ELSEVIER

Journal of Biotechnology 99 (2002) 187–198

JOURNAL OF Biotechnology

www.elsevier.com/locate/jbiotec

# Microbial technologies for the discovery of novel bioactive metabolites

Stefano Donadio [a,*], Paolo Monciardini [a], Rosa Alduina [b], Paola Mazza [a,1], Claudia Chiocchini [a,2], Linda Cavaletti [a], Margherita Sosio [a], Anna Maria Puglia [b]

[a] *Biosearch Italia, via R. Lepetit 34, 21040 Gerenzano, Italy*
[b] *Department of Cellular and Developmental Biology, University of Palermo, Palermo, Italy*

Received 15 October 2001; received in revised form 17 May 2002; accepted 24 May 2002

**Abstract**

Soil microbes represent an important source of biologically active compounds. These molecules present original and unexpected structure and are selective inhibitors of their molecular targets. At Biosearch Italia, discovery of new bioactive molecules is mostly carried out through the exploitation of a proprietary strain collection of over 50 000 strains, mostly unusual genera of actinomycetes and uncommon filamentous fungi. A critical element in a drug discovery based on microbial extracts is the isolation of unexploited groups of microorganisms that are at the same time good producers of secondary metabolites. Molecular genetics can assist in these efforts. We will review the development and application of molecular methods for the detection of uncommon genera of actinomycetes in soil DNA and for the rapid dereplication of actinomycete isolates. The results indicate a substantial presence in many soils of the uncommon genera and a large diversity of isolated actinomycetes. However, while uncommon actinomycete strains may provide an increased chance of yielding novel structures, their genetics and physiology are poorly understood. To speed up their manipulation, we have developed vectors capable of stably maintaining large segments of actinomycete DNA in *Escherichia coli* and of integrating site specifically in the *Streptomyces* genome. These vectors are suitable for the reconstruction of gene clusters from smaller segment of cloned DNA, the preparation of large-insert libraries from unusual actinomycete strains and the construction of environmental libraries.


*Keywords:* Soil microbes; Secondary metabolites; Actinomycete

* Corresponding author
*E-mail address:* sdonadio@biosearch.it (S. Donadio).
[1] Present address: Eberhard-Karls Universitaet Tuebingen, Biologisches Institut, Auf def Morgenstelle 28, 72076 Tuebingen, Germany.
[2] Present address: Biochemie-Fachberech Chemie, Philipps-University of Marburg, Hans-Meerwein Str., D-35032 Marburg, Germany.

0168-1656/02/$ - see front matter © 2002 Published by Elsevier Science B.V.
PII: S0168-1656(02)00209-2

## 1. Introduction

The search for new substances of pharmaceutical or agricultural importance can be conducted through screening large collections of diverse chemical entities (generally defined as a library), employing assays designed to detect modulators of pharmacologically or agriculturally relevant targets. Nowadays, screening is mostly conducted in an automated fashion and is referred to as high throughput screening (HTS). HTS is finding increased applications thanks to the availability of large libraries and to the large number of targets that are being identified through genome sequencing. General considerations about HTS can be found in a recent overview (Carrano and Donadio, 1999). A critical role in identifying novel bioactive substances lies in the quality of the library and in the assays employed for screening it. Some considerations about assays and the appropriate libraries to use in the antibacterial field, can be found in another paper in this issue (Donadio et al., 2002). In general terms, the ideal library should consist of as many as possible different chemical entities, without any bias in the type of chemical classes, synthetic routes or biological activities. In addition, it should contain a relevant fraction of unique compounds, unlikely to be found in other libraries. It should also consist of pure substances of known structure and specified concentration. Finally, increased amounts of interesting compounds found through screening should be readily obtained for further tests. Unfortunately, no single library can satisfy all these criteria.

Microbial secondary metabolites represent a large source of compounds endowed with ingenious structures and potent biological activities. Many of the products currently used for human or animal therapy, in animal husbandry and in agriculture are produced by microbial fermentation, or are derived from chemical modification of a microbial product. These products have been obtained after a few decades of intensive screening involving probably millions of microorganisms. These past successes make discovering new bioactive metabolites from microbial sources harder than ever, since thousands of compounds are described in the literature. However, different strains generally produce different compounds. Thus, new bioactive metabolites continue to be identified from microbial sources, thanks to the large variety of existing strains. However, not all microorganisms are equally capable of producing secondary metabolites. In fact, this capability is at the moment restricted to a few groups of bacterial or eukaryotic microbes. In particular, the ability to produce a large number of chemically different secondary metabolites is associated mostly with the filamentous actinomycetes, the myxobacteria, the pseudomonads and the cyanobacteria within the prokaryotic world, and mostly to the filamentous fungi for the eukaryotic microbes (G. Toppo, personal communication).

We believe that novel antibiotics and other bioactive secondary metabolites can still be discovered from microbial sources. In our opinion, the probability of finding novel bioactive compounds depends on a series of critical factors. On the one hand, there is the number of strains screened and their degree of diversity; on the other hand, their uniqueness and their potential to produce secondary metabolites. These last two criteria are extremely important, since intensively screened microbes are less likely to yield novel metabolites than unexploited groups. These critical factors must all be considered when embarking on a screening program for bioactive metabolites from microorganisms.

## 2. The strain collection and microbial extract bank at Biosearch Italia

Our assumption is that novel metabolites can be discovered by screening unusual or difficult to isolate strains belonging to the two most prolific groups of producers, the filamentous actinomycetes and the filamentous fungi. This assumption rests on the increased likelihood that they have not been intensively screened in the past and on their promise to be potentially capable of producing secondary metabolites, as demonstrated for some unusual actinomycetes (Sosio et al., 2000a). With this assumption in mind, Biosearch Italia has assembled a large library of processed fermentation broths ( $>150\,000$, at the time of this writing)

obtained from a proprietary collection of over 50 000 strains (Lazzarini et al., 2000). These processed fermentation broths are stored in a ready-to-screen format as dried, frozen aliquots in microtiter plates. Over 90% of them were obtained by fermenting actinomycetes other than *Streptomyces*, and filamentous fungi. Consequently, the strain collection consists mainly of unusual genera of actinomycetes (e.g. the easy-to-isolate streptomycetes account for about 1% of the actinomycetes isolated over the last 4 years; Fig. 1) and of slow-growing and endophytic fungi, with little emphasis on penicilli and aspergilli (e.g. about 8% of eukaryotic microbes belong to these two genera; data not shown).

The strain collection has been built through the application of isolation methods to a large number of different sources, which includes soil samples, plant materials, sediments, marine macroorganisms, etc. It is continuously enlarged by the addition of newly isolated strains from newly obtained specimens and a historical collection of $>10\,000$ dried soil samples. For isolating such a large number of diverse microbes, tens of different methods are applied, according to the source and the target group of microbes. The objective of an isolation program consists in obtaining, with the minimum effort, as large a number as possible of microbes that are unusual and difficult to isolate. We have developed a series of methods that can assist in the isolation of unusual groups of microorganisms. We describe here approaches that can be applied for the detection, dereplication and manipulation of unusual groups of actinomycetes.

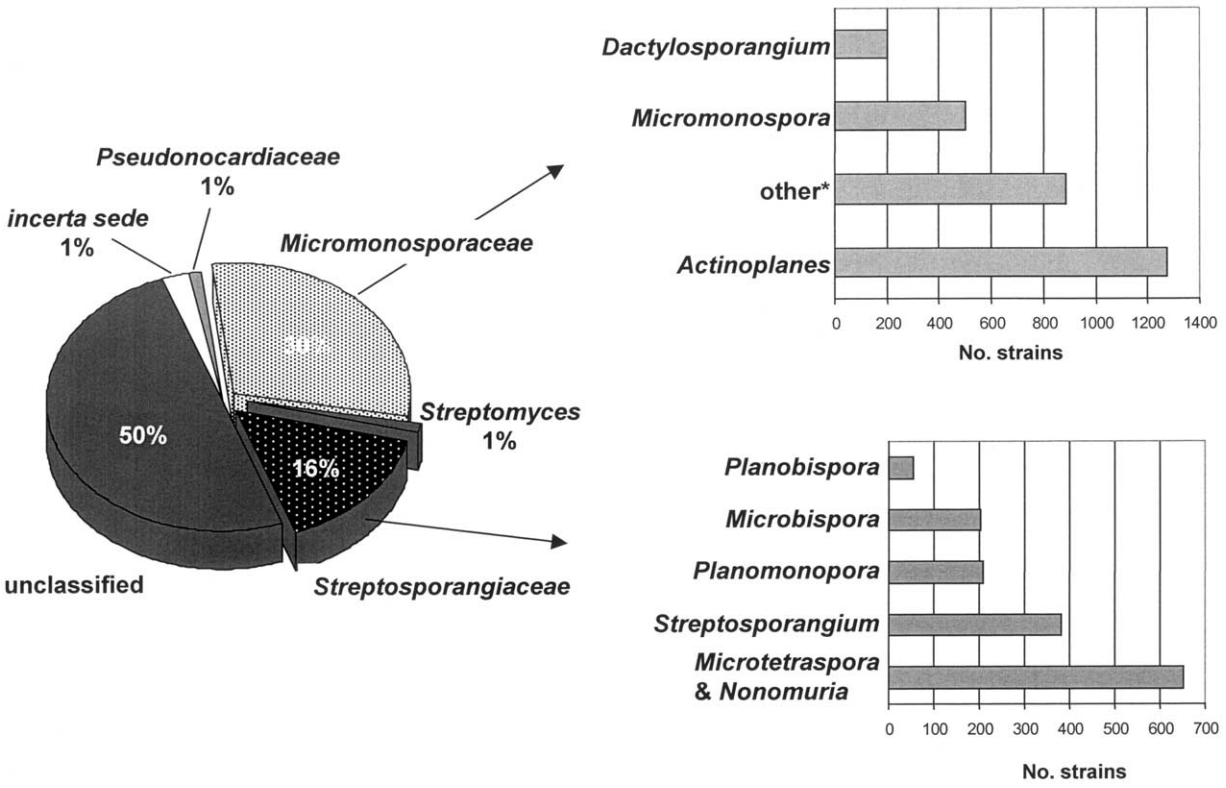



Fig. 1. Isolation effort in actinomycetes. The pie illustrates the distribution in major families of about 9300 strains isolated in the period 1996–2000. The portion unclassified refers to strains for which no classification was made. The enlargements show the actual number of strains belonging to various genera in the families *Micromonosporaceae* and *Streptosporangiaceae*.

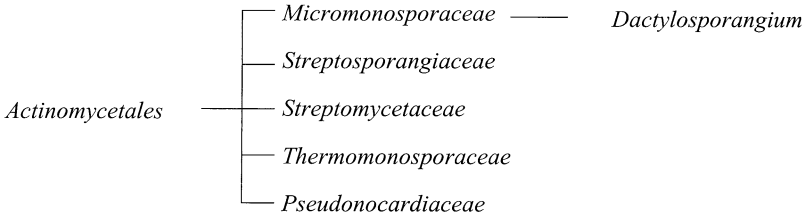
## 3. Random and directed isolation methods

Isolation methods are usually applied to soil samples or other specimens in a random way, without an a priori knowledge of the microbial composition of the source under investigation. Consequently, the inability to isolate a certain group of actinomycetes from a given soil, for example, can be due either to their absence from that soil, or to the use of inappropriate methods to recover them, or to their outcompetition by other, fast growing bacteria. It would be highly desirable to know in advance what groups of microorganisms are present in a natural source, their relative abundance and, possibly, their diversity, so that an appropriate effort can be devoted to that source.

The number of bacteria present in a soil sample can be extremely large, up to $10^9$ $g^{-1}$ of soil (Steffan et al., 1988; Weinbauer et al., 1998). Based on the use of cultivation-independent methods, it has also been estimated that over 99% of the strains present in a given soil do not find a match to described bacteria, suggesting that the large majority of bacterial species is uncultured and may be unculturable (Amann et al., 1995). These studies have mostly employed amplification methods based on 16S RNA. Indeed, this sequence is universally present in bacteria, highly conserved and a good phylogenetic marker (Amann et al., 1995; Woese, 1987). While the existence of invariant sequences allows 16S amplification from any bacterium, it is also possible to selectively amplify a 16S segment from a desired group through the use of specific primers (Knight, 2000). We reasoned that a similar approach could be applied to actinomycetes. In contrast to most of the work reported so far for these bacteria, which described conditions for genus- or species-specific amplification (Mehling et al., 1995; Moròn et al., 1999), we decided to develop methods for the amplification of related groups of genera, or families. This has been possible thanks to the extensive reclassification and phylogenetic grouping of actinomycetes, which has occurred during the last decade (Stackebrandt et al., 1997). Most genera of filamentous actinomycetes are described within families, and many of these families include a few unusual genera, suitable targets for our isolation efforts. In addition, this strategy leaves open the possibility of detecting 16S sequences that could not be ascribed to known genera.

Through the alignment of 16S sequences from desired families of actinomycetes, either present in public databases or ad hoc generated from the strains in our collection, we set up PCR conditions for selective amplification of strains belonging to the order *Actinomycetales*, to the families *Micromonosporaceae*, *Streptosporangiaceae*, *Thermomonosporaceae* and *Streptomycetaceae*, and to the genus *Dactylosporangium* (Monciardini et al., 2002). Three of the families include unusual genera of actinomycetes, as is for example *Dactylosporangium* in the family *Micromonosporaceae*. The family *Streptomycetaceae* includes mostly the relatively easy to isolate streptomycetes. The methods were also verified after sequencing the cloned PCR products obtained from a mixed population of strains, such as total soil DNA. Because the 16S sequence is highly predictive of the genus of the strain from which it originated, we could phylogenetically group the sequences from the soil clones with control actinomycete 16S sequences, thus determining the number of clones whose sequences grouped with the corresponding family or genus (Fig. 2). We can conclude that all the primer pairs tested detect between 1 and 5 pg DNA, corresponding to about 100–500 8-Mb genomes containing one copy of the 16S gene. In addition, all amplifications from soil DNA are $>90\%$-specific, as judged from the fraction of soil clones that grouped according to predictions.

Fig. 3 reports the results obtained after sequencing the products obtained from three soil samples with the primer specific for *Micromonosporaceae* (Monciardini et al., 2002). Over 75% of the sequences obtained are different, suggesting that each soil sample contains different *Micromonosporaceae*. While a certain correlation between soil and sequence grouping is evident, this is not absolute, suggesting that different genera are present in each soil. In addition, the *Dactylosporangium*-specific primers revealed a positive signal from one soil. The application of appropriate




| Target | Detection limit | Specificity |
|---|---|---|
| **Bacteria** | **5 pg** | **N.A.** |
| **Actinomycetes** | **0.5-5 pg** | **>97%** |
| ***Micromonosporaceae*** | **<0.5 pg** | **>97%** |
| ***Dactylosporangium*** | | **>96%** |
| ***Streptomyces*** | **<0.7 pg** | **>92%** |
| ***Streptosporangiaceae*** | **1 pg** | **>96%** |

Fig. 2. Selective amplification methods. The upper part shows a simplified dendrogram showing only the groups that were specifically amplified. The detection limit and specificity for each method are reported in the table in the bottom part (Monciardini et al., 2002).

isolation methods to this same soil allowed the recovery of *Dactylosporangium* isolates (L.C., unpublished results). This example illustrates how information on the microbial composition can be useful in directing isolation of desired genera.

## 4. Strain dereplication

One important factor in increasing the probability of finding novel metabolites is through the use of different strains. While each actinomycete strain has probably the genetic potential for

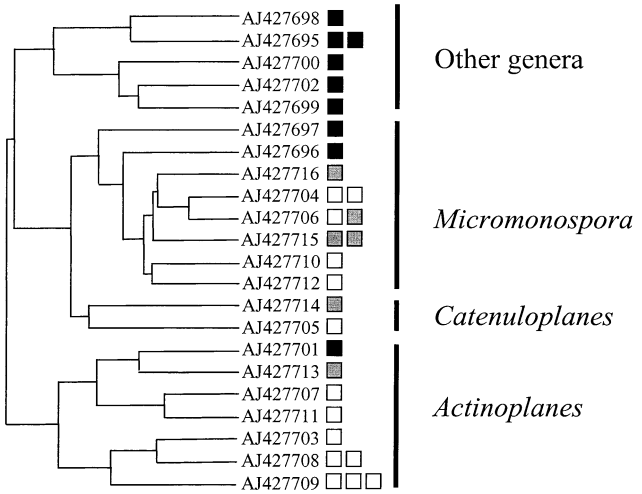



Fig. 3. Amplification of *Micromonosporaceae* sequences from soil DNA. A dendrogram of 16S sequences obtained from 29 cloned PCR products from three different soils. Each sequence is indicated by its accession number; black, gray and empty squares indicate sequences obtained from soil samples taken from Gerenzano (Italy), Siena (Italy) and Niger, respectively. The number of squares corresponds to the number of clones with identical sequences. Data are from Monciardini et al. (2002).

producing 10–20 secondary metabolite (Sosio et al., 2000a; Omura et al., 2001; Bentley et al., 2002), the probability of obtaining different metabolites is substantially higher by fermenting different strains than by repeated fermentations of the same strain. Consequently, an important step is strain dereplication, i.e. a rapid analysis of all the strains isolated from the same or similar sources to discard apparently identical isolates. This dereplication is usually accomplished by comparing the morphological characteristics of the isolated colonies, such as color, shape, consistency, etc. While this analysis can be extremely rapid and effective in many cases, it mostly relies on subjective parameters and is highly empirical. Furthermore, when all isolates are homogeneous for one of these characters, for example color, morphological differentiation becomes less predictive of the actual diversity of the isolates. An alternative approach may consist in dereplicating strains through the determination of a fingerprint for each strain. Several approaches have been described that make use of infrared (Naumann et al., 1991; Helm et al., 1991) or mass (Goodacre and Kell, 1996) spectroscopy to generate a chemical signature for each strain, or of DNA fingerprints (Gürtler and Stanisich, 1996; Doignon-Bourcier et al., 2000).

We have devised a molecular dereplication method, based on PCR-RFLP of 23S rDNA and of the 16S–23S intergenic region, that can be directly applied to colonies on an agar plate. The resulting fingerprints are compared among all the isolates. From a total of 4000 isolates obtained through a selective isolation method from a single Italian soil, consisting mostly of *Actinoplanes* and related genera, 182 strains were evaluated by PCR-RFLP (Fig. 4). Analysis of the fingerprints allowed the classification of the 182 strains into 99 groups, each giving a unique profile. Thus, at least for this group of actinomycetes and for one soil sample, genetically different strains can be obtained in large numbers from a single source. It remains to be seen whether this genetic diversity based on ribosomal sequences correlates to a genetic diversity for secondary metabolite formation.

## 5. Strain manipulation

The availability of a large number of diverse actinomycete strains is a prerequisite for a successful screening program based on microbial products. While strains belonging to unusual genera of actinomycetes may provide a higher probability of finding novel bioactive metabolites than *Streptomyces* strains, the mere fact that they are unusual implies little knowledge about their physiology and genetics. Consequently, the powerful tools that have been developed and can now be successfully applied for the genetic manipulation of *Streptomyces* (Kieser et al., 2000), are of limited value when dealing with other filamentous actinomycetes. This limitation results in the inability or inefficiency of genetically altering these strains for the production of novel metabolites by pathway manipulation or for an increase in metabolite yield. However, this limitation could be at least in part ameliorated if the genetic information for secondary metabolite production could be transferred from the original producer strain to a *Streptomyces* host of well-defined genetics and physiology, where these genes would be efficiently expressed. This transfer is technically conceivable, since the formation of secondary metabolites in actinomycetes is invariably controlled by gene clusters, discrete genomic segments carrying all the information required for synthesis of and, if necessary, resistance to the metabolite. The possibility of using surrogate hosts for studying and possibly exploiting the genetics of hard- or impossible-to manipulate microorganisms, is arising considerable interest in microbiology, with the large number of uncharacterized groups of bacteria that exist in the environment (Rondon et al., 2000; Beja et al., 2000a). Examples of heterologous expression of entire clusters in model actinomycetes have been reported (Kwon et al., 2001; Piel et al., 2000; Tang et al., 2000).

Toward this objective, we have developed bacterial artificial chromosomes that can be shuttled between an *Escherichia coli* and a *Streptomyces* host. These vectors have been given the generic name of ESAC, for *E. coli-Streptomyces* artificial chromosome. A typical vector is pPAC-S2, illustrated in Fig. 5. It is based on the PAC-series of

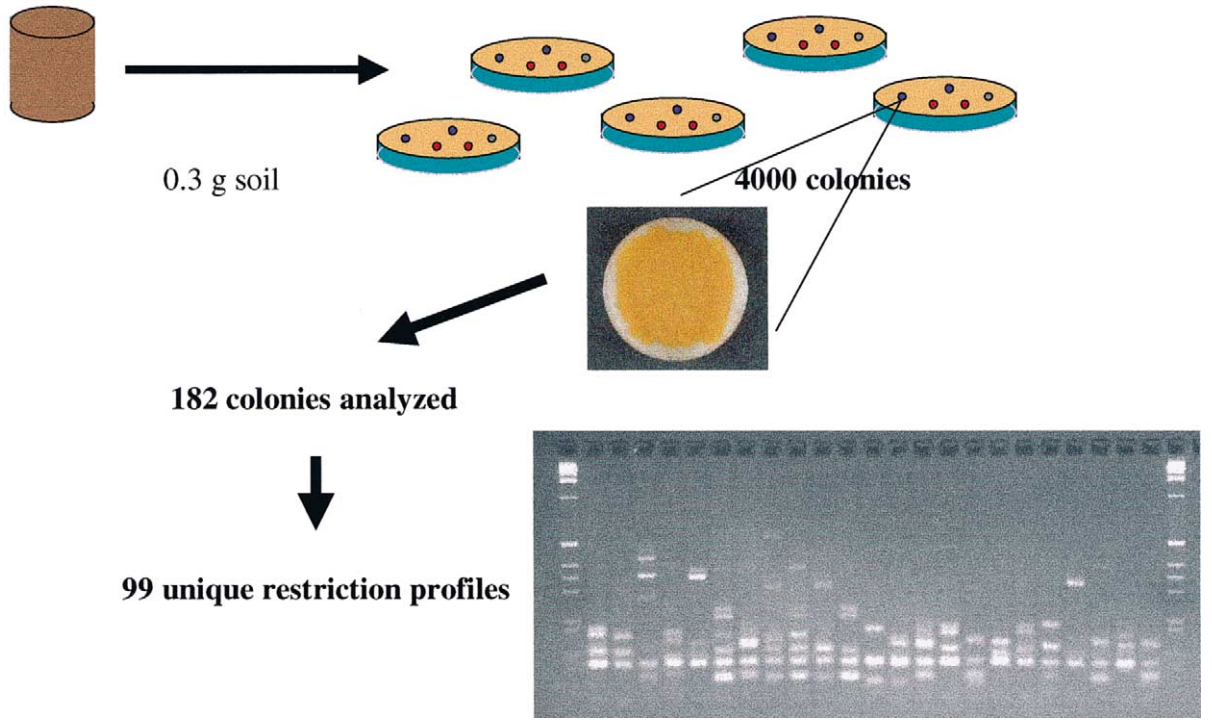



Fig. 4. Dereplication of actinoplanes isolates. A schematic of the 4000 colonies isolated, the typical aspect of many of them, and a summary of the results obtained after PCR-RFLP of 182 colonies. The inset shows the fingerprints obtained from 20 independent isolates after amplification of the 16S–23S spacer followed by *Hin*II digestion. The last lanes in the inset contain molecular weight markers.

bacterial artificial chromosomes (Ioannou et al., 1994) and can drive site-specific integration at the *attB* site of a *Streptomyces* genome by virtue of the phage ϕC31 *attP-int* system. In *Streptomyces*, it confers resistance to thiostrepton (Sosio et al., 2000b). An important characteristic is their ability to harbor large DNA fragments, in the order of 100-kb, both in an *E. coli* host, as do other BAC vectors, and in *Streptomyces lividans*, a model streptomycete. Indeed, we have been able to stably maintain ESACs carrying up to 120-kb of heterologous DNA into the *S. lividans* chromosome (Sosio et al., 2000b). This feature is extremely important, since the genetic information required for the synthesis of many secondary metabolites may in some cases approach or exceed the 100-kb size (August et al., 1998; Du et al., 2000).

In order to be versatile, these vectors should be amenable to different applications. For example, one should be able to make a large-insert library from a strain of interest, in order to isolate one or more large gene clusters. Alternatively, an available cluster, cloned as a set of smaller fragments, can be reconstructed as one piece in these vectors. Finally, one may wish to make a library from a mixture of strains, either generated in the lab or present in natural sources. These possible applications will be briefly summarized here.

The construction of a large-insert library from a donor strain can be made following substantially the published procedures for the construction of other BAC libraries (Birren and Lai, 1993). Particular care must be used in the generation of good quality vector and insert DNA, since there is no size-selection for clones and a sharp decrease in electroporation efficiency occurs with increasing DNA size (Sheng et al., 1995). Large molecular weight DNA is handled with extreme care, so it is normally obtained, digested and fractionated while embedded in an agarose gel. The vector also needs

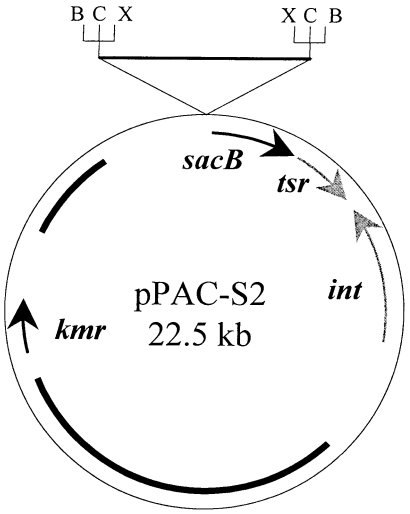



Fig. 5. A typical ESAC vector. The vector pPAC-S2 is depicted, with the relevant genes *sac*B (conferring resistance to sucrose when not expressed), *tsr* (conferring resistance to thiostrepton in *Streptomyces*), *int* (driving *attP* × *attB* site-specific integration into the chromosome) and *aph* (conferring kanamycin resistance in *E. coli*) indicated by arrows. The thick bars represent the DNA segments derived from bacteriophage P1. Unique restriction sites suitable for cloning are indicated as *B* for *Bam*HI, *C* for *Sca*I and *X* for *Xba*I.

to be of very good quality and sophisticated procedures for its preparation have been described (Osoegawa et al., 1998). The ligation mixture of vector and insert is then used to electroporate *E. coli* cells, which are selected for kanamycin and sucrose resistance. The latter phenotype results from the presence of insert DNA. A representative number of these colonies is analyzed for insert size after digestion with *DraI*, which in most cases releases the 70%-GC actinomycete DNA insert as a single fragment. Using this approach, we have been able to clone large DNA fragments from the actinomycetes *Streptomyces coelicolor*, *Planobispora rosea* and *Nonomuria* sp. The largest insert so far cloned has been a 200-kb *S. coelicolor* DNA fragment. We do not currently know whether this size represents an upper limit for high GC DNA, as suggested (Brosch et al., 1998; Gordon et al., 1999), or if larger inserts of actinomycete DNA will eventually be cloned with improved methodologies.

We have also explored the possibility of reconstructing an available cluster starting from smaller clones. Indeed, most antibiotic gene clusters have been obtained without the direct cloning in artificial chromosomes, usually from cosmid libraries. Thus, a 100-kb genomic segment would be covered by 3–4 overlapping cosmids. The reconstruction of such a segment as a single clone through in vitro DNA manipulation is technically challenging. In fact, the probability of finding suitably located restriction sites decreases with the size of the clusters. Several papers have recently appeared describing in vivo methods for manipulation of large DNA segments (Yang et al., 1997; Muyrers et al., 1999; Narayanan et al., 1999; Yu et al., 2000). We have developed a systematic in vivo approach to the reconstruction of a large gene cluster into a BAC, employing the ESAC vectors. This procedure makes use of iterative rounds of homologous recombination in *E. coli*. In each round, desired DNA fragment is transferred from one replicon into another through the formation and resolution of a plasmid cointegrate. Cointegrate formation is facilitated by the use of a temperature-sensitive (*ts*) replicon (Hamilton et al., 1989). The desired DNA fragments are transferred from the cosmid into the *ts* replicon, and from there into the ESAC vector. A typical reconstruction sequel is shown in Fig. 6. This procedure has been successfully used for the faithful reconstruction of a 90-kb segment from the *P. rosea* genome, which was originally encompassed by three separate cosmids (Sosio et al., 2001). This strategy for cluster assembly may be desirable in all those cases where a genomic segment of interest has been isolated and characterized, but the DNA of the original actinomycete strain is readily degraded upon gel-electrophoresis, as if often the case with this group of microorganisms (Ray et al., 1995).

Recently, the fascinating possibility has been discussed of directly accessing the genetic information present in complex sources, without the need to cultivate and isolate the corresponding microbes (Handelsman et al., 1998). The development of efficient methods for the construction of large-insert DNA libraries may provide the tantalizing possibility of obtaining whole pathways cloned

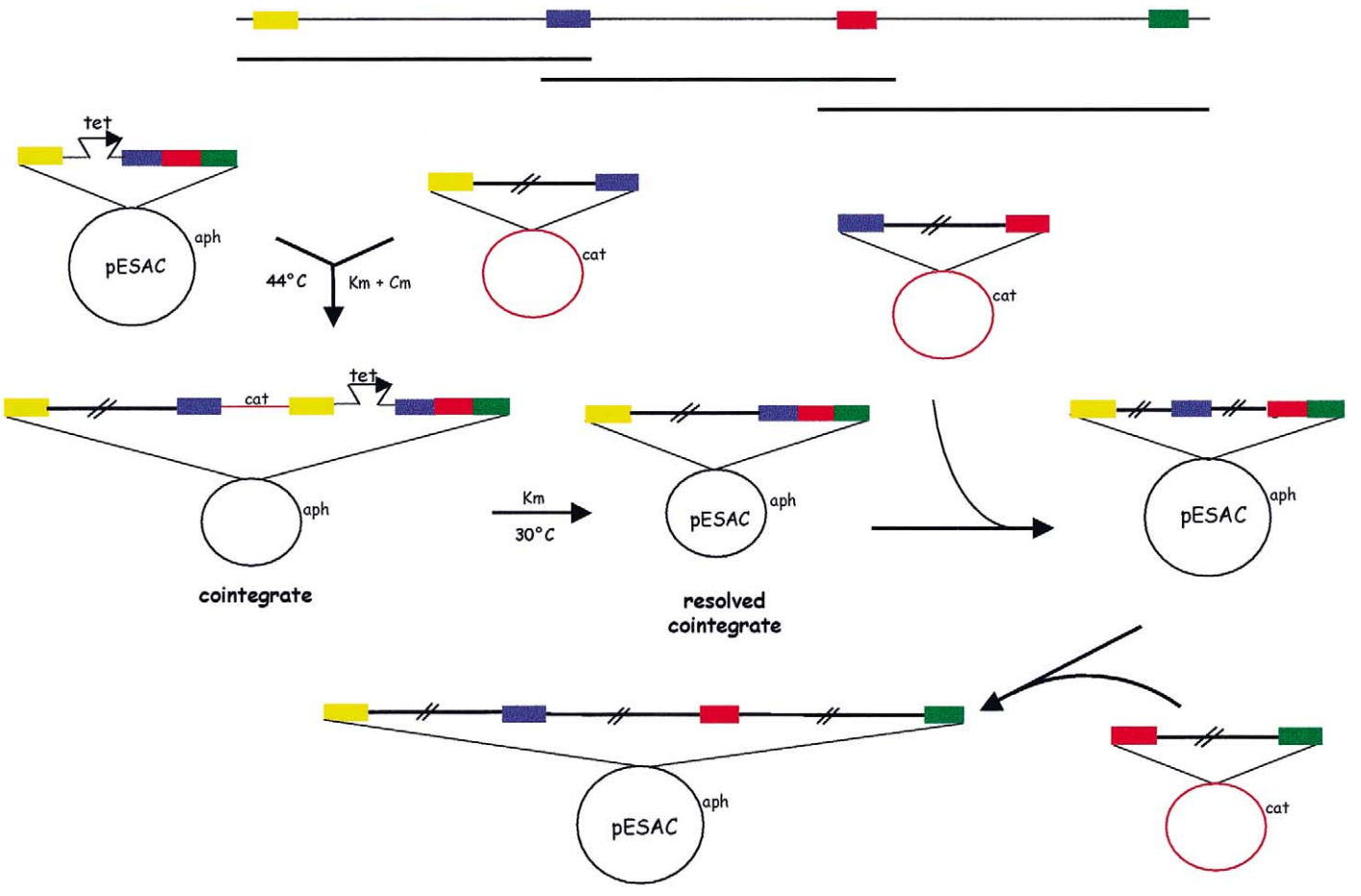



Fig. 6. Assembly of a cluster into ESAC. The figure depicts a strategy for assembly using multiple rounds of homologous recombination in *E. coli* (Sosio et al., 2001). The original cluster, present as three cosmids, is illustrated in the top. Four fragments, depicted in different colors, are necessary for constructing a four-fragment cassette in the ESAC vector. Two fragments denote the two ends of the clusters, while the other two fragments derive from the two regions of overlaps between cosmids. The insert is first transferred from each cosmid into a *ts* replicon (denoted by the red circle), and then into the ESAC vector, leading to its progressive enlargement. The insertion of the *tet* marker (conferring tetracycline resistance) facilitates the recognition of the desired constructs after resolution of the cointegrates. Other abbreviations: *aph* and *cat*, Km- and Cm-resistance gene, respectively; Cm, chloramphemicol; Km, kanamycin.

and expressed in suitable surrogate hosts (Rondon et al., 1999; Beja et al., 2000b). Thus, one may access gene clusters for the production of bioactive metabolites from the large number of uncultured bacteria present in the environment without the need to actually isolate them. Although we do not yet know the genetic potential for producing secondary metabolites of these unknown bacteria, the expression of environmental DNA may be the fastest way to provide an answer to this question.

The ESAC vectors may also be suitable for this application. In fact, *Streptomyces* species seem to have one of the most complex sets of promoters within the bacterial world, and these strains can efficiently transcribe heterologous promoters (Brawner, 1994). It is tempting to utilize the ESAC vectors also for constructing libraries utilizing DNA from environmental sources and expressing these clones in *Streptomyces*. Soil seems to be one of the richest sources of bacteria for this purpose. However, obtaining a sufficient quantity of good quality DNA for the construction of a large-insert size library is no trivial task. Using approaches directed at separating living cells from the bulk of the soil, we can obtain the first 'soil clones', some with $>50$-kb inserts. Remarkably, a good fraction of these clones contain high-GC DNA (C.C. et al., unpublished results). As we learn more about the putative phylogenetic affiliations of the unknown bacteria in soil, better predictions can be made about the possibility of effectively expressing large gene clusters from soil DNA in a surrogate host.

## 6. Conclusions and future perspectives

The discovery of novel bioactive metabolites from microbial sources is a challenging endeavor that can bring substantial rewards when successful. The main challenge stems from the large number of microbial products that have been already discovered. However, microorganisms possess a remarkable imagination in making chemical structures and in deploying intricate machineries for their synthesis, which results in a vast number of structurally original and potent bioactive compounds, difficult or impossible to find in any chemical or combinatorial library.

The application of molecular techniques can greatly improve the efficiency through which desired groups of microorganisms can be detected in environmental samples and isolated from them. Furthermore, even within the actinomycetes, a group extensively exploited in screening, sequences belonging to undescribed groups can be detected using family-specific primers (Monciardini et al., 2002). With reasonable effort, strains belonging to unreported groups of actinomycetes are likely to be isolated and cultivated. We believe that microbiologists skilled in the isolation of unusual microorganisms can effectively synergize with molecular biologists developing sophisticated detection and dereplication techniques for a full exploitation of the chemical diversity produced by unusual groups of actinomycetes. It is also likely that these methodologies can be successfully applied to other groups of microbes producing secondary metabolites. Finally, access to the genetics of previously uncultured or poorly characterized actinomycetes can be substantially improved through the expression of selected genes in a heterologous host. The ESAC technology offers the possibility of adding new tools for an effective manipulation of unusual actinomycetes detected and isolated from the environment.

## Acknowledgements

We are grateful to all our colleagues at Biosearch Italia for stimulating discussions and sharing unpublished data. This work was partially supported by a grant from the Italian Murst, legge 451 (to Biosearch Italia) and by a grant from the Italian CNR, PF Biotecnologie (to A.M.P.).

## References

Amann, R.I., Ludwig, W., Schleifer, K.-H., 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol. Rev. 59, 143–169.

August, P.R., Tang, L., Yoon, Y.J., Ning, S., Muller, R., Yu, T.W., Taylor, M., Hoffmann, D., Kim, C.G., Zhang, Z.H., Hutchinson, C.R., Floss, H., 1998. Biosynthesis of the

ansamycin antibiotic rifamycin: deductions from the molecular analysis of the rif biosynthetic gene cluster of *Amycolatopsis mediterranei* S699. Chem. Biol. 5, 69–79.

Beja, O., Aravind, L., Koonin, E.V., Suzuki, M.T., Hadd, A., Nguyen, L.P., Jovanovich, S.B., Gates, C.M., Feldman, R.A., Spudich, J.L., Spudich, E.N., DeLong, E.F., 2000a. Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. Science 289, 1902–1906.

Beja, O., Suzuki, M.T., Koonin, E.V., Aravind, L., Hadd, A., Nguyen, L.P., Villacorta, R., Amjadi, M., Garrigues, C., Jovanovich, S.B., Feldman, R.A., DeLong, E.F., 2000b. Construction and analysis of bacterial artificial chromosome libraries from a marine microbial assemblage. Environ. Microbiol. 2, 516–529.

Bentley, S.D., Chater, K.F., Cerdeno-Tarraga, A.M., Challis, G.L., Thomson, N.R., James, K.D., Harris, D.E., Quail, M.A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C.W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C.H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabbinowitsch, E., Rajandream, M.A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodwardm, J., Barrell, B.G., Parkhill, J., Hopwood, D.A., 2002. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). Nature 417, 141–147.

Birren, B., Lai, E., 1993. Pulsed Field Gel Electrophoresis: a Practical Guide. Academic Press, New York.

Brawner, M.E., 1994. Advances in heterologous gene expression by *Streptomyces*. Curr. Opin. Biotechnol. 5, 475–481.

Brosch, R., Gordon, S.V., Billault, A., Garnier, T., Eiglemeier, K., Soravito, C., Barrell, B.G., Cole, S.T., 1998. Infect. Immun. 66, 2221–2229.

Carrano, L., Donadio, S., 1999. Key ingredients for efficient high throughput screening. In: Miertus, S., Fassina, G. (Eds.), Combinatorial Chemistry and Technology: Principles, Methods and Applications. Marcel Dekker, New York, pp. 233–250.

Doignon-Bourcier, F., Willems, A., Coopman, R., Laguerre, G., Gills, M., De Lajudie, P., 2000. Genotypic characterization of *Bradyrhizobium* strains nodulating small senegalese legumes by 16S–23S rDNA intergenic gene spacers and amplified fragment length polymorphism fingerprint analyses. Appl. Environ. Microbiol. 66, 3987–3997.

Donadio, S., Carrano, L., Brandi, L., Serina, S., Soffientini, A., Raimondi, E., Montanini, N., Sosio, M., Gualerzi, C.O. 2002. Targets and assays for discovering novel antibacterial agents. J. Biotechnol. 99, 175–185.

Du, L., Sanchez, C., Chen, M., Edwards, D.J., Shen, B., 2000. The biosynthetic gene cluster for the antitumor drug bleomycin from *Streptomyces verticillus* ATCC15003 supporting functional interactions between nonribosomal peptide synthetases and a polyketide synthase. Chem. Biol. 7, 623–642.

Goodacre, R., Kell, D.B., 1996. Pyrolysis mass spectrometry and its applications in biotechnology. Curr. Opin. Biotechnol. 7, 20–28.

Gordon, S.V., Brosch, R., Billault, A., Garnier, T., Eiglmeier, K., Cole, S.T., 1999. Identification of variable regions in the genomes of tubercle bacilli using bacterial artificial chromosome arrays. Mol. Microbiol. 32, 643–655.

Gürtler, V., Stanisich, V.A., 1996. New approaches to typing and identification of bacteria using the 16S–23S rDNA spacer region. Microbiology 142, 3–16.

Hamilton, C.M., Aldea, M., Washburn, B.K., Babitzke, P., Kushner, S.R., 1989. New method for generating deletions and gene replacements in *Escherichia coli*. J. Bacteriol. 171, 4617–4622.

Handelsman, J., Rondon, M.R., Brady, S.F., Clardy, J., Goodman, R.M., 1998. Molecular biological access to the chemistry of unknown microbes: a new frontier for natural products. Chem. Biol. 5, R245–R249.

Helm, D., Labischinski, H., Schallehn, G., Naumann, D., 1991. Classification and identification of bacteria by Fourier transform infrared spectroscopy. J. Gen. Microbiol. 137, 69–79.

Ioannou, P.A., Amemiya, C.T., Garnes, J., Kroisel, P.M., Shizuya, H., Chen, C., Batzer, M.A., de Jong, P.J., 1994. A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. Nature Genet. 6, 84–89.

Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F., Hopwood, D.A., 2000. Practical Streptomyces Genetics. The John Innes Foundation, Norwich, UK.

Knight, I.T., 2000. Molecular genetic methods for detection and identification of viable but nonculturable microorganisms. In: Colwell, R.R., Grimes, D.J. (Eds.), Nonculturable Microorganisms in the Environment. ASM Press, Washington, DC, pp. 77–85.

Kwon, H.J., Smith, W.C., Xiang, L., Shen, B., 2001. Cloning and heterologous expression of the macrotetrolide biosynthetic gene cluster revealed a novel polyketide synthase that lacks an acyl carrier protein. J. Am. Chem. Soc. 123, 3385–3386.

Lazzarini, A., Cavaletti, L., Toppo, G., Marinelli, F., 2000. Rare genera of actinomycetes as potential producers of new antibiotics. Antonie Van Leeuwenhoek. 78, 399–405.

Mehling, A., Wehmeier, U.F., Piepersberg, W., 1995. Nucleotide sequence of streptomycete 16S ribosomal DNA: towards a specific identification system for streptomycetes using PCR. Microbiology 141, 2139–2147.

Monciardini, P., Sosio, M., Cavaletti, L., Chiocchini, C., Donadio, S., 2002. New PCR primers for the selective amplification of 16S rDNA from different groups of actinomycetes. FEMS Microbiol. Ecol., in press.

Moròn, R., Gonzàlez, I., Genilloud, O., 1999. New genus-specific primers for the PCR identification of members of the genera *Pseudonocardia* and *Saccharopolyspora*. Int. J. Syst. Bacteriol. 49, 149–162.

Muyrers, J.P., Zhang, Y., Testa, G., Stewart, A.F., 1999. Rapid modification of bacterial artificial chromosomes by ET-recombination. Nucleic Acids Res. 27, 1555–1557.

Narayanan, K., Williamson, R., Zhang, Y., Stewart, A.F., Ioannou, P.A., 1999. Efficient and precise engineering of a 200 kb beta-globin human/bacterial artificial chromosome in *E. coli* DH10B using an inducible homologous recombination system. Gene Ther. 6, 442–447.

Naumann, D., Helm, D., Labischinski, H., 1991. Microbiological characterizations by FT-IR Spectroscopy. Nature 351, 81–82.

Omura, S., Ikeda, H., Ishikawa, J., Hanamoto, A., Takahashi, C., Shinose, M., Takahashi, Y., Horikawa, H., Nakazawa, H., Osonoe, T., Kikuchi, H., Shiba, T., Sakaki, Y., Hattori, M., 2001. Genome sequence of an industrial microorganism *Streptomyces avermitilis*: deducing the ability of producing secondary metabolites. Proc. Natl Acad. Sci. USA 98, 12215–12220.

Osoegawa, K., Woon, P.Y., Zhao, B., Frenge, E., Tateno, M., Catanese, J.J., de Jong, P.J., 1998. An improved approach for the construction of bacterial artificial chromosome libraries. Genomics 52, 1–8.

Piel, J., Hoang, K., Moore, B.S., 2000. Metabolic diversity encoded by the enterocin biosynthesis gene cluster. J. Am. Chem. Soc. 122, 5415–5416.

Ray, T., Mills, A., Dyson, P., 1995. Tris-dependent oxidative DNA strand scission during electrophoresis. Electrophoresis 16, 888–894.

Rondon, M.R., Raffel, S.J., Goodman, R.M., Handelsman, J., 1999. Towards functional genomics in bacteria: Analysis of expression in *Escherichia coli* from a bacterial artificial chromosome library of *Bacillus cereus*. Proc. Natl Acad. Sci. USA 96, 6451–6455.

Rondon, M.R., August, P.R., Bettermann, P.D., Brady, S.J., Grossman, T.H., Liles, M.R., Loiacono, K.A., Lynch, B.A., MacNeil, I.A., Minor, C., Clardy, J., Handelsman, J., Goodman, R.M., 2000. Cloning of the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. Appl. Environ. Microbiol. 66, 2541–2547.

Sheng, Y., Mancino, V., Birren, B., 1995. Transformation of *Escherichia coli* with large DNA molecules by electroporation. Nucleic Acids Res. 23, 1990–1996.

Sosio, M., Bossi, E., Bianchi, A., Donadio, S., 2000a. Multiple peptide synthetase gene clusters in actinomycetes. Mol. Gen. Genet. 264, 213–221.

Sosio, M., Giusino, F., Cappellano, C., Bossi, E., Puglia, A.M., Donadio, S., 2000b. Artificial chromosomes for antibiotic producing actinomycetes. Nature Biotechnol. 18, 343–345.

Sosio, M., Bossi, E., Donadio, S., 2001. Assembly of large genomic segments in artificial chromosomes by homologous recombination in *Escherichia coli*. Nucleic Acids Res. 37, e37.

Stackebrandt, E., Rainey, F.A., Ward-Rainey, N.L., 1997. Proposal for a new hierarchic classification system, actinobacteria classis nov. Int. J. Syst. Bacteriol. 47, 479–491.

Steffan, R.J., Goskøyr, J., Bej, A.K., Atlas, R.M., 1988. Recovery of DNA from soils and sediments. Appl. Environ. Microbiol. 54, 2908–2915.

Tang, L., Shah, S., Chung, L., Carney, J., Katz, L., Khosla, C., Julien, B., 2000. Cloning and heterologous expression of the epothilone gene cluster. Science 287, 640–642.

Weinbauer, M.G., Beckmann, C., Höfle, M.G., 1998. Utility of green fluorescent nucleic acid dyes and aluminum oxide membrane filters for rapid epifluorescence enumeration of soil and sediment bacteria. Appl. Environ. Micorbiol. 64, 5000–5003.

Woese, C.R., 1987. Bacterial evolution. Microbiol. Rev. 51, 221–271.

Yang, X.W., Model, P., Heintz, N., 1997. Homologous recombination based modification in *Escherichia coli* and germline transmission in transgenic mice of a bacterial artificial chromosome. Nature Biotechnol. 15, 859–865.

Yu, D., Ellis, H.M., Lee, E.C., Jenkins, N.A., Copeland, N.G., Court, D.L., 2000. An efficient recombination system for chromosome engineering in *Escherichia coli*. Proc. Natl Acad. Sci. USA 97, 5978–5983.