were used to extrapolate the distribution of culture uniqueness across the entire collection.

### Results and Discussion

The aim of the study was to estimate the duplication of metabolically identical strains in the collection. It is reasoned that if two strains produce clearly different morphotypes, this represents *a priori* evidence that the two strains have different metabolic capabilities. Thus, if a strain was unique in its morphology within the sample group at a given Phase of the study, it was not a candidate for assessment in subsequent Phases. At the end of Phase 1, cultures of 32 of 49 marine-derived actinomycetes were morphologically unique and the rest comprised seven pairs and one triplet. At this stage only 27 cultures of 51 marine fungi were morphologically unique. There were nine pairs and two triplets in this group. The 50 endophytic fungal cultures were made up of 33 morphological uniques, seven

pairs and one triplet.

Plating onto multiple media in Phase 2 resolved only two of the pairs among the actinomycetes and three of the pairs from the endophytic fungi. This approach was much more rewarding with the marine-derived fungi where seven of the pairs and both of the triplets were completely resolved leaving only two pairs as morphologically unresolved.

The metabolite profiling was even more distinguishing. TLC of the extracts from the five remaining morphological duplicates and one triplet of marine-derived actinomycetes indicated that all of these extracts contained different metabolites except for two apparently identical strains coming from two of the cultures of the triplet. Similarly TLC analysis of the extracts from the four morphology pairs and one triplet of endophytic fungi distinguished between all of the pairs, however all three extracts from the triplet appeared identical. The four marine fungi (two pairs) that were morphologically similar all gave different extracts by TLC. These results are summarized in Table 1.

When the two cultures of the remaining undifferentiated pair of marine-derived actinomycetes were each allowed to grow on a Biolog SF-P MicroPlate<sup>TM</sup> and examined after 24 hours and 4 days, the growth patterns, colors and apparent cell density in identical address wells for the two plates were indistinguishable. A similar examination of all three of the apparently identical endophytic fungi also failed to produce any visible differences between these three cultures. The Biolog SF-P MicroPlates<sup>TM</sup> thus added no new information as to the duplication, or lack of it, of cultures within these selected subsets, but supported conclusions already drawn. The first three phases of our analysis had each contributed significant incremental evidence of diversity and this is shown in Figure 1.

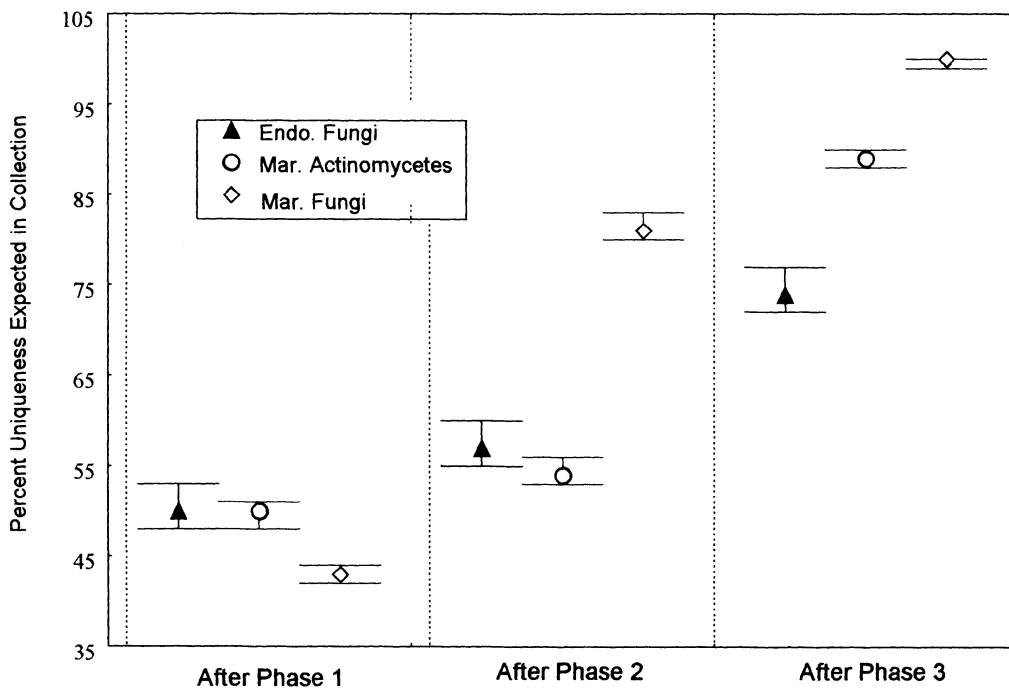
An analysis of the frequency distribution of plates with similar morphologies and metabolic profiles revealed that sample uniqueness is not significantly different from a Poisson distribution, as might reasonably be expected. The expected degree of culture uniqueness, which may be inferred across the entire collection, at each stage of the analyses, for each microbial type, and the 95% confidence limits on these is presented in Table 2.

Of the 7% expected non-unique cultures in the collection, Poisson predictions are that 3% of these would fall into the doublet category. This means that the total collection can be expected to have at least another 3% different cultures for a total expected 96% different cultures, and implying that our collection and isolation techniques and paradigm had resulted in only an estimated 4% level of metabolic potential redundancy. There would theoretically be additional cultures, different from the

Table 1. Extent of culture singularity detected at each stage of the study.

Sample size/ Total collection	Morphology on a single medium	Morphology on 3 additional media	Metabolic profile
<b>Marine derived actinomycetes</b>			
	32 individuals	4 individuals	11 individuals
49 / 2476	7 pairs	5 pairs	1 pair
	1 triplet	1 triplet	
<b>Marine derived fungi</b>			
	27 individuals	20 individuals	4 individuals
51 / 3852	9 pairs	2 pairs	
	2 triplets		
<b>Endophytic fungi</b>			
	33 individuals	6 individuals	8 individuals
50 / 776	7 pairs	4 pairs	
	1 triplet	1 triplet	1 triplet

Fig. 1. Comparison of expected percent uniqueness in the collection after three diagnostic phases.



Bars represent 95% confidence limits around the expected percentage unique value.

Table 2. Expected number of unique cultures per taxonomic group (95% confidence limits).

	After Phase 1	After Phase 2	After Phase 3
<b>Marine-derived</b>			
Actinomycetes	1229	1341	2209
(N=2,476)	(1192 to 1267)	(1304 to 1381)	(2181 to 2239)
<b>Marine-derived</b>			
Fungi	1648	3136	3840
(N= 3852)	(1604 to 1693)	(3092 to 3180)	(3834 to 3847)
Endophytic Fungi	388	443	575
(N=776)	(368 to 410)	(422 to 465)	(554 to 596)

uniques and doublets, among the higher multiplets, but as the expected frequency of higher multiplets is only approximately 1% of the collection, no attempt was made to calculate this.

Undoubtedly the analysis could be made more rigorous by a similar comparison of the entire library. However, such an analysis would almost certainly necessitate much more sophisticated analytical methods. Given the results obtained here, it seems reasonable to suggest that additional analytical methods are more likely to resolve doublets and multiplets down to metabolically unique cultures. Hence it seems reasonable to assume that any errors inherent in the study are more likely to be minor errors rather than major errors. (That is this study is more likely to have underestimated the overall diversity in Phytera's microbial collection, than have over-estimated it.)

### References

- 1) OKAMI, Y. & HAMA O UMEZAWA: The man and his dream. pp. 1~24. *In* Novel Microbial Products for Medicine and Agriculture, DEMA IN, A. L., G. A. SOMKUTI, J. C. HUNTER-CEVERA & H. W. ROSSMORE, *Eds.*, Elsevier, Amsterdam, New York, 1989
- 2) Dictionary of Natural Products on CD-ROM Version 10:1, August 2001, Chapman & Hall/CRC; ISBN 0-412-49150
- 3) VÉZINA, C.; A. KUDELSKI & S. N. SEGHAL: Rapamycin (AY22,989), a new antifungal antibiotic. I. Taxonomy of the producing streptomycete and isolation of the active principle. *J. Antibiotics* 28: 721~726, 1975
- 4) CHENG, Y. H.; J. HUANG, W. L. LIN & A. L. DEMA IN: Mutagenesis of the rapamycin producer *Streptomyces hygroscopicus* FC904. *J. Antibiotics* 54: 967~972, 2001
- 5) This presumption is supported by a comparative study of actinomycetes collected off the New Jersey shores with those isolated from in-state terrestrial samples. The study by A. ANDERSON *et al.* was reported at the SIM Annual Meeting in San Diego on July 27th, 2000, Paper No. S173
- 6) Notably among them were the Merck Sharp and Dohme Research Laboratories, MYCOsearch (most recently a subsidiary of OSI Pharmaceuticals) and the laboratory of Professor GARY STROBEL at Montana State University
- 7) Glucose (10.0 g), Difco tryptone (2.0 g), Difco beef extract (1.0 g), Difco yeast extract (1.0 g) agar (15.0 g)/liter sterile sea water with pH subsequently adjusted to 7.2
- 8) Soluble starch (20 g), glucose (10 g), Difco soytone (5.0 g), Difco yeast extract (5.0 g), CaCO<sub>3</sub> (1.0 g)/liter sterile sea water
- 9) Soluble starch (10.0 g), KNO<sub>3</sub> (2.0 g), Difco casamino acids (0.3 g), agar (20.0 g)/liter sterile sea water
- 10) Seed media broth was ATCC Number 5 without the agar
- 11) Test Panels for Actinomycetes & Fungi and other sporulating and filamentous microorganisms. SF-N MicroPlate™, SF-P MicroPlate™. a pamphlet published by Biolog, Inc. (www.biolog.com)