Actinobacteria

A Ul-Hassan and E M Wellington, University of Warwick, Coventry, England, UK

Glossary

co-metabolism The metabolic transformation of a substance by one organism to a second substance, which serves as a primary source of carbon for another organism.

mycelium The mass of hyphae that forms the vegetative and aerial parts of the streptomycete colony before sporulation.

phylogeny Evolutionary history of a group of organisms.

pseudogene A gene that has lost its protein-coding ability.

taxonomy Science of classification.

Abbreviations

ARDRA	Amplified ribosomal DNA Restriction analysis
CGH	comparative genomic hybridization
DF	dibenzofuran
DR	direct repeat
FAME	fatty acid methyl ester analysis
GITs	gastrointestinal tracts
HGT	horizontal gene transfer
hsp	heat shock protein
IS	insertion sequence
ISP	International *Streptomyces* Project
LFRFA	Low-frequency restriction fragment analysis
MLST	multilocus sequence typing
PAHs	polycyclic aromatic hydrocarbons
PCDOs	Polychlorinated dibenzo-*p*-dioxins
PCR-RAPD	PCR-randomly Amplified Polymorphic DNA
PFGE	pulse-field gel electrophoresis
PyMS	pyrolysis mass spectrometry
RFLP	restriction fragment length polymorphism
rRNA	ribosomal RNA
TEs	transposable elements
VNTRs	variable number tandem repeats

Defining Statement

The *Actinobacteria* form a distinct clade of Gram-positive bacteria which contains a large number of genera. This diverse and important group encompasses key antibiotic-producers, many soil bacteria critical for decomposition and resilient species capable of growing in hostile, polluted environments. A few are medically important pathogens including the causal agent of TB.

Introduction

Some of the earliest descriptions of *Actinobacteria* were those of *Streptothrix foersteri* in 1875 by Ferdinand Cohn and *Actinomyces bovis* in 1877 by Carl Otto Harz. Harz described *A. bovis*, causing a disease of cattle called 'lumpy jaw', as having thin filaments that ended in club-shaped bodies that he considered to be 'gonidia' resembling those of fungi, hence the name *Actinomyces* (Latin for ray fungus). In hindsight, though, the 'gonidia' were almost certainly host cells and so the resemblance to typical fungi was false. A number of other microorganisms were isolated, which had some of the same properties and were thought to belong to the same group as those mentioned above, including the causative agents of leprosy and tuberculosis. This group of organisms was officially recognized as *Actinomycetales* in 1916 and it became apparent that they comprised a large heterogeneous group, which vary greatly in their physiological and biochemical

properties, though their phylogenetic position as true bacteria was not established until the 1960s. The *Actinobacteria* are now considered to be one of the largest phyla of the bacterial kingdom. The GC content of these organisms ranges from 54% in some corynebacteria to more than 70% in streptomycetes. *Tropheryma whipplei*, the causative agent of Whipple's disease, has a GC content of 46.3%, which is the lowest so far reported for *Actinobacteria*.

The *Actinobacteria* are morphologically diverse and can range from coccoids (*Micrococcus*) and rods (*Mycobacterium*) to branching mycelium (*Streptomyces*), many of which can also form spores. Actinobacterial species are ubiquitous in the environment and can be isolated from both aquatic and terrestrial habitats. New species of *Actinobacteria* have been recovered from a diverse range of environments, including medieval wall paintings, desert soil, butter, marine sponges, and radon-containing hot springs. The ability of *Actinobacteria* to inhabit varied environments is due to their ability to produce a variety of extracellular hydrolytic enzymes, particularly in the soil, where they are responsible for the breakdown of dead plant, animal and fungal material, thus making them central organisms in carbon recycling. Some species can break down more complex, recalcitrant compounds, of which *Rhodococcus* species are a good example; they can degrade nitro-, di-nitrophenol, pyridine, and nitroaromatic compounds.

Actinobacteria are well-known for their ability to produce secondary metabolites, many of which have antibacterial and antifungal properties. Of all the antibiotics produced by *Actinobacteria, Streptomyces* species are responsible for ~80%, with smaller contributions by *Micromonospora, Saccharopolyspora, Amycolatopsis, Actinoplanes*, and *Nocardia*. A number of species have developed complex symbiotic relationships with plants and insects. A unique relationship has been reported between the European bee-wolf wasp, in which the female wasps carry *Streptomyces* species in their antennae and apply them to the brood cells. The bacteria are taken up by the larva and colonize the walls of the cocoon, where they protect it from fungal infestation. *Streptomyces* species also share beneficial relationships with plants, with *S. lydicus* being found to enhance pea root nodulation by *Rhizobium* species. The best-studied example of a *Streptomyces*-eukaryotic relationship is between pathogenic strains such as *Streptomyces scabies* that cause scab in potatoes, carrots, beets, and other plants. Strains of *Frankia* can fix nitrogen and are responsible for the nodulation of many dicotyledonous plants. Some members of the *Actinobacteria* are important human and animal plant pathogens. These include *Mycobacterium leprae* (leprosy), *Mycobacterium tuberculosis* (tuberculosis in humans), *Mycobacterium bovis* (tuberculosis in cattle), *Corynebacterium diphtheriae* (diphtheria), *Propionibacterium acnes* (acne), and *Streptomyces somaliensis*, and *Actinomadura* and *Nocardia* species (actinomycetomas).

Systematics of *Actinobacteria*

Traditional Phenotypic Analysis

The best-studied members of the *Actinobacteria* class belong to the genus *Streptomyces*, which was proposed by Waksman and Henrici in 1943. Members of this genus have high GC content DNA, being highly oxidative and forming extensive branching substrate and aerial hyphae. They also produce a variety of pigments responsible for the color of the substrate and aerial hyphae (*Bergey's Manual of Systematic Bacteriology*). *Streptomyces* species are prolific producers of antibiotics, and since the discovery of actinomycin and streptomycin produced by *S. antibioticus* and *S. griseus* respectively during the early 1940s, the interest in streptomycetes grew very rapidly. The discovery of *Streptomyces* species as rich sources of commercially important antibiotics led to new techniques for the cultivation of these organisms. However, due to the lack of standards for the classification and identification of new species, the new strains were described based on only small differences in morphological and pigmentation properties. This, along with the belief that one strain produced only one antibiotic, led to overspeciation of the genus, resulting in over 3000 species being described by the late 1970s. A number of methods were developed to overcome this problem, with the earliest being based on only a few subjectively chosen characters focusing largely on morphological and pigmentation properties that were rarely tested under standardized conditions. Subsequently, biochemical, nutritional, and physiological characters were included, but as these were only applied to selected species they did not necessarily reflect the phylogeny of streptomycetes. The International *Streptomyces* Project (ISP) was established in 1964 with an aim to describe and classify extant type strains of *Streptomyces* using traditional tests under standardized conditions. This study resulted in more than 450 species being redescribed, but an attempt to delineate the genera was futile.

The data collected by the ISP project were used by several researchers to develop computer-assisted identification systems. In 1962, Silvestri was the first to apply numerical taxonomy to the genus *Streptomyces*, where nearly 200 strains were tested for 100 unit characters. This study highlighted the fact that many of the characters used to describe *Streptomyces* species are highly variable and can be erroneously interpreted. Williams and colleagues carried out a more comprehensive study of the streptomycetes. The majority of the strains studied were from the ISP project, along with soil isolates and representative strains from 14 other Actinobacterial genera. Each strain was tested for 139 unit characters, including spore chain morphology, spore chain ornamentation, color of aerial mycelium, color of substrate and extracellular

pigments, production of extracellular enzymes, carbon and nitrogen source utilization, and resistance and sensitivity to certain antibiotics. Strains were clustered according to the observed similarities and this resulted in 19 major, 40 minor, and 18 single-strain clusters being identified, where the single-strain clusters were considered as species and the major clusters were referred to as species groups. The largest species group is cluster 1 *S. albidoflavus*, containing ~70 strains. This cluster is divided into three subclusters. Subcluster 1A contains strains like *S. albidoflavus, S. limosus*, and *S. fellus*, which form hooked or straight chains of smooth, yellow, or white spores and are melanin-negative. Subcluster 1B strains are similar in morphology and pigmentation to those in subcluster 1A and comprise strains such as *S. griseus, S. anulatus,* and *S. ornatus*. Strains in subcluster 1C produce gray, smooth spores and are melanin-negative; *S. olivaceus, S. griseolus*, and *S. halstedii* represent this group.

As in streptomycete taxonomy, early attempts to differentiate mycobacterial species were based on phenotypic properties such as growth rate and pigmentation. With the discovery of new species and the fact that pigment production can be temperature-dependent and that not all strains of a species share pigment-producing abilities, classification of mycobacterial species became less reliable. An alternative scheme was based on the pathogenic potential of a species, although this too was constantly changing as pathogenicity was being discovered in more species of *Mycobacterium*. The International Working Group on Mycobacterial taxonomy was set up in 1967 with an aim to standardize techniques used for the classification of these strains. This led to a numerical taxonomic approach to study mycobacterial strains. Closely related strains of *Corynebacterium, Rhodococcus*, and *Nocardia* were also included in these studies, which revealed that strains belonging to these four genera form distinct clades significantly separate from each other, with many strains being reclassified. Strains identified as *Corynebacterium equi* and *C. hoagii* were found to belong to *Rhodococcus*, and further chemical and genetic analysis confirmed the reclassification as they conformed to the original description of the genus. *Bacterionema matruchotti* was included in this study as its generic position was unresolved, being originally classified as a member of the Actinomycetaceae family. This strain was found to have all the characteristics of true corynebacteria and was renamed *Corynebacterium matruchotti*.

Attempts were made to resolve the confusion surrounding the classification of *Nocardia asteroides* and its relationship to *N. farcinica*. One hundred and forty-nine randomly selected *Nocardia* strains from various sources were analyzed. This study recovered seven major, nine minor, and twelve single-strain clusters. Two apparently identical strains of *N. farcinica* (NTCC 4524 and ATCC 3318) were found to group in two separate clusters. Strain NTCC 4524 clustered with *Mycobacterium* species, whereas 3318 grouped with *N. asteroides*. Numerical taxonomic studies have enabled the description of each species to be made and facilitated the development of new methods that would allow closely related species to be distinguished.

Characterization Based on Chemical Constituents

One of the drawbacks of numerical taxonomy is that it measures phenotypic similarities and differences and these do not always correlate with the genotype and thus only provide an estimate of relatedness between strains. Numerical taxonomy has largely been superseded by chemo- and molecular taxonomy. Chemotaxonomic methods have long been used to determine the relatedness between organisms. Goodfellow and colleagues used comparisons of fatty acid methyl ester analysis (FAME) between bacterial genera. Members of the Streptomycetaceae family have been described as having major amounts of either LL-diaminopimelic acid (LL-A_2pm) (*Streptacidiphilus* and *Streptomyces*) or *meso*-diaminopimelic acid (*meso*-A_2pm) (*Kitasatospora*) in their substrate mycelium and LL-A_2pm as the major diamino acid in aerial or submerged spores. Analysis of whole-cell sugar patterns revealed galactose or galactose and rhamnose (*Kitasatospora* and *Streptacidiphilus*). The presence of LL-A_2pm and glycine, with the absence of characteristic sugars, is typical of the *Streptomyces* cell wall type, which is characterized as Type I. These members also contain saturated iso- and anteiso- fatty acids with either seven or eight hydrogenated menaquinones with nine isoprene units as the predominate isoprenologues. They lack mycolic acids but contain the lipids phosphatidylglycerol, phosphatidylethanolamine, phophatidylinositol, and phosphatidylinositol mannosides. Chemotaxonomic characteristics of other families of the *Actinobacteria* class are summarized in **Table 1**.

Curie-point pyrolysis mass spectrometry (PyMS) has been applied for typing *Actinobacteria*. This method provides a fingerprint of the organisms, which can be used quantitatively to analyze differences between strains. Both FAME and PyMS require stringent standardization as changes in culture media and incubation can affect the results. Many examples exist in the literature where chemotaxonomy has been used successfully for the rapid characterization of new species as well as confirming the integrity of existing taxonomic clusters.

Genotypic Approaches to Determining Relatedness

DNA–DNA hybridization

Streptomyces coelicolor A3(2) and *S. lividans* 66 are members of the cluster 21 *Streptomyces violaceoruber* species group as defined by Williams and colleagues, which represents one of the well-defined species groups of the genus. Cluster 21

Table 1 Chemotaxonomic characteristics of selected families belonging to the class *Actinobacteria*

Family	*Phospholipid pattern*	*Fatty acids*	*Menaquinone*	*Diamino acid*	*Interpeptide bridge*	*Cell wall sugars*
Dermatophilaceae	PG, DPG, PI, PE, PC	$C_{16:0}$, $C_{15:0}$, $C_{14:0}$, $C_{17:0}C_{17:1}$ $C_{18:1}$	MK-8(H_4)	*meso*-A_2pm	None	
Dermacoccaceae	PG, DPG, PI, PE, PC	$C_{17:0}$, $C_{18:0}$, $C_{18:1}$, i-$C_{17:0}$i-$C_{17:1}$, ai-$C_{17:0}$	MK-8(H_4), MK-8(H_2), MK-8MK-9, MK-10	L-lysine	L-Ser_{1-2}-D-Glu/ L-Ser_{1-2}-L-Ala-D-Glu,D-Glu_2, L-Ser-D-Asp	
Cellulomonadaceae	PG, DPG, PI	ai-$C_{15:0}$, $C_{16:0}$	MK-9(H_4)	L-lysine/ornithine	L-Thr←D-Asp/L-Thr←D-Glu,D-Asp, D-Glu	Rhamnose
Micrococcaceae	CL, PG, DPG, PI, PL	i-$C_{15:0}$, ai-$C_{15:0}$,ai-$C_{17:0}$i-$C_{16:0}$	MK-7(H_2), MK-8(H_2), MK-9(H_2)MK-8, MK-9, MK-10	*meso*-A_2pm, Lysine LL-A_2pm, ornithine	L-lysine, L-alanine	Galactosamine, glucosamine
Corynebacteriaceae	DPG, PG, PI, PIM, PG	$C_{16:0}$, $C_{18:1}$, $C_{18:0}$	MK-8(H_2), MK-9(H_2)	*meso*-A_2pm		Arabinose/ galactose
Micromonosporaceae	PC, PE	i-$C_{15:0}$, i-$C_{15:1}$, $C_{17:1}$	MK-9(H_4), MK-10(H_4), MK-10(H_6) MK-12(H_4), MK-12(H_6), MK-12(H_8)	*meso*- or LL-A_2pm	L-glycine	Xylose/arabinose
Streptosporangiaceae	PG, PI, PE	i-$C_{16:0}$, $C_{17:0}$, i-$C_{18:0}$	MK9(H_4), MK-10(H_4)	*meso*-A_2pm		Madurose, glucose, ribose, mannose
Nocardioidaceae	PG, DPG, PE, PC PI, PIM	i-$C_{16:0}$, $C_{16:0}$, $C_{18:0}$, $C_{18:1}$ ai-$C_{15:0}$, i-$C_{14:0}$, $C_{18:1}$	MK-8(H_4), MK-9(H_4), MK-9(H_6) MK-9(H_8), MK-10(H_4)	LL-A_2pm	L-glycine	Glucose, ribose, mannose, galactose
Intersporangiaceae	PI, PIM, PG, DPG, PE	i-$C_{15:0}$, ai-$C_{15:0}$, i-$C_{14:0}$, i-$C_{16:0}$, $C_{17:0}$	MK-8, MK-8(H_4)	LL-A_2pm or *meso*-A_2pm	L-glycine	Glucose

DPG, diphosphatidylglycerol; PG, phosphatidylglycerol; PI, phosphatidylinositol; PIM, phosphatidylinositol mannosides; PE, phosphatidylethanolamine; PC, phosphatidylcholine; CL, cardiolipin; PL, unknown phospholipids; MK, menaquinone; *meso*-A_2pm, *meso*-diaminopimelic acid; LL-A_2pm, LL-diaminopimelic acid. Taken from Dworkin *et al*., 2006.

strains produce smooth gray spores and diffusible pigments, which are blue or red depending on the pH of the medium. *S. coelicolor* A3(2) and *S. lividans* 66 are model representatives of this cluster as they have been genetically, biochemically, and physiologically characterized. However, both strains have had a long and confused taxonomic history.

In 1908, Muller isolated an actinomycete as a contaminant that produced a soluble blue pigment and named it *Streptothrix coelicolor*. In 1916, Waksman and Curtis independently isolated an actinomycete culture from the soil, which also produced a red and blue pigment and was named *Actinomyces violaceoruber*. For a long time the two strains were considered to be synonyms and when the *Streptomyces* genus was established both were named *S. coelicolor* (Muller) Waksman and Henrici in the fourth edition of *Bergey's Manual of Systematic Bacteriology*. Subsequently isolated strains that produced blue pigments were either considered to be *S. coelicolor* (Muller) Waksman and Henrici or as different species altogether. Kutzner and Waksman reexamined all the strains that produced a blue, red, and purple pigment and clarified that the strains isolated by Muller in 1906 (*S. coelicolor* (Muller)) and Waksman and Curtis in 1916 (*S. violaceoruber*) are distinctly different species. Analysis of the blue pigments produced by these strains showed that they are chemically very different. *S. coelicolor* (Muller) is a member of cluster 1 streptomycetes, showing similarity to *S. griseus* and is not a member of the *S. violaceoruber* clade. Monson and colleagues (1969) confirmed the results of Kutzner and Waksman by DNA–DNA hybridization between *S. violaceoruber* and *S. coelicolor* (Muller).

DNA–DNA hybridization experiments are an acknowledged approach in determining the integrity of taxonomic clusters defined by numerical taxonomy, and the study of Monson and colleagues (1969) was one of the first to use this technique. Research based on numerical phenetic and DNA–DNA hybridization data has revealed high levels of congruence as the same taxonomic groups are recovered. The cluster 18 *S. cyaneus* species group is highly heterogeneous, with 9 out of 18 type strains being assigned to two DNA relatedness groups defined at or above the 70% relatedness level. The use of DNA–DNA hybridization experiments also demonstrated that the cluster 32 *S. violaceusniger* species group encompasses several genomic species when type strains are examined. The DNA relatedness groups were defined at similarity levels >70%, seven of which consisted of single members. The multimembered clusters were equated with *S. hygroscopicus* and *S. violaceusniger*. The latter two species were redescribed and a number of strains carrying different specific names reduced to synonyms of the newly redescribed taxa. A high degree of heterogeneity in the *Streptomyces lavendulae* (cluster 61) cluster was reported by Labeda and Labeda and Lyons. A number of strains were related at the species level as they shared high DNA relatedness values (>80%). *S. colombiensis* was reduced to a synonym of *S. lavendulae* as it showed 83% DNA homology to *S. lavendulae* type strain. A number of strains showed <45% relatedness and were therefore considered to belong to a different species group. The DNA relatedness studies discussed above outline the importance of evaluating numerical taxonomic clusters using taxonomic criteria. This is particularly important in the case of the *Actinobacteria* that comprise many species. In situations such as this there is a risk of grouping unrelated or partially related strains together in clusters using insufficient properties.

Members of the *M. tuberculosis* complex include the strains *M. tuberculosis, M. bovis, M. africanum, M. canettii*, *M. microti*, *M. caprae*, and *M. pinnipedii*. DNA–DNA hybridization analysis reveals that this complex comprises a single species. Subsequent analysis of the genomes has revealed little sequence variation among the members. DNA–DNA hybridization has been used to study *Nocardia* species and this allowed the differentiation between species of *N. asteroides* and other members of the genus. Stackebrandt and Fiedler studied 16 strains of *Arthrobacter* and two strains of *Brevibacterium*. The DNA of these strains was hybridized to the *Arthrobacter* type strain *A. globiformis*. This analysis indicates that *Arthrobacter* species share little homology between themselves, with values ranging from 11 to 55%. However, *Brevibacterium sulfureum* and *Brevibacterium protophormiae* showed relatively high homology to the *Arthrobacter* strains. Based on these results it was recommended that the two *Brevibacterium* strains be reclassified as *Arthrobacter* strains. DNA–DNA hybridization is routinely used to aid the characterization of novel species of *Actinobacteria* with many examples in the literature.

Comparative genomic analysis

The availability of whole genome sequences has allowed the development of microarrays, which have revolutionized functional and genomic analysis of organisms. Microarrays have been extensively used to analyze gene expression and regulation and have application in a number of disciplines, including immunology, oncology, forensic science, pharmacogenomics, and drug discovery. More recently, microarrays have been used to investigate the genome-wide analysis between closely related strains, in particular those of pathogens, with an aim to identify pathogenicity determinants. This comparative genomic hybridization (CGH) microarray analysis has been used to investigate a number of bacterial pathogenic species in relation to pathogenesis and host specificity. To date, there are ~115 actinobacterial genomes that are either completed or at various stages of completion (www.genomesonline.org). This has led to the development of microarrays for some sequenced genomes, including those for *S. coelicolor*, and *Mycobacterium* and *Corynebacterium* species. To date, many of the microarray

studies have focused on expression analysis of RNA, particularly with the pathogenic strains of *Mycobacterium* and *Corynebacterium*. Microarrays can also be used for DNA–DNA comparative genomic analysis between closely related strains, as was done for members of the cluster 21 *S. violaceoruber* strains. Using PCR-based microarrays Weaver and colleagues were able to identify 1-Mb amplification of the terminal regions of a number of laboratory strains of *S. coelicolor* A3(2) compared to the sequenced strain M145. The comparison of the cluster 21 streptomycetes also revealed 14 regions that were present in *S. coelicolor* M145 but absent in the other members. These regions encoded genes for biosynthesis of secondary metabolites and heavy metal resistance. All 14 regions were associated with transposon and insertion sequence (IS) elements and the fact they showed a much lower GC content than the rest of the *Streptomyces* chromosome strongly suggests these regions to have been acquired by horizontal gene transfer (HGT) by *S. coelicolor* M145. Ward and Goodfellow have reported a core set of genes from the comparative analysis of the Cluster 21 strains. These genes included those already identified as housekeeping genes and also those for some secondary metabolites; for example, genes for biosynthesis of actinorhodin were found to be conserved among the cluster 21 strains. Ul-Hassan (2006) used oligo-based microarrays for the comparative genomic analysis of soil strains identified as *S. violaceoruber*. The results revealed the strains to have undergone extensive deletions that could be correlated with their observed phenotypes. CGH microarray has been used to investigate the molecular taxonomy of a number of organisms, including *Saccharomyces* and *Yersinia* species and Clostridia. As more actinobacterial genomes are becoming available there is the potential to develop CGH microarray as a taxonomic tool to study environmentally and clinically important *Actinobacteria*.

Restriction digestion analysis of total chromosomal DNA

Restriction digestion analysis of total chromosomal DNA provides a fingerprint of the organisms being analyzed. Low-frequency restriction fragment analysis (LFRFA), restriction fragment length polymorphism (RFLP), and pulse-field gel electrophoresis (PFGE) have all been used to provide an indication of relatedness between strains. However, the results of these methods can be misleading if the strains have undergone large chromosomal deletions or amplifications. Amplified ribosomal DNA restriction analysis (ARDRA) requires the amplification of parts of the rRNA operon, including part of the 16S rRNA, the 16S–23S rRNA spacer region, and part of the 23S rDNA. The amplified PCR products are subjected to restriction enzyme digestion and electrophoresis, providing specific banding patterns for the strains being analyzed. ARDRA has been used to differentiate strains of *Arthrobacter* and *Microbacterium*. With the use of ARDRA in conjunction with PFGE, strain-specific restriction patterns for *Arthrobacter* and *Microbacterium* have been obtained. RFLP of the rRNA gene has been developed as a tool for identification of corynebacterial strains. The strains could be grouped based on the banding patterns and generally strains belonging to the same species clustered together. PCR-restriction enzyme pattern analysis (PRA) of Hsp65 (a heat shock protein) has been developed as a tool for differentiating between *Nocardia* species. However, when PRA analysis was applied to *Nocardia* species, the same banding patterns were recovered, demonstrating this technique not to be as useful for identification or differentiation of these species. Sequence analysis of the *Hsp65* gene displays sufficient polymorphic sites to allow identification. For methods which generate banding patterns that are subsequently used for identification, it is important to stress that a different banding pattern does not necessarily represent a new species. The differences may be attributed to genome rearrangements such as amplifications or deletions, which occur frequently.

PCR-randomly amplified polymorphic DNA (PCR-RAPD) involves PCR of the genome with an arbitrary set of primers to generate a characteristic fingerprint profile. The advantage of this method is that prior knowledge of the chromosomal sequence is not necessary, although stringent standardization is required. The drawback of this technique is that it is highly sensitive and variability in banding patterns may be observed depending on the type of reaction mixture, primers, and concentration of target DNA. This was used to analyze the relationship between *S. lavendulae* and *Streptomyces virginiae*, which were reported by Williams and colleagues to be synonyms. Although consistent results were obtained when comparing to DNA–DNA hybridization, LFRFA, and biochemical properties, the interspecific relationship of *S. lavendulae* and *S. virginiae* remained unresolved.

Analysis of the genomes of members of the *M. tuberculosis* complex demonstrates a number of repeat sequences, of which the best studied is the direct repeat (DR) region. This region consists of DR sequences (36 bp) interspersed with spacer sequences (34–41 bp), collectively termed direct variable repeat sequences. Spoligotyping was developed as a tool for analyzing the structure of the DR region. Spoligotyping patterns are produced by hybridization of sample DNA to oligonucleotides based on the specific sequences in the DR region. An international spoligotype database was set up and now consists of ~39 000 entries from research groups worldwide. The spoligotype patterns can be aligned, enabling researchers to group isolates based on these alignments into clades or strain families. Filliol and colleagues and Molhuizen and colleagues (1998) were able to identify distinct spoligotype patterns for nearly all of the defined *M. tuberculosis*

complex strains. However it could only provide limited discrimination of *M. bovis* strains. Analysis of multiple genomic regions that contain variable number tandem repeats (VNTRs) of different families of genetic elements has been proposed as an alternative tool to examine *M. bovis* strains. The studies of Roring and colleagues and Allix and colleagues have shown VNTR to be more discriminatory than both spoligotyping and RFLP for *M. bovis*.

Phylogeny of *Actinobacteria*

Molecular Analysis of 16S rRNA Sequences

Molecular methods are now used together with numerical and chemotaxonomic techniques to improve the understanding of species relatedness. Woese and Fox used molecular systematic analysis of ribosomal RNA (rRNA) molecules to provide an evolutionary classification of organisms. All 16S rRNAs have conserved primary structures, which allowed Edwards and colleagues, to design universal primers to amplify the entire 16S rRNA gene. Analysis of the 16S rRNA gene revealed regions that are genus-specific and more variable regions that can be used to infer relationships at a lower taxonomic order. Stackebrandt and colleagues identified three regions within the 16S rRNA gene that show variation: α-region (nt 982–998) (*S. ambofaciens* nomenclature) and β-region (nt 1102–1122), which can be used to resolve species to the genus level. The most variable region is the γ-region (nt 150–200), which is species-specific. Inferred relationships based on gene sequences are subject to a number of assumptions with the major condition being that the gene analyzed must not be subject to gene transfer. It is well-known that bacteria contain multiple copies of the 16S rRNA gene but transfer of 16S rRNA has not been determined definitively. It has been suggested that high levels of similarity can facilitate recombination between closely related species, resulting in strains containing chimeric molecules of 16S rRNA. Phylogenetic analysis using 16S rRNA genes can pose problem due to intraspecific variation and intragenomic heterogeneity. Another problem with using 16S rRNA genes is that it is a slowly evolving gene, making it more difficult to resolve the relationships between strains to the species level. Katakoa and colleagues was able to examine the γ-region of a number of streptomycete strains and, although being too variable for determining generic relationships, the inter- and intraspecies relationships were resolved. When Hain and colleagues investigated the *S. albidoflavus* group it was apparent that use of the 16S rDNA sequences were useful for species delimitation but not strain differentiation. The 16S–23S intergenic region was better in determining intraspecific relationships. Based on the similarities of the 16S rRNA gene, Wellington and colleagues included *Kitosatospora* into the *Streptomyces* genus. This was contested by Zhang and colleagues, who demonstrated that strains of *Kitosatospora* always form a stable monophyletic clade away from the streptomycetes when the entire 16S rDNA sequences were used. Streptomycete-specific primers have also been designed for the 23S and 5S rRNA gene. Sequence analysis of the 5S rDNA gene sequences were used to confirm that *Chainia, Elytrosporangium, Kitasatoa, Microellobosporia*, and *Streptoverticillium* as members of the *Streptomyces* genus. Stackebrandt and colleagues proposed a new hierarchic system of classification for *Actinobacteria,* which was based exclusively on 16S rRNA-rDNA sequences.

A comprehensive study on the relationships of members of the *Actinobacteria* class has not been done prior to this work. 16S rDNA sequences from representative strains from each genus of *Actinobacteria* were used to construct a phylogenetic tree (**Figure 1**). It is important to note that, as an up-to-date authoritative list of validly described *Actinobacteria* is not available (e.g., *Bergey's Manual of Systematic Bacteriology*), orders, families, and genera that have been identified were taken from those available in the public database. Four subclasses were recognized: *Actinobacteridae, Acidimicrobidae, Coriobacteridae*, and *Rubrobacteridae*, with *Actinobacteridae* being the largest, consisting of 11 suborders and 43 families (**Table 2**). When generating the tree, care was taken to use sequences of good quality and the cutoff for the length of the sequence was 1400 bp. Any sequences below this were not included in the phylogenetic analysis. Sequences containing ambiguous bases were also excluded from the study. The phylogenetic tree was constructed with PHYLIP, the neighbor-joining method using the kimura-2-parameter model of sequence evolution. Bootstrap was used to calculate the confidence in the groupings of strains and the significant bootstrap values (>80) are indicated on the nodes (**Figure 1**).

The phylogenetic tree shows clear, distinct groupings of strains of particular genera in their respective families that were expected; however, some irregularities can be observed. The genus *Amycolatopsis* was proposed by Lechevalier and colleagues and, based on phylogenetic and chemotaxonomic properties, this genus is placed within the family Pseudonocardiaceae. The genus currently contains ten validated species that include *A. fastidiosa.* However, **Figure 1** clearly demonstrated *A. fastidiosa* to group with *Actinokineospora diospyrosa*, which belongs to the family Actinosynnemataceae, with a significant bootstrap value. In previous phylogenetic studies of the *Amycolatopsis* genus, *A. fastidiosa* consistently grouped outside other members of *Amycolatopsis*. In these studies only *Amycolatopsis* strains have been used occasionally alongside *Pseudonocardia*, but members of the closely related family Actinosynnemataceae were excluded. It is recommended that further work be done to establish to

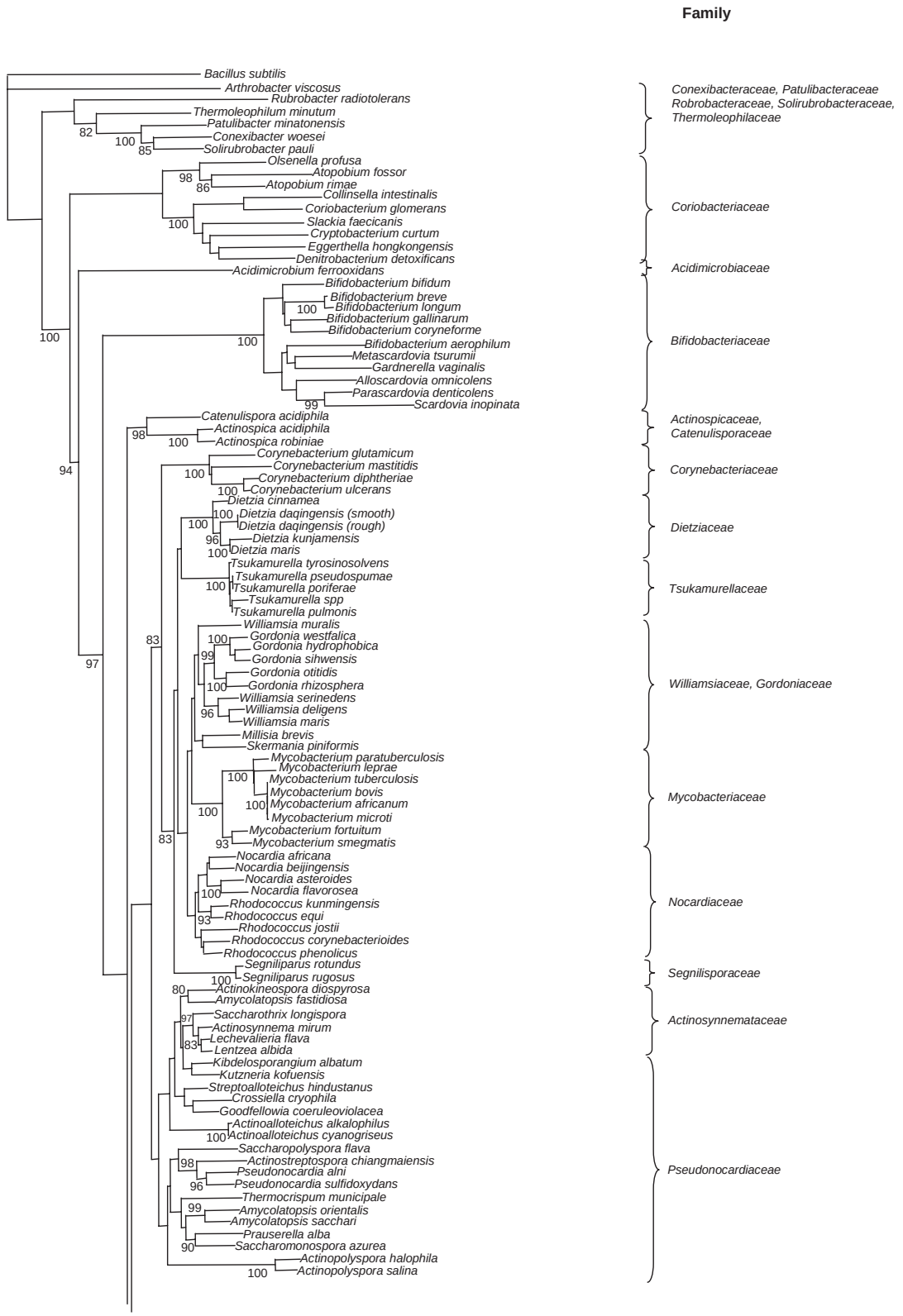

Figure 1 (Continued)

Figure 1 (Continued)

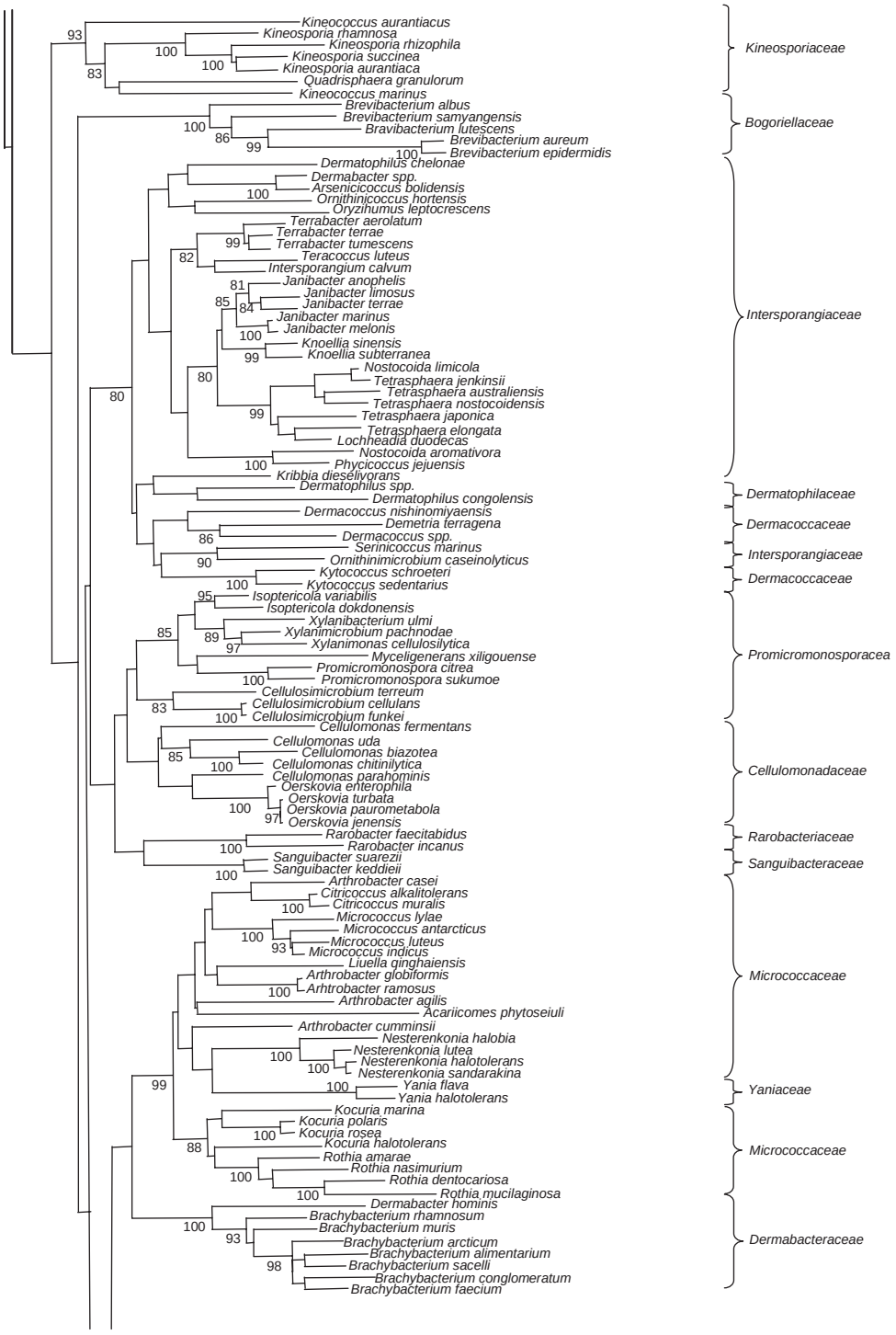

Figure 1 (Continued)

which family *A. fastidiosa* belongs, and in future studies of *Amycolatopsis* strains members of Actinosynnemataceae should be included. The genus *Parvopolyspora* was described by Liu and Lian; however, since *Parvopolyspora pallida* has not been described efficiently according to the rules of Bacterial Nomenclature, it is not considered to be a valid genus/species. Using chemotaxonomic, morphological, physiological, and DNA–DNA hybridization methods, Itoh and colleagues and Miyadoh and colleagues showed *P. pallida* to be closely related to *Actinomadura* with the latter study proposing *P. pallida* to be transferred to the genus *Actinomadura*. Using 16S rRNA gene

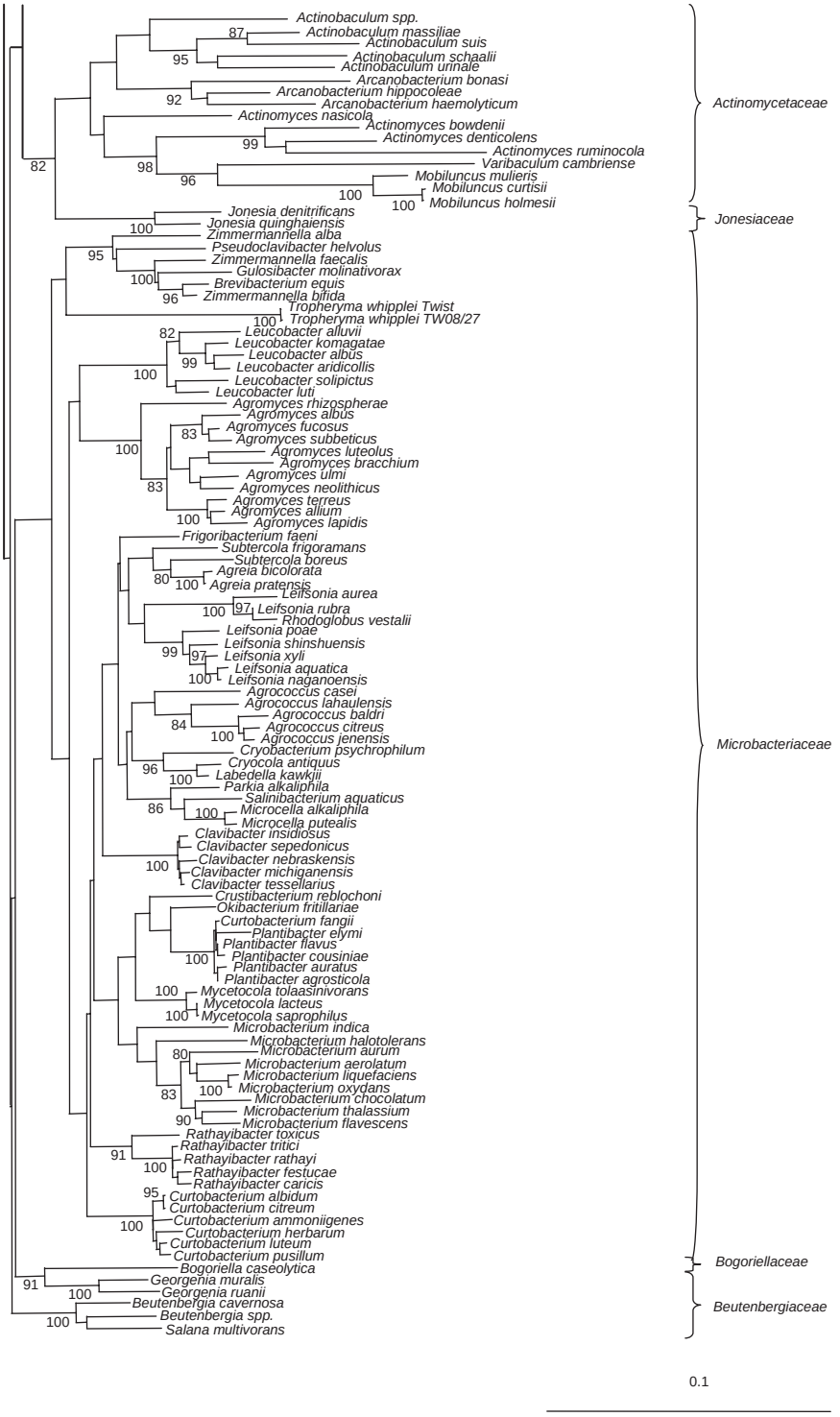

Figure 1 Phylogenetic analysis of species belonging to the class *Actinobacteria* using the near entire 16S rDNA gene sequences. The phylogenetic tree was construced using the neighbour-joining method and Kimura-2-parameter of sequence evolution. Numbers on the branches indicate the percentage bootstrap value of 100 replicates. The scale bar indicates 10% nucleotide dissimilarity (10% nucleotide substitutions per 100 nucleotides).

sequences, Tamura and Hatano confirmed previous work that favored the transfer to *Actinomadura*. In the current study, *P. pallida* groups with *Actinomadura* strains, further confirming earlier work on *P. pallida*, and calls for an emended description of the genus *Actinomadura* to be made to accommodate *P. pallida* as *Actinomadura pallida*.

Table 2 Members comprising the class *Actinobacteria*

Class	*Subclass*	*Order*	*Suborder*	*Family*
Actinobacteria	Acidimicrobidae	Acidimicrobiales	Acidimicrobineae	Acidimicrobiaceae (1)
	Actinobacteridae	Actinomycetales	Actinomycineae	Actinomyceaceae (5)
			Catenulisporineae	Actinospicaceae (1)
				Catenulisporaceae (1)
			Corynebacterineae	Corynebacteriaceae (1)
				Dieziaceae (1)
				Gordoniaceae (3)
				Mycobacteriaceae (1)
				Nocardiaceae (3)
				Segniliparaceae (1)
				Tsukamurellaceae (1)
				Williamsiaceae (1)
			Frankineae	Acidothermaceae (1)
				Frankiaceae (1)
				Geodermatophilaceae (4)
				Kineosporiaceae (3)
				Nakamurellaceae (1)
				Sporichthyaceae (1)
			Glycomycineae	Glycomycetaceae (2)
			Micrococcineae	Beutenbergiaceae (3)
				Bogoriellaceae (1)
				Brevibacteriaceae (1)
				Cellulomonadaceae (3)
				Dermabacteraceae (2)
				Dermacoccaceae (3)
				Dermatophilaceae (2)
				Intrasporangiaceae (14)
				Jonesiaceae (1)
				Microbacteriaceae (24)
				Micrococcaceae (10)
				Promicromonosporaceae (7)
				Rarobacteraceae (1)
				Sanguibacteraceae (1)
				Yaniaceae (1)
			Micromonosporineae	Micromonosporaceae (18)
			Propionibacterineae	Nocardioidaceae (11)
				Propionibacteriaceae (8)
			Pseudonocardineae	Actinosynnemaaceae (6)
				Pseudonocardiaceae (17)
			Streptomycineae	Streptomycetaceae (8)
			Streptosporangineae	Nocardiopsaceae (4)
				Streptosporangiaceae (13)
				Thermomonosporaceae (4)
		Bifidobacteriales		Bifidobacteriaceae (3)
	Coriobacteridae	Coriobacteriales	Coriobacterineae	Coriobacteriaceae (8)
	Rubrobacteridae	Rubrobacterales	Rubrobacterineae	Conexibacteraceae (1)
				Paulibacteraceae (1)
				Rubrobacteraceae (1)
				Solirubrobacteraceae (1)
				Thermoleophilaceae (1)

Reproduced from sequences available from NCBI. Numbers in parenthesis indicate the number of genera in each family.

Microbispora bispora is currently recognized as a member of the Pseudonocardiaceae family, yet in **Figure 1** *M. bispora* is next to *Thermopolyspora flexuosa* within the family Streptosporangiaceae, distinctly separate from Pseudonocardiaceae. Grouping outside both *M. bispora* and *T. flexuosa* is *Thermomonospora chromogena*, which belongs to the family Thermomonosporaceae. These three strains form a separate group from the Pseudonocardiaceae, Thermomonosporaceae, and Streptosporangiaceae families, being closer to the last. The three strains also share many chemotaxonomic traits, so re-examination of these strains is advised to determine their exact positions within the three

families. *Prauseria hordei* is also recognized as a member of Pseudonocardiaceae, though this strain has not been validly described in the literature. Analysis of its 16S rRNA gene has placed it among the Nocardiopsaceae to a strain of *Nocardiopsis* with a bootstrap value of 100, with the 16S rDNA sequences of the two strains showing 99% sequence similarity.

The phylogenetic position of *T. whipplei* has been under some considerable debate. This has largely been due to the lack of chemotaxonomic studies as this organism has been difficult to culture. Initially, *T. whipplei* was placed as a deep-branching strain of the Cellulomonadaceae family. Only recently, La Scola and colleagues gave a detailed description of the Whipple's disease bacillus with regard to its cultivation and morphology. The new genus and species name of *T. whipplei* was based solely on the 16S rRNA gene sequence data as it was not possible to study any chemotaxonomic traits. In a phylogenetic tree presented in *The Practical Streptomyces Handbook*, *T. whipplei* grouped closer to Microbacteriaceae, although not placed within it. Chater and Chandra constructed a phylogenetic tree using the 16S rRNA genes of sequenced *Actinobacterial* genomes, which included two strains of *T. whipplei*. In this study, *T. whipplei* grouped with *Leifsonia* and also showed a close relationship to *Streptomyces* strains. **Figure 1** demonstrates that *T. whipplei* is a member of the Microbacteriaceae family, as indicated by Chater and Chandra, but is not as close to the *Streptomyces* genus as shown in their tree.

Attention is also drawn to the families Dermabacteraceae, Intrasporangiaceae, Dermacoccaceae, and Dermatophilaceae. The last accommodates two genera, *Dermatophilus* and *Kineosphaera*. The genera *Dermacoccus, Kytococcus*, and *Demetria* make up the family Dermacoccaceae and 14 genera have been described for Intrasporangiaceae with Dermabacteraceae containing two genera, *Brachybacterium* and *Dermabacter*. It is evident from **Figure 1** that some strains from these families have been misclassified. For example, a *Dermabacter* species groups with a member of Intrasporangiaceae, *Arsenicicoccus bolidensis*, with a high bootstrap value. The classification of members of these families, in particular Dermatophilaceae and Dermacoccaceae, needs to be studied again to clarify their taxonomic positions.

Speciation of Genera Using Protein-Coding Genes

A number of studies have made use of other housekeeping genes to support the phylogeny derived from the 16S rRNA gene. The genes chosen for phylogenetic analysis must fulfill certain criteria in that they must be essential and distributed among the genera, therefore reducing the possibility of HGT. They must also have an evolutionary rate higher than the 16S rRNA gene, thus providing better resolution of closely related strains. Examples of the housekeeping genes that have been used in conjunction with the 16S rRNA gene include *recA, gyrB, trpB, rpoB, secA1, hsp65, sodA*, and *trpB*. In many of these studies, the use of alternative genes resolved the relationships between closely related species. Ul-Hassan examined the *S. violaceoruber* cluster and phylogenetic analysis was done using 16S rRNA, *gyrB, recA*, and *trpB* genes. The strains formed a tight monophyletic cluster in the partial 16S rRNA gene tree. The results of the *gyrB, recA*, and *trpB* analysis correlated with the 16S rRNA analysis as the topology of the trees and grouping of the strains were identical. Phylogenetic histories for the housekeeping genes were generated and the relative separation in phylogenetic tree space was examined using the Robinson–Foulds distance metric. This confirmed that the genes *gyrB, trpB*, and *recA* show a faster evolutionary rate than 16S rRNA, and therefore being good choices for use alongside the 16S rRNA gene. Analysis of the *S. violaceoruber* strains with *recA, gyrB*, and *trpB* showed no further resolution of the intrageneric relationships between the closely related species. These results were in agreement with those of Duangmal and colleagues, who used the entire 16S rRNA sequence to study type members of the *S. violaceoruber* cluster, as well as soil isolates, and showed that members of the *S. violaceoruber* cluster are highly homogeneous.

The study of Ul-Hassan was similar to a multilocus sequence typing (MLST) approach. MLST is a method for the genotypic characterization between closely related species using the allelic mismatch of a number of housekeeping genes (usually seven). This is a powerful tool which has been used in molecular epidemiology for phylogenetic analysis of bacterial pathogens. *Bifidobacterium* strains have been shown to have high levels of sequence similarity of the 16S rDNA gene ranging from 87.7 to 99.5%, with some strains possessing identical sequences, thus making it difficult to identify and characterize strains. Ventura and colleagues developed MLST to study strains of the *Bifidobacteria* genus. Analysis using the 16S rRNA sequences allowed the discrimination of most species within the genus, but it was more difficult to do so between subspecies. For MLST analysis the genes *clpC, dnaB, dnaG, dnaJi, purF, rpoC*, and *xlp* were used. The phylogenetic tree generated from the concatenated sequences showed a significant increase in the discriminatory power between the strains.

Genome Structure and Evolution

The linearity of the streptomycete chromosome was first determined in *S. lividans* and has subsequently been seen in other members of the genera. Redenbach and colleagues set out to analyze whether large linear chromosomes were a distinct feature of *Streptomyces* species. Linearity of

the chromosomes was determined by PFGE. The results of this study concluded that members of the genera which undergo a complex cycle of morphological differentiation (e.g., *Streptomyces, Micromonospora, Actinoplanes*, and *Nocardia*) possess large linear chromosomes, whereas actinobacterial strains with simpler life cycle (*Mycobacterium, Corynebacterium*, and *Rhodococcus*) have smaller, circular chromosomes. Genomes of 19 medically and industrially important *Actinobacteria* have been completely sequenced and annotated. Analyses of whole genome sequences have provided an insight into how different *Actinobacteria* have become adapted to their particular ecological niches.

The *Streptomyces* Genome

Of the streptomycetes, complete genomes of *S. coelicolor* A3(2) and *S. avermitilis* are available. *S. coelicolor* A3(2) is genetically the best-known representative of the genus as nearly all major achievements in streptomycete genetics and physiology have been done in this model organism. *S. avermitilis* is an important organism in the pharmaceutical industry as it is a major producer of avermectins, which are antiparasitic agents used in human and veterinary medicine. Both *S. coelicolor* A3(2) and *S. avermitilis* have linear chromosomes of 8.7 and 9 Mb respectively. *Streptomyces* species are the predominant *Actinobacteria* in soil, which is a highly heterogeneous matrix composed of organic, inorganic, and gaseous material. Organisms must withstand extremes of temperature and moisture, particularly in the upper layers of the soil profile. Streptomycetes have a saprophytic lifestyle and their ability to successfully colonize the soil is due to the production of a variety of extracellular hydrolytic enzymes. These include nucleases, lipases, amylases, xylanases, proteinases, and chitinases, thus making streptomycetes central organisms in decomposition.

Analysis of the types and location of the genes in the *S. coelicolor* A3(2) chromosome suggests that it comprises a central core region and a pair of chromosomal arms. Genes with an essential function such as in DNA replication, transcription, and translation are located in the core region whereas those with a nonessential function (e.g., secondary metabolism) are located in the arm regions. Significant synteny between the core of the *S. coelicolor* A3(2) genome and the whole genomes of *M. tuberculosis* and *C. diphtheriae* was observed, suggesting these to have a common ancestor, whereas the arms of the *S. coelicolor* A3(2) chromosome consisted of acquired DNA.

Genetic instability of the Streptomyces *genome*

The extreme variability of *Actinobacteria* is a well-known phenomenon first demonstrated by the work of Lieske, and is clearly evident when examining culture plates. In nearly all cases, genetic instability has a pleiotropic effect and can result in the loss of antibiotic biosynthesis and resistance, pigment production, and aerial mycelium formation. Genes can be lost in various frequencies from 10^{-4} to 10^{-2} per spore; these deletions can remove up to 25% of the genome. For a chromosome of 8 Mb this can be up to 2 Mb, which exceeds the size of a small bacterial genome. Initially, genetic instability was considered to be a consequence of the linear structure of the chromosome. However artificially circularized, chromosomes were found to be more unstable than the parent chromosome and it was only after the deletion of the terminal regions that the circular chromosomes became stable structures. Lin and Chen (1997) proposed the high numbers of transposable elements (TEs) in the terminal regions to be responsible for genetic instability. Approximately 40% of TEs are located in the terminal regions of the *S. coelicolor* A3(2) genome, with a similar pattern being seen in *S. avermitilis*. TEs can be found in multiple copies in the genome and this can often lead to DNA rearrangements in the form of transpositions, insertions, deletions, and gene transfer events. The structure of the *Streptomyces* chromosome and the distribution of the essential and nonessential genes and TEs provides some benefit to the organism in that it allows a certain amount of plasticity to the genome. As the terminal regions contain only a few essential genes they are more tolerant to DNA rearrangements and acquisitions. Using DNA microarrays, 14 regions were identified in the *S. coelicolor* A3(2) genome that were absent in its close relatives and of these regions, 11 were located in the arms of the chromosome. Prior to sequencing of the *S. coelicolor* A3(2) genome, gene clusters for actinorhodin, undecylprodigiosin, calcium-dependent antibiotic, and the *whiE* cluster had been analyzed. Sequencing of the chromosome revealed ~20 more clusters encoding putative secondary metabolites. These include clusters for coelichelin, coelibactin, geosmins, desferrioxamines, and hopanoids (**Table 3**).

Reductive Genomes

The ability to easily acquire and lose DNA without causing detrimental effects to the organisms plays a major role in the evolution of these free-living saprophytic bacteria. In contrast to this, obligate intracellular pathogens occupy a stable environmental niche, so gene transfer does not play such a crucial role in the evolution and adaptation of these organisms. A process of reductive evolution is seen to occur in these pathogens where a number of gene functions become redundant as the host will supply these needs. The presence of pseudogenes in a genome gives an estimation of gene decay. A gene for a particular function will be made redundant when the functional constraint is relaxed, thus making it prone to inactivating mutations. When these mutations become fixed in a population the gene becomes a

Table 3 Secondary metabolites produced by *S. coelicolor* A3(2)

Secondary metabolites	*Location on chromosome*
Known structures	
Antibiotics	
Actinorhodin	SCO5071-SCO5092
Calcium dependant antibiotic	SCO3210-SCO3249
Prodiginines	SCO5877-SCO5898
Methylenomycin	SCP1 plasmid
Siderophores	
Coelibactin	SCO7681-SCO7691
Coelichelin	SCO0489-SCO0499
Desferrioxamines	SCO2782-SCO2785
Pigments	
Isorenieratene	SCO0185-SCO0191
Tetrahydroxynaphthalene	SCO1206-SCO1208
TW95a (*whiE*) spore pigment	SCO5314-SCO5320
Lipids	
Eicosapentaenoic acid	SCO0124-SCO0129
Hopanoids	SCO6759-SCO6771
Other molecules	
Butyrolactones	SCO6266
Geosmin	SCO6073
Unknown structure	
Chalcone synthases	SCO7669-SCO7671, SCO7222
Deoxysugar synthases/ glycosyl transferases	SCO0381-SCO0401
Nonribosomal peptide synthetases	SCO6429-SCO6438
Sesquiterpene cyclase	SCO5222-SCO5223
Siderophore synthetase	SCO5799-SCOSCO5801
Type I polyketide synthases	SCO6273-SCO6288, SCO6826-SCO6827
Type II fatty acid synthase	SCO1265-SCO1273

Taken from Bentley *et al*. (2002) and Challis and Hopwood (2003).

pseudogene. These genes will either remain in the genome and be subjected to further mutations to such an extent that they are no longer recognizable or are completely removed.

M. leprae

The best documented example of reductive evolution is seen in *M. leprae*. The complete genomes of *M. leprae* and *M. tuberculosis* have been sequenced and are much smaller than the streptomycete genomes: 3.2 Mb (*M. leprae*) and 4.4 Mb (*M. tuberculosis*). When the two mycobacterial strains are compared they show a large difference in GC content with *M. leprae* having an average GC of 57.8% while that of *M. tuberculosis* is 65.6%. The most striking feature of the *M. leprae* genome is that it contains 49.5% protein-coding genes (1604 genes) compared to 90.8% (3959) protein-coding genes in *M. tuberculosis*. The number of pseudogenes in *M. leprae* is 1116, with only six being found in *M. tuberculosis*, indicating massive gene decay in *M. leprae*. It is now proposed that *M. leprae* has evolved to have the natural minimal gene set for Mycobacteria and, unlike *M. tuberculosis*, has a limited metabolic repertoire and host range. Reduction in the genome size is also connected with the observation that intracellular pathogens make extensive use of host cellular processes.

T. whipplei

The genome of *T. whipplei* is the most extreme example of an Actinobacterium that has undergone genome reduction. *T. whipplei* is the causative agent of Whipples disease, which is characterized by malabsorption and is a systemic infection affecting any part of the body. Like *M. leprae*, *T. whipplei* was also difficult to culture and it was only in 2000 that it was grown in human fibroblasts and exhibits a slow doubling time of 17 days comparable to 14 days for *M. leprae*. The genome sequence of the *T. whipplei* Twist strain has recently become available and shows it to have a small circular chromosome of only 0.92 Mb. The average GC content of the genome is 46%, which is considerably lower than that of streptomycetes and mycobacterial strains and is in contrast to the reduced *M. leprae*. Analysis of the genome has shown that enzymes involved in information processing (DNA/RNA polymerases and gyrases) are present. In *M. leprae*, genes for biosynthesis of amino acids were present, suggesting that these are limiting in their environment, but in *T. whipplei* it became apparent that complete and partial losses of some amino acid biosynthesis gene clusters have occurred. This implies that amino acids are obtained from the host. At least two amino acids and peptide ABC transport systems were identified in the genome of *T. whipplei*. An interesting observation from the genome included the identification of *whiA* and *whiB*, which in *S. coelicolor* are involved in sporulation. Spores have not been reported in *T. whipplei* under laboratory conditions, although they may perhaps arise in the environment. The genome shows little evidence of gene acquisition, which is a common property of reduced genomes. In other examples of reduced genomes like those of *M. leprae* and *Rickettsia prowazekii*, high numbers of pseudogenes are present, suggesting that these genomes are still in the process of downsizing. In contrast, the *T. whipplei* genome contains a few pseudogenes, indicating that no further genome decay is occurring.

P. acnes

P. acnes is a commensal bacterium found on human skin with preference to sebaceous follicles, but it can be an opportunistic pathogen and can cause acne. *P. acnes* contains a single circular chromosome of 2.5 Mb encoding 2333 genes. Putative functions were assigned to 68% of the genes with 20% sharing no significant similarity to any database entries. The presence of 35 pseudogenes containing frameshift mutations or premature stop codons leads the author to

suggest this gene decay to be a recent event. Analysis of the whole genome led to ten regions being identified as possibly being of foreign origin, one of which contained genes for the biosynthesis of lanthionine. Other genes identified were involved in substrate uptake and pathogenicity. With regard to the physiology of *P. acnes*, all genes of the Embden–Meyerhof and pentose-phosphate pathway are present. *P. acnes* can grow anaerobically on a number of substrates, and enzymes required for this have been identified.

Prior to the genome being sequenced, GehA was recognized as an extracellular triacylglycerol lipase, which degrades skin tissue components. The ability of lipases to degrade human skin lipids results in the production of fatty acids that assist in bacterial adhesion and colonization of the follicles. Many other lipases and esterases have been identified, including endoglycoceramidase that breaks down glycosphingolipids, which exist in the cell membranes of all vertebrates. Three putative sialidase enzymes have been identified, along with sialic acid transport proteins, suggesting that *P. acnes* cleaves sialoglycoconjugants to obtain sialic acid as a source of carbon and energy. Homologs of CAMP factors have also been recognized in *P. acnes* and these had only previously been detected in streptococcal species. CAMP factors are secreted proteins that are known as pathogenic determinants and have lethal effects when given to mice and rabbits. CAMP factors have been suggested to act as pore-forming toxins. Three enzymes with putative hemolytic activities show some resemblance to hemolysin III of *Bacillus cereus*. Many surface proteins that can act as antigens have been identified and can trigger an inflammatory response that is seen during acne.

An interesting feature of some of the genes is the presence of continuous stretches of 12–16 guanine or cytosine residues, either in the promoter region or at the 5′end. The length of this poly(C)/(G) tract is variable and is generated during replication by slipped-strand mispairing. These tracts are involved in phase variation, which serves as an adaptation mechanism whereby the organism can change its phenotype to evade immune responses or rapidly adapt to environmental changes.

Bifidobacteria longum

Members of the *Bifidobacterium* genus comprise 3–6% of the adult fecal flora and the presence of these organisms is thought to provide health benefits. This has led to the increase in the use of *Bifidobacterium* species in health-promoting foods. *Bifidobacterium* species are obligate anaerobes and the majority of strains have been isolated from mammalian gastrointestinal tracts (GITs). They are among the first to colonize GITs of newborn babies until weaning, when *Bacteroides* take over. This system of successive colonization is thought to play a major role in the build-up of immune system tolerance. Thus a complex balance of microflora is needed for a normal and healthy digestive system. This becomes evident after antimicrobial therapy, where the incidence of GIT disorders greatly increases. However, little is known about the physiology and genetics of the organisms and the mechanism of host-microbe interaction. Sequencing the genome of *B. longum* has provided important clues into the adaptation of *Bifidobacterium* to GITs of humans. The genome of *B. longum* is 2.3 Mb and it is estimated that 86% of the genome is protein coding. No aerobic or anaerobic respiratory components were identified, confirming *B. longum* to be a strict fermentative anaerobe. Homologs of the enzymes needed for the fermentation of glucose, including the fructose-6-phosphate shunt and a partial Embden–Meyerhof pathway, are present. Enzymes that are needed to feed many sugars into the fructose-6-phosphate are present, further confirming the ability of *B. longum* to ferment a large variety of sugars.

B. longum is able to ferment amino acids through the use of 2-hydroxyacid and other predicted deaminases and dehydratases. More than 20 putative peptidases have been predicted and these enable *B. longum* to obtain amino acids from proteinaceous material in the GIT where carbohydrates are less abundant. *Bifidobacterium* species colonize the lower GIT, which tends to be poor in mono- and disaccharides as they are taken up by the host and the microflora in the upper GIT. As a result, more that 8.5% of the predicted proteins are dedicated to carbohydrate transport and metabolism. Many glycosyl hydrolases have been predicted in the genome and these cover a wide range of substrates including di-, tri-, and higher order oligosaccharides. Oligosaccharide transporters have also been identified, which may aid *B. longum* to compete for the uptake of structurally diverse oligosaccharides. Unlike most bacteria, *B. longum* makes great use of negative transcriptional control to regulate gene expression, with nearly 70% of its transcriptional regulators being repressors. Negative repressors are thought to allow for a more precise response to changes in the environment, which is consistent with the need of *B. longum* to adapt to the constantly changing conditions of the GIT.

Industrially Important Phenotypes of *Actinobacteria*

The best-studied example of an important industrial application of *Actinobacteria* is the production of antibiotics by streptomycete strains. As mentioned previously, the discovery of actinomycin and streptomycin in the early 1940s led many large pharmaceutical companies in different parts of the world to initiate large screening programs in a hope of finding novel antibiotic compounds. This resulted in a rapid increase in the rate at the discovery of new compounds between the 1940s and the 1960s. These years are now considered to be the Golden Age of antibiotic discovery, as after 1960s the

rate at which new compounds were discovered decreased sharply. Initial methods used in isolating new compounds were based on simple plating procedures where the soil samples were mixed with water and subsequent filtering to remove large soil particles. The extract was then plated on nutrient medium. Representative colonies were isolated and studied further for antibiotic production. It is now well known that antibiotic production is part of secondary metabolism and occurs only under certain nutritional conditions. One method of discovering new compounds is by isolating new organisms, and Takahashi and Ōmura provide an excellent review on different isolation methods that have been developed based on this rationale.

Considerable interest has been applied in screening marine organisms for the discovery of novel compounds. The study of Okazaki and colleagues and subsequent research by the group has reported the isolation of an actinobacterial strain from the Sagami bay area producing a novel bioactive compound. This isolate would only produce this compound in selective sea water containing Japanese seaweed. This study highlighted the fact that normal culture media are not sufficient for these organisms as they have adapted to producing bioactive compounds under marine-specific nutritional conditions. It is therefore important to study and understand the physiology of marine *Actinobacteria* to develop effective techniques for their isolation. *Actinobacteria*-specific bacteriophage have been successfully used for isolation and identification of novel or rare actinobacterial strains from terrestrial environments and to determine the relatedness of actinobacterial strains. Kurtboke (2005) developed an improved method for detecting marine *Actinobacteria*. This method uses the actinophage to reduce the number of common marine organisms that tend to outgrow any rare *Actinobacteria*, increasing the likelihood of isolating new actinobacterial strains that potentially produce novel bioactive compounds. Actinophage can be used for host identification at the genus and the species level. In general, streptomycetes phage are genus-specific, although some cross-reactivity has also been detected with other genera, including *Nocardia, Streptosporangium*, and *Mycobacterium*.

Sequencing of mycobacterial strains is done mainly because they are important human and animal pathogens, but the genomes of two nonpathogenic mycobacterial strains have been sequenced. *Mycobacterium* sp. KMS was isolated from soil sites that had been polluted by creosols, pentachlorophenol, and polycyclic aromatic hydrocarbons (PAHs), which contain up to four aromatic rings. Another strain of *Mycobacterium vanbaalenii* PYR-1 was found to possess the remarkable ability to degrade PAHs, including alkyl- and nitro-substituted PAHs such as naphthalene. Both these mycobacterial species use dioxygenases and monooxygenases for the oxidation of the ring component of these compounds. PAH compounds are toxic and have carcinogenic properties, and microbial degradation of these compounds is the most effective method of remediation of the contaminated soil. Vinyl chloride is a potential carcinogen and tends to accumulate as an end-product of dechlorination of solvents such as perchloroethylene and trichloroethene.

A strain of *Nocardioides* sp. JS614 was isolated from an industrial soil site that was contaminated with vinyl chloride and 1,2-dichloroethane. Higher growth yields were obtained when *Nocardioides* sp. JS614 was grown on media containing vinyl chloride than without; the strain is unusually sensitive to vinyl chloride starvation. This strain possess a 300-kb plasmid carrying genes encoding monooxygenases and epoxyalkane: coenzyme M transferase, thought to be involved in degradation of vinyl chloride. *Nocardioides* species are known to degrade other aromatic compounds, including 2,4,6-trinitrophenol, phenanthrene, and dibenzofuran (DF). Carbazole is an N-heterocyclic aromatic compound derived from creosote and crude and shade oil and is known to be both toxic and mutagenic. It is widely used as a raw material for the production of dyes, medicines, and plastics. *Nocardioides aromaticivorans* IC177 is able to degrade carbazole. Inoue and colleagues were able to clone and partially sequence the *car* genes responsible for carbazole degradation. The sequences showed similarities to genes found in *Pseudomonas* and *Sphingamonas* strains.

Arsenic is a highly toxic metal and its presence in the environment is largely from a geochemical source (rocks and minerals), although anthropogenic action has led to its increase. Many organisms have been documented to be able to transform arsenic, through either reduction or oxidation reactions; these include *Cenibacterium arsenoxidans, Alcaligenes faecalis, Agrobacterium tumefaciens, Bacillus*, and *Shewanella*. An arsenic-defence mechanism is present in all organisms studied for arsenic degradation. Members of the genus *Corynebacterium* are of great biotechnological importance, especially for the large-scale production of amino acids such as L-glutamate and L-lysine. Members of the coryneform bacteria (*C. glutamicum* and *C. lactofermentum*) are resistant to arsenic and genes involved in this process are contained in two operons, *ars1* and *ars2*. Research of Mateos and colleagues aims to make use of genetically engineered strains of *C. glutamicum* for bioremediation of arsenic from heavily contaminated water sites.

Members of the genus *Gordonia* belong to the same family as Corynebacteria (Corynebacteriaceae). A strain was isolated from polluted water taken from the inside of a deteriorated rubber tyre. By using chemotaxonomy, DNA–DNA hybridization and 16S sequence analysis this strain was found to represent a new species within the genus *Gordonia* and was given the name *Gordonia westfalica*. It could utilize natural and synthetic components of rubber,

including *cis*-1,4-polyisoprene. *Janibacter terrae* strain XJ-1 was found to have the ability to degrade DF. Polychlorinated dibenzo-*p*-dioxins (PCDOs) and polychlorinated DF (PCDFs) are common pollutants in the environment and are released as contaminants in pesticides and herbicides. These are highly toxic compounds and tend to accumulate in the body fat of animals. A DF-degrading strain of *J. terrae,* which can use DF as a sole source of carbon and energy, was also isolated. *J. terrae* contains the *dbdA* (DF dioxygenase) gene cluster and sequence analysis demonstrated it to be nearly identical to the cluster found on a large plasmid of *Terrabacter* sp. YK3, which utilizes DF in a similar manner.

Among the *Actinobacteria, Rhodococcus* species are well-known for their ability to biodegrade and transform a wide range of complex organic compounds. It is for this reason they have been referred to as 'masters of catabolic versatility' by Larkin and colleagues. Many of the genes required for degradation of xenobiotic compounds are encoded on plasmids, including those for polychlorinated biphenyls, isopropylbenzene, and indene. Genes associated with virulence in pathogenic strains (*R. equis*) are also encoded on plasmids. Readers are recommended the articles by Larkin and colleagues, Sekine and colleagues, and Gurtler and colleagues for the detailed analysis of the overall metabolic diversity and genetics of these organisms. Pesticides are composed of compounds with varying chemical structures, including organochlorides, *s*-trazines, triazinones, organophosphates, and sulfonylureas. Members of *Actinobacteria* play a major role in biotransformation and biodegradation of these chemicals, where a single strain can degrade more than one compound though co metabolism. De Schrijver and De Mot provide a comprehensive review on the degradation of pesticides by *Actinobacteria.*

It is clear from the examples discussed above that members of the *Actinobacteria* play a major role in bioremediation and biodegradation of complex xenobiotic compounds in the environment. They have the genetic capabilities to either utilize these compounds as sources of energy or break them down to simpler forms, which in turn can be used by other organisms (cometabolism). *Actinobacteria* also have an enormous biotechnological potential. As mentioned before members of the *Actinobacteria*, in particular the genus *Streptomyces*, are major producers of medically important antibiotics. Examples of other uses of Actinobacteria are listed in **Table 4**.

Table 4 Uses of actinobacterial strains in biotechnology

Organism	*Biotechnological uses*
Mycobacterium(nonmedical strains)	• Biotransformation of steroids • Removal of vinyl chloride from industrial waste • Production of optically active epoxides which are subsequently used for chemical synthesis of optically active pharmaceutical compounds
Corynebacterium (nonmedical strains)	• *C. glutamicum* used for the large-scale production of L-glutamic acid and L-lysine. These strains can also be modified to produce threonine, isoleucine, tylosine, phenylalanine, and tryptophan • Fermentative production of nucleotides which are used as flavor enhancers in foods
Microbispora rosea	• Produces D-xylose isomerase, which converts glucose into fructose, which is subsequently used to produce high-fructose syrup
Micrococcus species	• Used for processing of fermented meats to improve color, aroma, flavor and keeping quality • Synthesis of long-chain aliphatic hydrocarbons that have the potential to be processed into lubricating oils or other petroleum substitutes
Rhodococcus species	• Rhodococcal nitrile converting enzymes used to convert nitriles into their corresponding higher value acids and amides, which can be used as polymers in dispersants, flocculants, and superabsorbents • Commercial production of biosurfactants. • Capacity to degrade a diverse range of hydrocarbons, including halogenated and long chain as well as aromatic compounds • Biopurification of coal and crude oil by removing contaminating organosulfur compounds
Frankia	• Along with other actinorhizal organisms are used where it is necessary to rapidly establish a plant cover
Cellulomonas species	• Used for single-cell protein production from a variety of waste products • Mixed cultures can be used to convert xylan into methane via hydrolysis, acidogensis, and methanogensis • Photoevolution of molecular hydrogen using cellulose as sole carbon source
Micromonospora	• Commercial scale production of amylases and cellulases • Vitamin B_{12} production
Brevibacterium	• Used for cheese ripening
Nocardioides	• Used for their ability to perform chemical and enzymatic modifications of complex compounds and production of industrially important enzymes

Concluding Remarks

The application of chemotaxonomy, numerical taxonomy, and DNA–DNA hybridization methods have provided a basis for studying the taxonomy of *Actinobacteria*. New advances in DNA technology have contributed considerably to bacterial taxonomy, in particular sequence analysis of small subunit rRNA. Sequence analysis of the 16S rRNA is routinely used in conjunction with analysis of chemotaxonomic traits to identify and describe existing and newly isolated strains. Phylogenetic analyses of families comprising the *Actinobacteria* class have been studied in great detail with regard to the genera they accommodate; however, in many of these studies members of other closely related families are not included. The aim of this chapter was to examine the phylogenetic relationships of all the genera that have been described as *Actinobacteria* using the entire 16S rDNA gene sequences available in the public database. The results of this study correlated well with previous analyses, and clusters that have already been defined by numerical taxonomy were retained. This study also highlights the importance of using a combination of traditional methods along with newer molecular based techniques for taxonomic purposes. Also when generating the phylogeny of strains, the analysis of more strains from closely related families is essential to allow a more accurate discrimination between strains.

Bacterial genomes are under constant selection pressure whether they are in the environment in the presence of toxic chemicals, living a saprophytic lifestyle in the soil, or as obligate intracellular pathogens. *Actinobacteria* represent a heterogeneous group of organisms that have the ability to adapt to their particular ecological niches. Free-living species such as *Streptomyces* have made use of HGT, tolerating acquisition and loss of genes, thus allowing them to acclimatize to their fluctuating environment more rapidly. In contrast, pathogenic bacteria (*Mycobacterium* and *Tropheryma*) have chosen the path of reductive evolution where nearly all genes not essential for growth are lost. *Actinobacteria* represent an important group of organisms for the bioremediation of water and soil sites that have been polluted with toxic recalcitrant compounds. The increasing availability of whole genome sequences of actinobacterial strains and their ongoing analysis has revealed the enormous genetic capabilities of this important group of bacteria.

See also: Antibiotic Production; Ecology, Microbial; Horizontal Transfer of Genes between Microorganisms; *Streptomyces*; Tuberculosis: Molecular Basis of Pathogenesis

Further Reading

Chater KF and Chandra G (2006) The evolution of development in *Streptomyces* analysed by genome comparisons. *FEMS Microbiology Reviews.* 30: 651–672.

De Schrijver A and De Mot R (1999). *Degradation of pesticides by Actinomycetes. Critical Reviews Microbiology.* 25: 85–119.

Dworkin M, Falkow S, Rosenberg E, Schleifer K, and Stackebrandt E (eds.) (2006) *The Prokaryotes*, 3rd edn., vol. 3, Archaea, Bacteria: Firmicutes, Actinomycetes. New York: Springer.

Goodfellow M, Ferguson EV, and Sanglier JJ (1992) Numerical classification and identification of *Streptomyces* species – a review. *Gene.* 115: 225–233.

Hopwood DA (2007) *Streptomyces in nature and medicine: The Antibiotic Makers New York*. Oxford University Press, Inc.

Larkin MJ, Kulakov LA, and Allen CC (2005) Biodegradation and *Rhodococcus* – masters of catabolic versatility. *Current Opinion Biotechnology* 16: 282–290.

Lechevalier MP and Lechevalier H (1970) Chemical composition as a criterion in the classification of aerobic Actinomycetes. *International Journal of Systematic Bacteriology* 20: 435–443.

Lin YS and Chen CW (1997) Instability of artificially circularized chromosomes of *Streptomyces lividans*. *Molecular Microbiology* 26: 709–719.

Loria R, Kers J, and Joshi M (2006) Evolution of plant pathogenicity in *Streptomyces*. *Annual Reviews Phytopathology* 44: 469–487.

Stackebrandt E, Rainey FA, and Ward-Rainey NL (1997) Proposal for a new hierarchic classification system. *Actinobacteria* classis nov. *Journal of Systematic Bacteriology* 47: 479–491.

Takahashi Y and Omura S (2003) Isolation of new actinomycete strains for the screening of new bioactive compounds. *Journal of General and Applied Microbiology* 49: 141–154.

Ward AC and Bora N (2006) Diversity and biogeography of marine *Actinobacteria*. *Current Opinion Microbiology.* 9: 279–286.

Wellington EM and Toth IK (1996) Studying the ecology of actinomycetes in the soil rhizosphere. In: Hall GS (ed.) *Methods for Examination of Organismal Diversity in Soils and Sediments* pp. 23–41. CAB International, UNESCO and IUB.

Williams ST, Goodfellow M, and Alderson G (1989) Genus *Streptomyces* Waksman and Henrici 1943, 339^{AL}. In: Williams ST, Sharpe ME, and Holt JG (eds.) *Bergey's Manual of Systematic Bacteriology,* vol. 4, pp. 2452–2492. Baltimore: Williams and Wilkins.

Williams ST, Goodfellow M, Alderson G, *et al.* (1983) Numerical classification of *Streptomyces* and related genera. *Journal of General Microbiology.* 129: 1743–1813.

Relevant Website

http://www.genomesonline.org – GOLD Genomes OnLine Database v 2.0