Genome sequence of the Fleming strain of *Micrococcus luteus*, a simple free-living actinobacterium

Michael Young[1]*, Vladislav Artsatbanov[2], Harry R. Beller[3], Govind Chandra[4], Keith F. Chater[4], Lynn G. Dover[5], Ee-Been Goh[3], Tamar Kahan[6], Arseny S. Kaprelyants[2], Nikos Kyrpides[7], Alla Lapidus[7], Stephen R. Lowry[7], Athanasios Lykidis[7], Jacques Mahillon[8], Viktor Markowitz[9], Konstantinos Mavrommatis[7], Galina V. Mukamolova[10], Aharon Oren[11], J. Stefan Rokem[12], Margaret C. M. Smith[13], Danielle I. Young[1] and Charles L. Greenblatt[12]

[1] *Institute of Biological Environmental and Rural Sciences, Aberystwyth University, Aberystwyth, Ceredigion, SY23 3DD, UK.*

[2] *Bakh Institute of Biochemistry, Leninsky Pr. 33, Moscow 119071, Russia.*

[3] *Joint BioEnergy Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, USA.*

[4] *John Innes Centre, Norwich NR4 7UH, UK.*

[5] *Biomolecular and Biomedical Research Centre, School of Applied Sciences, Northumbria University, Newcastle upon Tyne, NE1 8ST, UK.*

[6] *Bioinformatics Unit, Faculty of Medicine, The Hebrew University of Jerusalem, Ein Kerem, Jerusalem, Israel.*

[7] *DOE-Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA94958, USA.*

[8] *Laboratory of Food and Environmental Microbiology, Université Catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium.*

[9] *Biological Data Management and Technology Center, Lawrence Berkeley National Laboratory, Berkeley, CA, USA.*

[10] *Department of Infection, Immunity and Inflammation, University of Leicester, PO Box 138, Leicester, LE1 9HN, UK.*

[11] *Department of Plant and Environmental Sciences, The Institute of Life Sciences, and the Moshe Shilo Minerva Center for Marine Biogeochemistry, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel.*

[12] *Department of Microbiology & Molecular Genetics, IMRIC, Hebrew University Hadassah Medical School, P.O.B, 12272, IL-91120 Jerusalem, Israel.*

[13] *Institute of Medical Sciences, University of Aberdeen, Foresterhill, Aberdeen AB252ZD, UK.*

*Corresponding author. Mailing address: Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Penglais Campus, Aberystwyth, Ceredigion SY23 3DD, UK. Phone: 44-1970-622348. Fax: 44-1970-622354. Email miy@aber.ac.uk

Abstract

Micrococcus luteus (NCTC2665, "Fleming strain") has one of the smallest genomes of free-living actinobacteria sequenced to date, comprising a single circular chromosome of 2,501,097 bp (G+C content 73%) predicted to encode 2403 proteins. The genome shows extensive synteny with that of the closely related organism, *Kocuria rhizophila*, from which it was taxonomically separated relatively recently. Despite its small size, the genome harbors 73 IS elements, almost all of which are closely related to elements found in other actinobacteria. An IS element is inserted into the *rrs* gene of one of only two *rrn* operons found in *M. luteus*. The genome encodes only four sigma factors and fourteen response regulators, indicative of adaptation to a rather strict ecological niche (mammalian skin). The high sensitivity of *M. luteus* to β-lactam antibiotics may result from the presence of a reduced set of penicillin-binding proteins and the absence of a *wblC* gene, which plays an important role in antibiotic resistance in other actinobacteria. Consistent with the restricted range of compounds it can use as a sole source of carbon for energy and growth, *M. luteus* has a minimal complement of genes concerned with carbohydrate transport and metabolism and its inability to utilize glucose as a sole carbon source may be due to the apparent absence of a gene encoding glucokinase. Uniquely among characterized bacteria, *M. luteus* appears to be able to metabolize glycogen only via trehalose, and to make trehalose only via glycogen. It has very few genes associated with secondary metabolism. In contrast to other actinobacteria, *M. luteus* encodes only one resuscitation-promoting factor (Rpf) required for emergence from dormancy and its complement of other dormancy-related proteins is also much reduced. *M. luteus* is capable of long-chain alkene biosynthesis, which is of interest for advanced biofuel production; a three-gene cluster essential for this metabolism has been identified in the genome.

INTRODUCTION

Micrococcus luteus, the type species of the genus *Micrococcus* (family *Micrococcaceae*, order *Actinomycetales*) (117), is an obligate aerobe. Three biovars have been distinguished (138). Its simple, coccoid morphology delayed the recognition of its relationship to actinomycetes, which are typically morphologically more complex. In the currently accepted phylogenetic tree of the actinobacteria, *Micrococcus* clusters with *Arthrobacter* and *Renibacterium*. Some other coccoid actinobacteria originally also called *Micrococcus*, but reclassified into four new genera (*Kocuria*, *Nesterenkonia*, *Kytococcus* and *Dermacoccus*), are more distant relatives (121). The genus *Micrococcus* now includes only five species, *M. luteus*, *M. lylae*, *M. antarcticus*, *M. endophyticus* and *M. flavus* (20, 69, 70, 121).

In this paper we report the genome sequence of *Micrococcus luteus* NCTC2665 (= DSM 20030^{T}), a strain of historical interest, since Fleming used it to demonstrate bacteriolytic activity (due to lysozyme) in a variety of body tissues and secretions (29, 30), leading to its designation as *Micrococcus lysodeikticus* until its taxonomic status was clarified in 1972 (59). *M. luteus* has been used in a number of scientific contexts. The ease with which its cell wall could be removed made it a favored source of bacterial cell membranes and protoplasts for investigations in bioenergetics (28, 34, 89, 93). Because of the exceptionally high GC content of its DNA, *M. luteus* was used to investigate the relationship between codon usage and tRNA representation in bacterial genomes (51, 52, 61). Although it does not form endospores, *M. luteus* can enter a profoundly dormant state, which could explain why it may routinely be isolated from amber (39). Dormancy has been convincingly demonstrated under laboratory conditions (53-55, 83) and a secreted protein (Rpf) with muralytic activity is involved in the process of resuscitation (81, 82, 84, 85, 87, 125, 133).

Micrococci are also of biotechnological interest. In addition to the extensive exploitation of these and related organisms by the pharmaceutical industry for testing and assaying compounds for antibacterial activity, micrococci can synthesize long-chain alkenes (1, 2, 127). They are also potentially useful for ore dressing and bioremediation applications, since they are able to concentrate heavy metals from low grade ores (26, 66, 67, 116).

Given its intrinsic historical and biological importance, and its biotechnological potential, it is perhaps surprising that the genome sequence of *Micrococcus luteus* was not determined previously (130). Here we consider the strikingly small genome sequence in these contexts and also in relation to the morphological simplicity of *M. luteus* as compared with many of its

actinobacterial relatives, which include important pathogens as well as developmentally complex, antibiotic-producing bacteria with some of the largest bacterial genomes.

METHODS

Genome sequencing, assembly and gap closure. The genome of *M. luteus* was sequenced at the Joint Genome Institute (JGI) using a combination of 8-kb and fosmid (40-kb) libraries and 454 pyrosequencing. All general aspects of library construction and sequencing performed at the JGI can be found at http://www.jgi.doe.gov/. Pyrosequencing reads were assembled using Newbler assembler (Roche http://www.454.com/downloads/protocols/1_paired_end.pdf). Large Newbler contigs were chopped into 2793 overlapping fragments of 1000 bp and entered into assembly as pseudo-reads. The sequences were assigned quality scores based on Newbler consensus q-scores with modifications to account for overlap redundancy and adjust-inflated q-scores. Hybrid 454/Sanger assembly was performed with Arachne assembler (8). Possible mis-assemblies were corrected and gaps between contigs were closed by custom primer walks from sub-clones or PCR products. The error rate of the completed genome sequence of *M. luteus* is less than 1 in 50,000.

Genome annotation was performed with the Integrated Microbial Genomes Expert Review (IMG-ER) annotation pipeline (72). Predicted coding sequences (CDSs) were additionally manually evaluated using JGI's Quality Assurance pipeline (http://geneprimp.jgi-psf.org/).

Genome analysis. Comparative analysis of *M. luteus* with related organisms was performed mainly using a set of tools available in IMG. The cutoff for the minimal size of an open reading frame (ORF) was set to 30 residues. Unique and homologous *M. luteus* genes were identified by using BLASTp (cutoff scores of $E < 10^{-2}$). Reciprocal hits were calculated based on these values. Comparisons between genes for the identification of common genes were conducted using BLAST similarities of e value 10^{-5} and similarity > 20% (3). For the determination of orthologs in selected other genomes BLASTp (3) was employed with an expect value threshold of $1e^{-4}$. Two proteins were considered orthologous only if they were reciprocal best hits of each other in BLAST databases consisting of all proteins encoded by their respective genomes. In some situations where detailed "manual" evaluation was made, apparent orthologs which showed less than 50% amino acid identity or covered less than 80% of the longer protein were flagged as doubtful, and rejected unless they were part of a syntenous segment. Signal peptides were identified using SignalP 3.0 (10) and TMHMM (62) at default values.

Following automated annotation of the completed genome sequence, a workshop was held in Jerusalem (April 13th – 18th, 2008), as part of the IMG expert review process, to compare *M.*

luteus with other closely related actinobacteria and illuminate its relationship to *Arthrobacter* and *Kocuria rhizophila* (formerly *Sarcina lutea*) (123), which are two of its closest relatives. Comparative genomic studies were performed in the context of the Integrated Microbial Genomes system (IMG) v.2.2 (73). Habitat information was retrieved from the Genomes OnLine Database (68).

Nucleotide sequence accession numbers. The sequence data described here have been deposited in GenBank (NC_012803) and the Genomes OnLine Database accession number Gc01033.

RESULTS AND DISCUSSION

General features and architecture of the genome. The genome consists of one circular chromosome of 2,501,097 bp with a 73% GC content. It is one of the smallest actinobacterial genomes sequenced to date. The origin of replication was identified in a region upstream of *dnaA*, where a significant shift in the GC skew value was observed (Fig. 1). In common with *K. rhizophila*, there is no evidence of the global GC skew (preference for G on the leading strand) that is commonly observed in many other bacterial genomes, including those of some of its close relatives, including *Arthrobacter aurescens* and *Renibacterium salmoninarum* (78, 123, 137). Of 2458 predicted genes, 2403 encode proteins. Putative functions were assigned to 74.2% of genes while the remaining 25.8% were annotated as hypothetical proteins. The ATG start codon is used most frequently (1718 times) followed by GTG and TTG (665 & 20 times, respectively). There is an even stronger bias to the use of TGA as stop codon (2238 times), with TAG and TAA being used only 115 & 50 times, respectively.

The properties and statistics of the *M. luteus* genome are summarized in Table 1. There are two rRNA operons with the typical order of 16S, 23S and 5S RNA genes. However, in one of them (Mlut_14290), the *rrs* gene (encoding the 16S rRNA) is interrupted by an IS*L3* family transposase (see below). Apart from the transposon insertion, Mlut_03860 has an additional 'A' residue at position 1437 that is absent in Mlut_14290. A BLAST search of the nucleotide data base with Mlut_03860 showed that of the top 100 hits (classified as *M. luteus, Micrococcus* sp. or uncultured bacterium) 95 clearly do not have the A at position 1437. The transposon insertion has therefore inactivated the *rrs* gene that mostly closely resembles that of the species. Possibly the presence of both functional *rrs* genes was detrimental to the functionality of the ribosome and the transposon insertion in Mlut_14290 rescued fitness. Forty-eight tRNA genes were identified encoding tRNAs theoretically capable of translating all codons including AGA, which had previously been found to be untranslatable using an *in vitro* system derived from a *M. luteus* strain (61). Eighteen additional RNA genes were predicted using models from the RNA families database (Rfam) (40). The substantially reduced number of paralogous genes present (9.2%) as compared with *Arthrobacter* spp. (57.6% for *Arthrobacter aurescens* TC1 and 61.3% for *Arthrobacter* sp FB24) correlates with the reduced size of the *M. luteus* genome. This indicates that most of the gene reduction or expansion that post-dates the last common ancestor of *Micrococcus* and *Arthrobacter* has affected paralogous gene families.

Genes found in a 0.9 Mbp segment including the presumed origin of replication show no particular strand bias, whereas genes located in the remainder of the genome (0.5 – 2.1 Mbp) are more abundant on the presumed leading strand (Fig. 1). RNA genes are more abundant on one chromosome arm (0 – 1.25 Mbp) than on the other (1.25 – 2.5 Mbp).

In order to evaluate the functional content of the genomes of various members of the *Micrococcaceae*, the numbers of genes assigned to different COG functional categories are compared in Table 2 and Fig. 2. Generally speaking, the relative proportions of genes in different functional categories are in line with expectation, based on its small genome size, as described previously (60). *M. luteus* devotes a greater proportion of its genome to core processes of translation, replication, and repair than do the other members of the *Micrococcaceae* except *K. rhizophila*, which has a reduced complement of genes concerned with replication and repair. On the other hand, *Arthrobacter* spp. and *Renibacterium salmoninarum* devote a greater proportion of their genomes (between 10% and 11%) to transcription and its regulation than do either *M. luteus* or *K. rhizophila* (6% and 7%, respectively), and have a greater repertoire of sigma factors and associated regulatory proteins. Other prominent differences are the reduced number and proportion of genes concerned with carbohydrate metabolism in *M. luteus* (and *K. rhizophila*) as compared with other members of the *Micrococcaceae* and also, the presence of a very large number of genes encoding transposases in the *M. luteus* genome (see below). In line with their small genome sizes, both *M. luteus* and *K. rhizophila* have fewer genes concerned with secondary metabolism than *R. salmoninarum* and the two *Arthrobacter* spp. Genes within the other COG functional categories shown in Table 2 (amino acid metabolism, lipid metabolism, energy production and ion transport) increase in number in proportion to genome size.

Dot plots comparing the positions of genes in *M. luteus* with their putative orthologs in other actinobacteria reveal extensive synteny with other members of the *Micrococcaceae*, with evidence for one and two inversions about the presumed replication origins in the comparisons with *K. rhizophila* and *Arthrobacter* sp. FB24, respectively (Fig. 3). Synteny, though interrupted by many more inversions about the origin, was also evident with more distantly related organisms, such as *Clavibacter michiganensis*, *Renibacterium salmoninarum* and *Mycobacterium tuberculosis*, and even with *Streptomyces coelicolor* A3(2), despite the linearity and ca. three-fold larger size of this streptomycete genome (11).

TTA-containing genes are exceptionally rare in *M. luteus*. The high G+C content of actinobacterial genomes is correlated with a paucity of A+T-rich codons. This is particularly

marked for the TTA codon, one of six encoding leucine, to the extent that in streptomycetes the codon is found only in genes that are non-essential for growth. For example, *S. coelicolor* has only 145 TTA-containing genes, and the determinant for the cognate tRNA can be deleted without impairing vegetative growth. It has been proposed that the translation of UUA-containing mRNA is subject to checkpoint control by regulation of the availability of charged cognate tRNA (Ventura et al., 2007; Chandra and Chater, 2008; Chater and Chandra, 2008). Remarkably, an even smaller proportion of *M. luteus* genes, just 24 out of 2403, contain a TTA codon. It would be of considerable interest to find out whether the elimination of the relevant tRNA determinant has any phenotypic consequences.

Mobile genetic elements. Although *M. luteus* has one of the smallest actinobacterial genomes, it harbors more than 70 IS elements, or their remnants. Thirty distinct IS elements are present, representing eight of the 23 well-characterized IS families, *viz*: IS*3*, IS*5*, IS*21*, IS*30*, IS*110*, IS*256*, IS*481* and ISL*3* (Table 3). No elements related to the Class II transposon, Tn*3*, were found. Although the IS*3* family elements show the greatest diversity (8 distinct elements are present), only one of them is intact. There are five distinct types of IS*256* family elements, most of which (19/24 copies) are intact. Some regions contain several elements, including examples of one IS being inserted into another. For example, there are three IS*3* elements into which IS*256* family members (ISMlu*1*, ISMlu*2* or ISMlu*11*) have inserted. With only one exception (*Burkholderia mallei*), all ISMlu transposases have their closest relatives in other actinomycetes, *viz*: *Brevibacterium linens*, *Corynebacterium diphtheriae*, *Corynebacterium jeikeium*, *Corynebacterium striatum*, *Leifsonia xyli*, *Mycobacterium avium*, *Mycobacterium branderi*, *Mycobacterium marinum*, *Mycobacterium smegmatis*, *Rhodococcus aetherivorans*, *Rhodococcus erythropolis*, *S. coelicolor*, and *Terrabacter* sp. One copy of ISMlu*4* (ISL*3* family) has inserted into one of the two *M. luteus rrs* genes (Mlut_14290) (see above). There is evidence from *Escherichia coli* that such insertions are not necessarily polar, so the genes encoding the 23S and 5S ribosomal RNA downstream of Mlut_14290 may be expressed (15, 79). Another element, ISMlu*9* (IS*481* family), is located just downstream of the corresponding 5S rRNA.

A complete compilation of the "ISome" of *M. luteus* is given in the ISfinder database (115). The plethora of IS elements in *M. luteus* may be responsible, at least in part, for the intra-species heterogeneity that has been described previously (90, 138).

A search for integrated genetic elements associated with integrase/recombinase proteins revealed 3 serine recombinases, Mlut_16170, Mlut_06210 and Mlut_00100. Mlut_00100 and

Mlut_06210 are members of the family of large serine recombinases which are common in amongst the actinomycetes and low GC% Gram-positive bacteria (118). These recombinases can be phage-encoded or present on integrating conjugative elements (ICEs) where they excise an element from the chromosome to form a circle of DNA that then undergoes conjugation to a new host (16, 88). In the recipient the circular molecule integrates, often site-specifically via the action of the recombinase. Mlut_00100 lies at one end of 59 genes, generally of low GC%, that interrupt a region with synteny with *Arthrobacter* (Fig. 4). This element, IEMlut1, may be conjugative as it carries a putative relaxase and DNA primase which could act to initiate conjugal transfer of an excised circle of DNA. Paralogues of *dnaK, grpE, dnaJ, clpB* and thioredoxin metabolism genes encoded on IEMlut1 may have a role in overcoming environmental stress.

A second element, IEMlut2, appears to have integrated into a flavin dependent oxidoreductase represented by the gene fragments Mlut_06220 and Mlut_06130. When these two fragments are spliced together and used to search the protein database, they align well with close homologues e.g. from *Arthrobacter*. The large serine recombinase, Mlut_06210, is at one end of this element and is probably responsible for the integration in the ORF. There is a putative relaxase gene, albeit annotated as a pseudo-gene, suggesting that this element might once have been conjugative. IEMlut2 encodes a putative mercuric reductase and a MerR-like regulator.

IEMlut3 and IEMlut4 were probably mobilized by a conserved gene triplet acting together. An alignment of the XerD homologues (which are members of the tyrosine recombinase family of site-specific recombinases) shows that Mlut_06590 and Mlut_20700 are almost identical as are Mlut_06600 and Mlut_20690. In addition Mlut_06610 is 83% identical to Mlut_20680. These 6 genes therefore represent two examples of a conserved triplet of genes that can also be seen in *Mycobacterium* sp. KMS, *M. smegmatis, Arthrobacter*, an organism denoted *Vibrio angustum* S14 in the Genomes OnLine Database (http://genomesonline.org/index2.htm) and others. In *M. luteus* the triplets are located at each end of two low GC rich regions, IEMlut3 and IEMlut4 (Fig. 4). IEMlut3 encodes one of the few sigma factors in *M. luteus* so acquisition of this element may have had a global effect on gene expression. The extent of this element is currently defined only by the lower than average GC% as it lies in a region with little synteny with *Arthrobacter*. IEMlut4 contains putative arsenic and cadmium resistance determinants (Mlut_20620 & Mlut_20660 respectively).

Two of the four putative integrated elements may therefore have been the vehicles for introducing mercury, arsenic and cadmium resistance into *M. luteus*. In addition, Mlut_18330

(a DDE family transposase; no other paralogues in *M. luteus*) is adjacent to a putative copper resistance protein (68% GC; Mlut_18340) suggesting the possible presence of a small, fifth element within *M. luteus*.

Regulation and signal transduction. Detailed analysis of individual genes revealed that only 103 *M. luteus* genes, i.e. 4.4%, including its four sigma factors (see below), encode likely DNA-interacting regulatory proteins, while 27 others (1.1%) have roles in signal transduction. One might expect that as genome size approaches some minimal level for a fully functional organism, the proportion of regulatory genes that have orthologs in related genomes of larger size would approach (though not reach) 100%. However, the suite of *M. luteus* regulatory genes includes many that do not appear to be conserved in the genomes of other members of the *Micrococcaceae* (*K. rhizophila* and *Arthrobacter* sp. strain FB24). To evaluate this more closely, we carried out a limited "manual" analysis of apparent orthologs, based on the relatively stringent criteria of reciprocal BLASTp hits plus at least 50% amino acid identity over at least 80% of the length of the longer protein. On this basis, 69 of the 99 genes (excluding sigma factors) were considered to be orthologous with genes in one or more of the other genomes (see supplementary material Table 1S).

Of the four genes likely to encode sigma factors, Mlut_13280 encodes the principal sigma factor (RpoD). It is very similar to the principal sigma factor (HrdB) of *S. coelicolor* (84% identity over the C-terminal 344 residues; 35% identity over the N-terminal 153 aa). The other three appear to encode ECF sigma factors, which are usually involved in responses to extra-cytoplasmic stresses or stimuli. One of these, Mlut_07700, is widely conserved and corresponds to SCO5216 of *S. coelicolor,* whose product (SigR) plays a key role in responses to disulfide stress. Mlut_07700 is located immediately upstream of an ortholog of *rsrA,* which in other actinomycetes encodes an anti-sigma that antagonizes SigR. The micrococcal RsrA retains the conserved cysteine residues involved in sensing disulphide stress. Another sigma factor gene, Mlut_06410, is species-specific, and located within IEMlut3 (see above and Fig. 4), and the third, Mlut_14900, appears to be present in *Arthrobacter* and possibly other actinobacteria. The absence of any Class III sigma genes (i.e., related to sigma B of *Bacillus subtilis*), and of any genes encoding the relevant classes of anti-sigmas or anti-anti-sigmas, is unusual among Gram-positive bacteria. In general, members of that class of sigmas are involved either in non-nutritional stress responses or sporulation. There was no gene identified for a sigma-54 – this class of sigmas appears to be confined to Gram-negative bacteria. Since sigma factors underpin nearly all cellular developmental programs in bacteria, the simplicity of

the sigma factor profile suggests that there is no undiscovered developmental program in *M. luteus* and that its range of stress responses may be exceptionally narrow.

The genome encodes 14 response regulators, accounting for an unusually high proportion (14%) of the total suite of regulatory proteins (only about 6% of *Streptomyces coelicolor* regulatory proteins are response regulators). Eleven of them form clear two-component systems with an adjacent gene encoding the cognate histidine kinase, two of which are widely conserved. PhoRP (Mlut_03740 & Mlut_03750) probably respond to phosphate limitation and MtrAB (Mlut_14770 & Mlut_14760) probably play a role in cell-cycle progression (but here without the widely conserved adjacent gene *lpqB*, which appears to play an accessory role in the action of MtrAB in other actinobacteria). In addition, the Mlut_14100 & Mlut_14110 genes are probably involved in the regulation of citrate/malate metabolism, whereas the Mlut_04120-Mlut_04130 gene pair, of unknown function, is present in many simple actinobacteria, but absent from *S. coelicolor*. Three response regulator genes are "orphans" not located very close to genes for histidine protein kinases, while just two of the 13 genes for histidine protein kinases are located away from any response regulator gene. All but one (Mlut_03350) of the "orphan" response regulators retain the highly conserved aspartate expected to be phosphorylated, as well as other conserved residues in the typical "phosphorylation pocket", suggesting that their regulation is via phosphorylation. Atypical response regulators that seem unlikely to be regulated by conventional phosphorylation are found in other organisms, e.g. *S. coelicolor* (45). Among the three orphans, one is widespread among the actinobacteria (Mlut_11030) but there appears to be no information about the functions of the orthologs, at least among streptomycetes. One is confined to the *Micrococcinaceae* (Mlut_03350) and one (Mlut_21850) is genome-specific.

M. luteus has three *pkn* genes encoding serine/threonine protein kinases (STPKs): Mlut_00760, Mlut_00750, and Mlut_13750, which correspond to *pknA*, *pknB*, and *pknL* found in *My. leprae*, *C. glutamicum*, and *My. tuberculosis* (which has 11 *pkn* genes) (6, 22, 23, 96). The three STPK genes (see also below, under cell division, morphogenesis, and peptidoglycan biosynthesis), together with a gene encoding a partial STPK sequence apparently fused to a protein of unknown function, represent about 4% of the "regulatory" genome (a similar proportion as in *S. coelicolor*).

Unusual small iron-sulfur cluster-containing regulatory proteins resembling the archetypal WhiB sporulation protein of streptomycetes (hence the term Wbl, for WhiB-like) have been found in all actinobacteria, and no other bacteria. *M. luteus* has an unusually small number

(two) of them. In general, orthologs of some of these small genes are widely conserved, though their small size makes it difficult to be confident about orthology. One of the two (Mlut_05270) appears to be orthologous with the near-ubiquitous archetypal *whiB,* which is important not only for development in streptomycetes, but also for cell-division in mycobacteria (38, 119). The other is similar to *wblE,* which is also near-ubiquitous, but whose role is less clear, though it has been implicated in the oxidative stress response in corynebacteria (58). The absence of a *wblC* gene may account for the sensitivity of *M. luteus* to many antibiotics. In mycobacteria and streptomycetes WblC is a pleiotropic regulator that plays an important role in resistance to diverse antibiotics and other inhibitors (80).

The presence of two *crp*-like genes (Mlut_09560 & Mlut_18280), and a putative adenylate cyclase gene (Mlut_05920), indicates that cAMP plays a signaling role in *M. luteus*, as in most other actinobacteria. Mlut_18280 is conserved among other actinobacteria, so it probably plays the major role in sensing cAMP.

The *M. luteus* genome contains a strikingly high representation (11) of genes related to *merR*, whose products typically exert their regulatory effects by compensating for aberrant spacing between the -10 and -35 regions of the promoters they regulate (95). There are two putative *hspR* genes for the regulation of the heat-shock response (Mlut_00590 & Mlut_18780), both located next to *dnaJ*-like genes, and three genes for ArsR-like regulators. The *arsR*-like genes are adjacent to genes encoding a cation (zinc, cobalt, cadmium) diffusion facilitator (Mlut_13910), an arsenic resistance protein (Mlut_20620) and a cadmium transporter (Mlut_20660) and they are all located in regions mainly comprising *M. luteus*-specific genes (see above). One of the *merR*-like genes (Mlut_06140) is close to genes involved in mercury resistance (e.g. mercuric reductase, Mlut_06150) and another (Mlut_20770) is related to certain excisionases. At least four of the MerR/ArsR proteins are likely to play a role in metal-resistance, which is of interest, since *M. luteus* may have potential utility for gold recovery from low-grade ores (66, 67).

Cell division, morphogenesis, and peptidoglycan biosynthesis. Many of the known genes dedicated to peptidoglycan synthesis, cell division, and morphogenesis are conserved between *M. luteus* and *My. tuberculosis*. *M. luteus* peptidoglycan is of subgroup A2, in which an L-Ala-D-Glu-L-Lys-D-Ala stem peptide (with a glycyl modification of the D-Glu component) is cross-linked to the D-Ala residue of its counterpart by an identical tetrapeptide (107). All but one of the expected genes required for production of a UDP-N-acetylmuramate-pentapeptide-N-acetyl glucosamine precursor were readily identified (Table 4), the exception being that for the enzyme responsible for the glycinyl modification of the D-glutamate residue of the stem peptide.

Schleifer (106, 107) proposed that the use of a stem peptide as a functional interpeptide, which is a feature of *M. luteus* peptidoglycan, could involve cleavage between the *N*-acetylmuramate moiety of one PG monomer and the L-Ala of its peptide component after that peptide had been directly cross-linked to the L-Lys of a neighboring chain. Thereafter, the terminal D-Ala would be linked to the ε-amino group of the L-Lys of a third stem peptide. Either a single transpeptidase (TP) with broad acceptor specificity, or two specific TPs are required. The Mlut_16840 product, which is probably a member of the amidase_2 superfamily (pfam 01510, $9e^{-15}$ BITS score 53.7), appears to be a strong candidate for the *N*-acetylmuramoyl-L-alanine amidase. It bears a twin-arginine transporter type N-terminal signal sequence in conjunction with a Cys residue occupying position 33, suggesting that it is a lipoprotein, as required for this proposed function.

According to Ghuysen (35), class A high-molecular-mass-penicillin-binding proteins (HMM-PBPs) can perform all the basic functions required for PG polymerization. Many bacteria possess several class A HMM-PBPs (37) that may functionally substitute for each other (57), but *M. luteus* possesses only one, encoded by Mlut_18460.

Class B HMM-PBPs possess transpeptidase activity and contain additional modules that may mediate protein-protein interactions (44) or assist with protein folding (37). They are involved in septation, lateral wall expansion and shape maintenance (48, 100, 120, 136). *M. luteus* and *My. tuberculosis* share similar complements of class B HMM-PBPs. As is seen with the cognate *My. tuberculosis* genes, Mlut_13660 (*ftsI*) is associated with the division/cell wall (DCW) cluster and Mlut_00770 is clustered with other genes encoding cell division and cell shape-determining factors such as FtsW and the regulatory elements PknAB (37). *M. luteus* lacks an ortholog of the third *My. tuberculosis* PBP, Rv2864c, which may contribute to its well-known β-lactam sensitivity.

The products of Mlut_01190 and Mlut_16800, potentially D-alanyl-D-alanine carboxypeptidases, may be involved in cell wall remodeling or provide the extra TP potentially required to incorporate the inter-peptide unit of PG. Mlut_16800 is close to Mlut_16840, encoding the putative MurNAc-L-alanine amidase that may also be involved in this process. Similarly, the putative soluble murein transglycosylase encoded by Mlut_13740 probably participates in cell wall remodeling.

The DCW cluster is present in many organisms, and a hypothetical archetypal cluster has been defined (92, 114) (Fig. 5) that has been broadly maintained in bacilli. It has been dispersed or rearranged in other lineages such as coccal firmicutes, or has been modified to accommodate developmental processes such as sporulation (75). Most of the genes that have dispersed from the DCW clusters of Gram-positive organisms encode enzymes that supply cytoplasmic PG precursors. The organization of the *M. luteu*s DCW cluster appears almost identical to that of its rod-shaped relative, *My. tuberculosis*. This might seem paradoxical given the dispersal that has occurred in other lineages (e.g. firmicutes), in which representatives have developed a coccoid morphology. However, unlike *B. subtilis* and *E. coli*, actinomycetes (including the rod-shaped *My. tuberculosis*) do not show appreciable intercalary insertion of new PG; this activity is more or less restricted to the cell poles (126). It is not known whether the spherical shape of *M. luteus* cells masks a hidden polarity in relation to growth and division.

The Div1B/FtsQ proteins encoded by the DCW clusters found in several cocci possess an extended hydrophilic region immediately preceding the largest hydrophobic region of the protein towards the N-terminus (98). Although FtsQ from several actinobacteria (*My. tuberculosis C. glutamicum*, *Arthrobacter* spp., *K. rhizophila*, *R. salmoninarum*) also possesses this extended hydrophilic region, it is absent from the *M. luteus* FtsQ homolog encoded by Mlut_13580, which is some 90 residues shorter. Moreover, the N-terminus of the predicted *M. luteus* protein lacks 33 amino acid residues compared with *B. subtilis* Div1B, inviting the speculation that these differences in protein architecture may have functional significance in the organism's transition to a coccoid form.

Two of the STPK determinants (see above), the *pknAB* genes, are part of a conserved actinobacterial gene cluster implicated in cell division and morphogenesis that also includes *rodA* (Mlut_00780), whose product is involved in the control of cell shape, *pbpA* (Mlut_00770) encoding a transpeptidase involved in PG cross-linking, and *pstP* (Mlut_00790) encoding a protein phosphatase that dephosphorylates PknA and PknB in *My. tuberculosis* (13, 21). In *My. tuberculosis*, the genes in this cluster have a single transcriptional start site and the

start and stop codons of successive genes overlap, suggesting transcriptional and translational coupling. This relationship is apparently conserved in *M. luteus*. Furthermore, the extracellular domain of PknB has been described as a penicillin-binding and Ser/Thr kinase-associated (PASTA) domain that is also found in the bifunctional HMM-PBPs involved in PG synthesis. This domain may bind both penicillins and PG-related analogues (141) as well as muropeptides, effectively coupling cell envelope synthesis to other core processes including transcription and translation (111). One of the phosphorylation targets of *My. tuberculosis* PknA is the product of *wag31* (50). This essential gene (104), also found in *M. luteus* (Mlut_13520) with a conserved genetic context and neighboring the DCW cluster, encodes a homolog of DivIVA that controls placement of the division septum in *B. subtilis* (18) but appears to differ functionally in actinomycetes. Recent studies using *C. glutamicum* and other actinobacteria suggest its role in polar peptidoglycan synthesis is more significant than its involvement in septation (65, 105).

Anionic wall polysaccharides. The cytoplasmic membrane of *M. luteus* bears a α-D-mannosyl-(1→3)-α-D-mannosyl-(1→3)-diacylglycerol (Man_2-DAG) glycolipid and a succinylated lipomannan (sucLM) based on it (64, 94, 97). The lipomannan components of *Corynebacterium glutamicum* and *M. luteus* have structural similarities (76, 77, 124), and the genes encoding sucLM biosynthesis in *M. luteus* were identified by comparison with the cognate genes from *C. glutamicum*. The product of Mlut_04450 is a strong candidate for one of the mannosyltransferases that forms Man_2-DAG. Genes encoding homologs of MptA (Mlut_09700) and MptB (Mlut_09690) form a cluster with another gene (Mlut_09710) that encodes a GT-C family glycosyltransferase, suggesting that Mlut_09710 is also involved in sucLM biosynthesis. These three genes are co-transcribed and probably translationally coupled, since each overlaps its predecessor by 4 nucleotides, suggesting coordinate regulation through a polycistronic mRNA, whereas their homologs in corynebacteria and mycobacteria are widely dispersed. Assuming Mlut_09690 and Mlut_09700 are MptBA orthologues, they would, in concert, provide an α-(1→6) linked mannosyl backbone, a common feature in other lipoglycans. Mlut_09710 might then provide either the 2- or 3-linked mannose residues reported in early characterizations (108). However, it is also possible that each of these three GT-C glycosyltransferases produces a distinct linkage and that this operon provides all of the biosynthetic capability to produce the bulk of the sucLM mannan. A homolog of mycobacterial and corynebacterial polyprenyl monophosphomannose synthases, necessary to provide mannosyl donors to MptAB, is encoded by Mlut_12000. Interestingly, a gene

encoding a C-N hydrolase commonly found immediately downstream or, in the case of *My. tuberculosis*, fused in a continuous reading frame, is absent from *M. luteus* (36, 41).

The *M. luteus* cell wall is decorated with a teichuronic acid (TUA) consisting of repeating disaccharide units of *N*-acetyl-mannosaminuronic acid (ManNAcU) and glucose (Glc) (42, 47). The polymer is attached to PG *via* the phosphate group of a reducing terminal trisaccharide consisting of two ManNAcU residues and an *N*-acetylglucosamine (GlcNAc) phosphate residue (33). The TUA operon of *B. subtilis* provides few useful search queries and the genes concerned with TUA biosynthesis in most other organisms have not been well characterized. *M. luteus* contains three homologs of UDP-ManNAcU dehydrogenase (Mlut_05630, Mlut_08960 & Mlut_08980), which might produce the ManNAcU nucleotide donor. Mlut_05630 is clustered with a single putative glycosyltransferase (Mlut_05650) and two other genes of unknown function, while Mlut_08960 and Mlut_08980 lie within a large cluster encoding several putative GTs as well as other functions necessary for TUA biosynthesis and export. The individual glycosyl residues of the repeating polymer unit are both derived from UDP-glycosyl donors (43). A likely biosynthetic route to UDP-ManNAcU, predominantly based in this cluster, is apparent. UDP-GlcNAc may be formed by the Mlut_05450 product, a homolog of *E. coli* GlmU (a bifunctional enzyme with glucosamine-1-phosphate *N*-acetyltransferase and GlcNAc-1-phosphate uridyltransferase activities). UDP-GlcNAc could then be transformed to UDP-ManNAc *via* a putative UDP-GlcNAc 2-epimerase encoded by Mlut_09080. The oxidation of this precursor to UDP-ManNAcU could be achieved by either of the UDP-*N*-acetyl-D-mannosaminuronate dehydrogenase homologs encoded by Mlut_08960 or Mlut_08980. Mlut_08720 appears to encode a UDP-Glc 6-dehydrogenase; together with the apparent aminosugar preference of both Mlut_08960 and Mlut_08980, this suggests that there may be greater heterogeneity in TUA biosynthesis than is currently recognized. As both Mlut_08960 and Mlut_08980 appear to be twinned with GT genes as immediate upstream neighbors, these might reflect potentially interchangeable functional units for the introduction of aminosugar-derived glycuronic acid residues into a TUA polymer.

The complement of GTs necessary for TUA biosynthesis can be estimated from biochemical data. A polyprenyl pyrophosphate-GlcNAc-$(ManNAc)_2$ acceptor is synthesized initially and then extended by separable Glc and ManNAcU transferases that intervene alternately in polymer elongation (43, 122). Acceptor synthesis will require TagO (Mlut_08100), a polyprenyl phosphate-dependent GlcNAc phosphate transferase that is implicated universally in the attachment of polymers to PG together with (probably) two ManNAcU transferases. The

only partially characterized GT involved in *M. luteus* TUA biosynthesis is the glucosyltransferase responsible for the deposition of the α-D-Glc residues within the polymer; two apparent subunit types of mass 54 kDa and 52.5 kDa with an estimated pI~5 for the octameric active enzyme were described previously (25). These parameters show a good match with the predicted GT product of Mlut_09020 (molecular mass = 56.8 kDa and pI = 5.94). Finally, the secretion of the TUA may be accomplished by the products of Mlut_09060 and Mlut_09070, which together constitute a predicted ABC family polysaccharide export system.

The Mlut_08990, Mlut_09000 and Mlut_09010 genes that complete this cluster encode products related to aminosugar-*N*-acetyltransferases, a pyridoxal phosphate-dependent aminotransferase and a dehydrogenase, respectively. These activities are consistent with the provision of a further *N*-acetylated amino sugar for glycuronic acid formation. Interestingly Mlut_09000 is a pseudogene; analysis of the *in silico* translated sequence reveals homology with several predicted UDP-4-amino-4-deoxy-L-arabinose-oxoglutarate aminotransferases over the full sequence length with the incorporation of a stop codon through a nucleotide substitution in codon 121. Thus, the structure of this gene cluster is entirely consistent with the biosynthesis of the TUA described for *M. luteus* and taken together with the apparent redundancy in *N*-acetylglycosaminuronate provision and transfer, the apparent loss of Mlut_09000 function suggests a previously more diverse TUA profile.

Energy Metabolism. The *M. luteus* genome contains genes encoding the various respiratory chain components including complex I (NADH-quinone oxidoreductase: Mlut_18970 – Mlut_18940, Mlut_11050, Mlut_11060 & Mlut_05010), complex II (succinate dehydrogenase with cytochrome *b* and a Fe-S cluster: Mlut_04850-04820) and complexes III and IV (quinol-cytochrome *c* oxidoreductase with cytochrome *b* and a Fe-S center and the cytochrome *c* - cytochrome aa_3 oxidase complex: Mlut_12150-12120, Mlut_12210 & Mlut_12220). The *M. luteus* respiratory chain also contains the quinol-oxidase complex with cytochrome *bd* (Mlut_13220 & Mlut_13210), which is usually responsible for the growth of bacteria under low oxygen conditions (128). Genes encoding a membrane-bound malate-menaquinone oxidoreductase (Mlut_08440) and transmembrane L-lactate–menaquinone oxidoreductase (Mlut_21510) are present and the corresponding activities have been demonstrated experimentally in isolated membrane particles (5). The *M. luteus* respiratory chain may also contain a membrane-bound pyruvate dehydrogenase (cytochrome; Mlut_02710), glycerol-3-phosphate dehydrogenase (Mlut_23030), and L-proline dehydrogenase (Mlut_19430) as in *C. glutamicum* (14).

Central carbon metabolism. Genes encoding a complete set of enzymes of the citric acid cycle are present. The genes for the α and β subunits of succinyl–CoA synthetase lie adjacent to each other (Mlut_04280 & Mlut_04270) and genes for the succinate dehydrogenase subunits A-D form a cluster (Mlut_04820-04850). The other genes encoding citrate synthase (Mlut_15490), aconitase (Mlut_13040), isocitrate dehydrogenase (Mlut_04530), the E1 and E2 components of 2-oxoglutarate dehydrogenase (Mlut_07730 & Mlut_13330), fumarase (Mlut_05170) and malate dehydrogenase (Mlut_00950) are located separately.

M. luteus can use acetate as a sole source of carbon for energy and growth (V. Artsatbanov, unpublished results), so it must contain the isocitrate lyase and malate synthase A components of the glyoxylate shunt (and, despite its small genome, must have all the necessary pathways to synthesize amino acids, nucleotides, carbohydrates, and lipids from acetate). The isocitrate lyase and malate synthase A components are functionally active (E. Salina, unpublished results) and they are encoded by adjacent genes (Mlut_02080 & Mlut_02090). Interestingly, the isocitrate lyase is considered as a "persistence factor" in *My. tuberculosis* where it "allows net carbon gain by diverting acetyl-CoA from β-oxidation of fatty acids into the glyoxylate shunt pathway" (74, 112).

All of the main enzymes of glycolysis are present other than the first enzyme, glucokinase, responsible for the phosphorylation of glucose. This is consistent with the observed inability of

M. luteus to grow with glucose as sole carbon source (140), but it may not serve to explain it, since the product of Mlut_13470 is predicted to be a polyphosphate glucokinase. Other closely related species e.g. *R. salmoninarum*, *Arthrobacter* spp. and *Rhodococcus* sp. do have glucokinase genes. *M. luteus* is able to synthesize all glycolytic intermediates (presumably by gluconeogenesis) from phosphoenolpyruvate or pyruvate obtained from oxaloacetate through the agency of phosphoenolpyruvate carboxykinase (Mlut_03380) or pyruvate carboxylase (Mlut_13810). It has a complete complement of enzymes concerned with triose metabolism, and genes for all of the pentose phosphate pathway enzymes are annotated except for 6-phosphogluconolactonase. While this open cycle could produce all intermediates, pentoses would presumably be synthesized via the non-oxidative pentose phosphate pathway that is not coupled to the reduction of $NADP^+$.

Does glycogen serve mainly as a biosynthetic intermediate for trehalose in *M. luteus*? One remarkable feature of some actinobacteria is the importance of the glucose-related storage metabolites, trehalose and glycogen, and the complexity of associated metabolic processes: in some organisms, this area of metabolism appears to be essential for viability (9, 17). There are at least three ways in which mycobacteria and streptomycetes can make trehalose (17) and some interconnections between trehalose and glycogen metabolism (110).

M. luteus has strikingly fewer genes concerned with carbohydrate metabolism than its close relatives (Table 2). All the genes for conventional glycogen biosynthesis from central metabolism are present: *pgm* (Mlut_01060), *glgA* (Mlut_11690), *glgB* (Mlut_03850), *glgC* (Mlut_11680), together with a debranching enzyme gene, *glgX* (Mlut_16760) and the incompletely characterized, glycogen-related gene *glgE* (Mlut_03840). Remarkably, though, there is no obvious candidate gene for glycogen phosphorylase, a highly conserved enzyme that is almost universal and which provides the major route for glycogen breakdown. This raises the question of whether *M. luteus* can metabolize glycogen.

Equally surprising is the absence of the *otsAB* genes for the conventional synthesis of trehalose. However, in the last 10 years, another pathway (present in many actinobacteria) has been discovered, in which the terminal reducing glucosyl residue on chains of α-1,4-linked glucose polymers is “flipped” by the TreY trehalose malto-oligosyl trehalose synthase enzyme so that it is now α-1,1-linked, and then this terminal trehalosyl disaccharide is cleaved off by the TreZ malto-oligosyltrehalose trehalohydrolase enzyme. This pathway is present in *M. luteus* (Mlut_03980, *treY*; Mlut_03990, *treZ*), and could therefore provide a route for glycogen breakdown to give trehalose. In other *Micrococcaceae* greater numbers of trehalose

biosynthetic genes have been annotated (4 in *Arthrobacter* sp FB24, 6 in *K. rhizophila* and 8 in *R. salmoninarum* and *A. aurescens* TC1), and in more distantly related organisms such as *C. glutamicum*, multiple trehalose biosynthetic pathways may be present (139). The degradation of trehalose to yield glucose is catalyzed by trehalase. Although it is generally difficult to recognize trehalases by homology, Mlut_17860 encodes a protein with 35% identity to a trehalase recently characterized in *My. smegmatis* (17). It seems possible that the breakdown of glycogen in *M. luteus* may well take place via trehalose.

Another pathway for trehalose biosynthesis, found in many actinobacteria, uses the trehalose synthase enzyme to interconvert maltose and trehalose – see (110), for references. There is no close homolog of *treS* in *M. luteus*. Thus, the only obvious route for trehalose synthesis appears to be via glycogen. In this respect, it is relevant to note that the *treYZ* genes of *M. luteus* are separated from *glgE* and *glgB* only by a group of genes peculiar to *M. luteus* that were therefore probably acquired by lateral gene transfer. The *glg* and *tre* genes may previously have been adjacent, consistent with the idea of a combined glycogen-trehalose biosynthetic pathway. We are not aware of a comparable system in any other organism.

***M. luteus* has relatively little capacity for secondary metabolism.** The genome of *M. luteus* includes only 39 genes (2.2%) annotated as being concerned with secondary metabolism (Table 2). Among these are 11 clustered genes (Mlut_21170-21270) implicated in carotenoid synthesis (Table 5). Yellow pigmentation has long been an important for the identification of *M. luteus*, suggesting that the genes concerned with pigment production would show a restricted distribution. The phytoene synthetase (Mlut_21230), phytoene desaturase (Mlut_21220) and polyprenyl transferase (Mlut_21210) genes lie in the centre of the cluster and have homologs in many different organisms, including the photosynthetic bacteria (*Synechococcus*), where the carotenoids might function to protect against radiation damage, as has been proposed in *M. luteus* (4). Other genes towards the extremities of the cluster show a more restricted distribution (Table 6). Although homologs of most of these genes are present in other high G+C Gram-positive bacteria, a high level of synteny is restricted to only a few organisms, *viz*: *L. xyli*, *Corynebacterium efficiens*, *C. glutamicum* and *C. michiganensis*.

There is little evidence from the genome annotation for the presence of other secondary metabolic functions. There is a cluster of genes involved in siderophore transport (Mlut_22080 – Mlut_22120) as well as single genes that according to the annotation, might be involved in non-ribosomal peptide synthesis (COG1020) and benzoate catabolism (protocatechuate 3,4 dioxygenase – COG3485). The genome is among the minority of actinobacterial genomes

(including *C. diphtheriae* and *Tropheryma whipplei*) that encode no obvious cytochromes P450 – for example, free-living mycobacteria and streptomycetes generally encode more than ten (130). There is no evidence for a significant repertoire of genes that might be involved in polyketide production (only Mlut_20260 & Mlut_22550) or xenobiotic catabolism. Finally, *M. luteus* uses the non-mevalonate pathway for C5 isoprenoid biosynthesis (Table 7). The genes are dispersed about the bacterial chromosome with Mlut_03780 encoding a bifunctional enzyme comprising IspD (2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase, EC:2.7.7.60) and IspF (2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase EC:4.6.1.12) as is also the case in several other actinobacteria, e.g. *Kytococcus sedentarius*, *C. michiganensis*, *L. xyli* and *T. whipplei* (58%, 48%, 46% & 42% identity, respectively).

Osmotolerance. *M. luteus* is salt tolerant and grows in rich medium, such as nutrient broth, containing 10% NaCl, although pigment production is abolished at concentrations above 5% NaCl (G. Price and M. Young, unpublished). Heterotrophic bacteria that are salt-requiring or salt-tolerant generally use organic solutes such as amino acids (glutamate, proline), glycine betaine, ectoine, and trehalose for osmotic balance.

When grown in rich media, most heterotrophic bacteria accumulate glycine betaine (46). However, very few heterotrophs are capable of *de novo* synthesis of betaine; they tend to accumulate the compound from the medium (yeast extract is a good source of betaine). Accordingly, *M. luteus* encodes an ABC transporter system for proline/glycine betaine with four components clustered in the genome (Mlut_15720-15750). In addition, two choline/carnitine/betaine transporters have also been annotated (Mlut_01900 & Mlut_16530). All of these genes have homologs throughout the actinobacteria.

Although an ectoine synthetase gene is annotated (Mlut_02920), the remainder of the operon required for production of this molecule is apparently absent. The *M. luteus* gene has greatest similarity to genes found in *Stigmatella aurantiaca* and *Burkholderia ambifaria* (55% and 54% identity, respectively), in which the remaining genes required for ectoine synthesis are also apparently absent. Complete ectoine operons are found in the genomes of many actinobacteria including *Brevibacterium linens* BL2 and *Streptomyces avermitilis*.

As noted above, *M. luteus* has two genes potentially concerned with trehalose production from glycogen: a trehalose malto-oligosyl trehalose synthase (Mlut_03980) and a malto-oligosyltrehalose trehalohydrolase (Mlut_03990). This suggests that of the common compatible solutes, *M. luteus* can only produce glutamate, proline and trehalose (the last of

these only via glycogen) though, like many heterotrophic bacteria, it does have transporters for other compatible osmoprotectants.

Long-chain alkene biosynthesis by *M. luteus*. Four decades ago, two research groups studying a close relative of *M. luteus*, *Sarcina lutea* ATCC 533 (now *Kocuria rhizophila*), reported the biosynthesis of *iso*- and *anteiso*-branched, long-chain (primarily C_{25} to C_{29}) alkenes (1, 2, 127). Although the biosynthetic pathway was postulated to involve decarboxylation and condensation of fatty acids, the underlying biochemistry and genetics of alkene biosynthesis have remained unknown until very recently. *M. luteus* strains have also been shown to produce primarily branched C_{29} monoalkenes; this includes the sequenced strain (based upon analysis by gas chromatography-chemical ionization-time of flight mass spectrometry; H. R. Beller, E. B. Goh, J. D. Keasling, submitted for publication) and strain ATCC 27141 (31).

Recently, Friedman and Rude (international patent application WO 2008/113041) reported that heterologous expression of *oleACD* from a range of bacteria (including *Stenotrophomonas maltophilia*, *Xanthomonas axonopodis*, and *Chloroflexus aggregans*) resulted in long-chain alkene biosynthesis, and *in vitro* studies indicated that the alkenes indeed result from fatty acyl-CoA or fatty acyl-ACP precursors. Beller and co-workers have heterologously expressed the three-gene *M. luteus* cluster comprising *oleA* (Mlut_13230), *oleBC* (a gene fusion; Mlut_13240), and *oleD* (Mlut_13250) in *E. coli* and observed C_{27} and C_{29} alkenes that were not detectable in negative controls with a plasmid lacking any *M. luteus* genes (H. R. Beller, E. B. Goh, J. D. Keasling, submitted for publication). *In vitro* studies with *M. luteus* OleA indicated that it catalyzes decarboxylation and condensation of activated fatty acids (H. R. Beller, E. B. Goh, J. D. Keasling, submitted for publication).

Some close relatives of *M. luteus* that also produce long-chain alkenes, such as *K. rhizophila* and *A. aurescens* TC1, have similar *oleABCD* gene organization (Fig. 6) and a relatively high degree of sequence identity. To illustrate, the OleA, OleBC, and OleD sequences of *K. rhizophila* and *A. aurescens* TC1 share 45-57% identity with those of *M. luteus*. In contrast, *Arthrobacter* sp. FB24, which does not produce alkenes (31), appears to lack an *oleB* gene and has OleA and OleD sequences that share only 25-27% sequence identity with those of *M. luteus*. It is possible that the divergence of *ole* sequences in strain FB24 from those in the other strains discussed here explains its inability to produce long-chain alkenes. In the Gram-negative, alkene-producing bacterium *S. maltophilia*, *oleB* and *oleC* are separate genes, in contrast to the Gram-positive species represented in Fig. 6, but there is still relatively strong

similarity between the *S. maltophilia* and *M. luteus* sequences (e.g., OleA and OleD both share 39% sequence identity between these two organisms).

Although BLASTp searches with *M. luteus* OleA (Mlut_13230) showed best hits to β-ketoacyl-ACP-synthase III (FabH), a key enzyme involved in fatty acid biosynthesis, the true *fabH* in the *M. luteus* genome, Mlut_09310, falls in a cluster of genes critical to the biosynthesis of branched-chain fatty acids, including a putative branched-chain α-keto acid decarboxylase (Mlut_09340), malonyl-CoA:ACP transacylase (*fabD*; Mlut_09320), acyl carrier protein (ACP; Mlut_09300), and β-ketoacyl-ACP-synthase II (*fabF*; Mlut_09290).

Dormancy.

M. luteus can enter a profoundly dormant state from which it can be resuscitated by a secreted protein called resuscitation-promoting factor (Rpf) (85). Amongst its closest relatives, *Arthrobacter* sp. FB24, *R. salmoninarum* and *K. rhizophila*, all encode a secreted protein with an N-terminal transglycosylase-like domain and a C-terminal LysM domain that closely resembles *M. luteus* Rpf (123, 137). A second protein belonging to the highly conserved RpfB family (99) is encoded by genes found in *Arthrobacter* sp. FB24 and *R. salmoninarum*. RpfB is also present in *Arthrobacter aurescens* TC1, which does not encode a protein with a similar domain structure to that of *M. luteus* Rpf.

Dormancy has been extensively studied in *My. tuberculosis*, which has five *rpf* genes involved in controlling culturability and resuscitation (12, 27, 86, 102, 113, 129) and the remainder of this section will focus on a comparison with this organism. In *My. tuberculosis*, a state of growth arrest often referred to as dormancy is occasioned by hypoxia *in vitro* (135). This leads to the expression of a cohort of ca. 50 genes (the Dos regulon) under the control of the *devRS* (*dosRS*) two-component system (101). Microarray studies have been accomplished using five different dormancy models and many genes belonging to the Dos regulon are up-regulated in these datasets (see supplementary material – Table 2S). Throughout these datasets, between 45% and 50% of the genes that are significantly up-regulated have homologs in *M. luteus* (Table 8). Several of the *M. luteus* genes in these lists have multiple paralogs in *My. tuberculosis*, the most striking example being Mlut_01830. This gene encodes the universal stress protein UspA. It is represented once in the *M. luteus* genome, whereas there are six homologs in *My. tuberculosis* (Rv1996, Rv2005c, Rv2028c, Rv2623, Rv2624c and Rv3134c). Out of a cohort of 17 genes that are up-regulated in all five *My. tuberculosis* dormancy models in Table 2S, nine have homologs in *M. luteus* (Table 9). Notable among these are genes

encoding the universal stress protein UspA (Mlut_01830), ferredoxin (Mlut_15510), an erythromycin esterase homolog (Mlut_05460), an Hsp20 family heat shock protein (Mlut_16210) and a zinc metalloprotease (Mlut_11840) that lies within a cluster including putative proteasome components. This cluster shows a high degree of synteny with *Arthrobacter* sp. FB24.

Recent studies of the enduring hypoxic response of *My. tuberculosis* (103) revealed that more than two hundred genes are up-regulated, only five of which belong to the Dos regulon, and only two of these five were up-regulated more than threefold (103). However, comparison of the 47 genes up-regulated more than 3-fold in the enduring hypoxic response with up-regulated genes from the other 5 dormancy models (Table 8) showed that 22 of them (47%) were up-regulated more than 3-fold in at least one of those models. 62% of these 47 genes have homologs in the *M. luteus* genome (Table 3S).

M. luteus therefore contains many genes similar to genes up-regulated in different *My. tuberculosis* dormancy models. Among them are members of the Dos regulon, including the *devRS* (*dosRS*) genes encoding the sensory histidine kinase and response regulator that control the regulon (possibly Mlut_18530 & Mlut_18540, although other gene pairs such as Mlut_16250 & Mlut_16240 or Mlut_21850 & Mlut_21860 might fulfill this role). The total number of *M. luteus* dormancy-related genes revealed by these various comparisons is roughly proportional to the two-fold difference in genome size between *M. luteus* and *My. tuberculosis*. The elevated number of dormancy-related genes in *My. tuberculosis* is accounted for, in part at least, by the presence of multiple paralogs that do not exist in *M. luteus*, indicating that the dormancy machinery of *M. luteus* is highly minimized in comparison with that of *My. tuberculosis,* though it clearly remains fully functional (24, 27, 49, 54, 56, 113, 131, 132, 134).

Conclusion. The *M. luteus* genome is very small compared with those of other free-living actinobacteria, raising speculation that this may be connected with both its simple morphology and a restricted ecology. Soil-dwelling organisms typically have a substantial capability for environmental responses mediated through two-component systems and sigma factors, but the *M. luteus* genome encodes only fourteen response regulators and four sigma factors, indicative of adaptation to a rather strict ecological niche. We therefore speculate that its primary adaptation is to (mammalian?) skin, where it is often found, and that its occasional presence elsewhere (water, soil) might possibly arise from contamination by skin flakes. The somewhat minimized nature of the genome may also provide opportunities to evaluate the roles of conserved genes that, in other actinobacteria, are members of substantial paralogous families.

Despite its small size, the *M. luteus* genome contains an exceptionally high number of transposable elements. These do not seem to have resulted in large-scale genome rearrangements, since the genome sequences for the phylogenetically closest actinobacteria available show only one or two multigene inversions spanning the *oriC* region, and none that do not include *oriC,* when compared with *M. luteus*. The possibility remains, however, that some of the transposable elements could have played a part in the contraction of the *M. luteus* genome from a larger ancestral version. Such elements are present at a number of points of discontinuity between the genomes of *M. luteus* and its closest characterized relatives (not shown).

The simple morphology of *M. luteus* is reflected in the absence of nearly all genes known to be concerned with developmental decisions in more complex actinobacteria, and the confinement of genes for cell division and cell wall biosynthesis to a minimal set. The only obvious exception to this is the presence of orthologues of the *Streptomyces* sporulation genes *whiA* and *whiB*, but this is not surprising since *whiA* orthologues are found in virtually all Gram-positive bacteria, both firmicutes and actinobacteria, and *whiB* orthologues are nearly universal among actinobacteria. The roles of these two genes in such a simple organism as *M. luteus* merit exploration. The presence of only a single class A PBP, and only two of class B, may be connected with the adoption of spherical morphology, and with high sensitivity to β-lactams and lysozyme. The well-known sensitivity of *M. luteus* to diverse other antibiotics may possibly be due to its lack of a *wblC* gene: this gene confers increased resistance to a wide range of antibiotics and other inhibitors on streptomycetes and mycobacteria, apparently by affecting the expression of a large number of genes that include many predicted to affect resistance (80).

1 ACKNOWLEDGEMENT

2 We are grateful to the British Council, the Israeli Ministry of Science and Technology and the
3 Hebrew University for funding a workshop on the *M. luteus* genome held in Jerusalem 13th –
4 18th April, 2008. For HRB and EBG, this work was part of the DOE Joint BioEnergy Institute
5 (http://www.jbei.org) supported by the U. S. Department of Energy, Office of Science, Office
6 of Biological and Environmental Research, through contract DE-AC02-05CH11231 between
7 Lawrence Berkeley National Laboratory and the U. S. Department of Energy. MY thanks the
8 UK BBSRC for financial support and VA and ASK thank the MCB RAS program for financial
9 support.

FIGURE LEGENDS

FIG. 1. Circular representation of the *M. luteus* chromosome. Genome coordinates are given in Mbp. From outside to inside, the various circles represent: genes on the forward strand; genes on the reverse strand, RNA genes (tRNAs green, rRNAs red, other RNAs black), GC content and GC skew. Genes are color coded according to their COG category.

The color code of function category for top COG hit is shown below.

COG Code	COG Function Definition
[A]	RNA processing and modification
[B]	Chromatin structure and dynamics
[C]	Energy production and conversion
[D]	Cell cycle control, cell division, chromosome partitioning
[E]	Amino acid transport and metabolism
[F]	Nucleotide transport and metabolism
[G]	Carbohydrate transport and metabolism
[H]	Coenzyme transport and metabolism
[I]	Lipid transport and metabolism
[J]	Translation, ribosomal structure and biogenesis
[K]	Transcription
[L]	Replication, recombination and repair
[M]	Cell wall/membrane/envelope biogenesis
[N]	Cell motility
[O]	Posttranslational modification, protein turnover, chaperones
[P]	Inorganic ion transport and metabolism
[Q]	Secondary metabolites biosynthesis, transport and catabolism
[R]	General function prediction only
[S]	Function unknown
[T]	Signal transduction mechanisms

[U]	Intracellular trafficking, secretion, and vesicular transport
[V]	Defense mechanisms
[W]	Extracellular structures
[Y]	Nuclear structure
[Z]	Cytoskeleton
NA	Not assigned

FIG. 2. Percentage of genes assigned to different COG categories in *M. luteus* and related organisms.

FIG. 3. Synteny between actinobacterial genomes. For each genome the first gene is *dnaA*, except in the case of the linear *S. coelicolor* genome, in which *dnaA* is located centrally. Each dot represents a reciprocal best match (BLASTp) between proteins in the genomes being compared. Dots are positioned according to their genome locations. See Methods for further details. Abbreviations: Mlu, *Micrococcus luteus*; Krh, *Kocuria rhizophila* (123); Art, *Arthrobacter* sp. strain FB24; Cmm, *Clavibacter michiganensis* subsp. *michiganensis* (32); Mtb, *Mycobacterium tuberculosis* (22); Sco, *Streptomyces coelicolor* (11); Rsa, *Renibacterium salmoninarum* (137).

FIG. 4. Proposed integrated elements (IEs) in *M. luteus* Fleming. Block arrows containing numbers represent ORFs and their %GC values. The proposed function of the gene products is shown where predictions from database searches are informative. All four of the proposed IEs are within regions of lower than average %GC for *M. luteus* and three of the elements (IEMlut1, IEMlut2 & IEMlut4) interrupt regions with good synteny with *Arthrobacter*. The ORFs are colored as follows: Brown indicates synteny of gene order with *Arthrobacter*; grey indicates that the gene product might be involved in plasmid replication or transfer; green is a transposase or fragment thereof; red is used to highlight the putative metal resistance genes. A and B. IEMlut1 (approximate coordinates 11840-72798) and IEMlut2 (approximate coordinates 672329-680904) may have been integrated via the serine integrases, Mlut_00100 and Mlut_06210, respectively. The putative DnaK, GrpE, DnaJ and the ClpB-like chaperone

(Mlut_00560 - Mlut00580, Mlut_00600) have been included in IEMlut1 as they appear to have been acquired horizontally. Their closest relatives are not the paralogous genes on the *M. luteus* chromosome (Mlut_11810, Mlut_11800, Mlut_11790, Mlut_18660) but genes from other actinomycetes such as *Streptomyces* sp, *Catenulispora* and *Gordonia*. On the other hand, the closest relatives of Mlut_11790, Mlut_11800 and Mlut_18660 are from the phylogenetically close *Arthrobacter* and *Kocuria*. IEMlut2 appears to have inserted into a putative oxidoreductase to yield two gene fragments, Mlut_06130 and Mlut_06220 (yellow). C and D. IEMlut3 (approximate coordinates 695571-717542) and IEMlut4 (approximate coordinates 2,223,379-2,238,868) may have integrated via the action of the conserved triplet of genes that includes two tyrosine recombinases (closest homologues are either purple, Mlut_06600 and Mlut_20690, or light blue, Mlut_06590 and Mlut_20700) and a conserved hypothetical (CH; colored blue-green) (Mlut_06610 & Mlut_20680).

FIG. 5. Conservation of Division/Cell Wall (DCW) clusters. The DCW clusters of several bacteria are schematically represented. Coding sequences are not drawn to scale, in order to facilitate alignment. The triangles denote the positions of single gene insertions unless numerals are present that indicate the insertion of multiple genes. Where orthologues are absent from the cluster but retained at a locus nearby, they are placed to one side. The key to the gene symbols placed above each cluster are as follows: Z, *mraZ*; W, *mraW*; L, *ftsL*; I, *ftsI*; I', *spoVD*; E, *murE*; Y, *mraY*; D, *murD*; Fw, *ftsW*; G, *murG*; C, *murC*; B, *murB*, Dd, *ddlB*; Q, *ftsQ*; A, *ftsA*; Fz, *ftsZ*.

FIG. 6. Organization of the *ole* (olefin synthesis) genes in *M. luteus* and other bacteria.

REFERENCES

1. **Albro, P. W.** 1971. Confirmation of the identification of the major C-29 hydrocarbons of *Sarcina lutea*. J Bacteriol **108:**213-8.
2. **Albro, P. W., and J. C. Dittmer.** 1969. The biochemistry of long-chain, nonisoprenoid hydrocarbons. I. Characterization of the hydrocarbons of *Sarcina lutea* and the isolation of possible intermediates of biosynthesis. Biochemistry **8:**394-404.
3. **Altschul, S. F., T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research **25:**3389-3402.
4. **Anwar, M., T. H. Khan, J. Prebble, and P. F. Zagalsky.** 1977. Membrane-bound carotenoid in *Micrococcus luteus* protects naphthoquinone from photodynamic action. Nature **270:**538-40.
5. **Artsatbanov, V., G. V. Tikhonova, and D. N. Ostrovskii.** 1983. [Generation of membrane potential by aerobic bacteria *Micrococcus lysodeikticus*. Correlation between coupled and uncoupled respiration]. Biokhimiia **48:**1568-79.
6. **Av-Gay, Y., and M. Everett.** 2000. The eukaryotic-like Ser/Thr protein kinases of *Mycobacterium tuberculosis*. Trends Microbiol **8:**238-44.
7. **Bacon, J., B. W. James, L. Wernisch, A. Williams, K. A. Morley, G. J. Hatch, J. A. Mangan, J. Hinds, N. G. Stoker, P. D. Butcher, and P. D. Marsh.** 2004. The influence of reduced oxygen availability on pathogenicity and gene expression in *Mycobacterium tuberculosis*. Tuberculosis (Edinb) **84:**205-17.
8. **Batzoglou, S., D. B. Jaffe, K. Stanley, J. Butler, S. Gnerre, E. Mauceli, B. Berger, J. P. Mesirov, and E. S. Lander.** 2002. ARACHNE: a whole-genome shotgun assembler. Genome Res **12:**177-89.
9. **Belanger, A. E., and G. F. Hatfull.** 1999. Exponential-phase glycogen recycling is essential for growth of *Mycobacterium smegmatis*. J Bacteriol **181:**6670-8.
10. **Bendtsen, J. D., H. Nielsen, G. von Heijne, and S. Brunak.** 2004. Improved prediction of signal peptides: SignalP 3.0. J Mol Biol **340:**783-95.
11. **Bentley, S. D., K. F. Chater, A.-M. Cerdeno-Tarraga, G. L. Challis, N. R. Thomson, K. D. James, D. E. Harris, M. A. Quail, H. Kieser, D. Harper, A. Bateman, S. Brown, G. Chandra, C. W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T. Hornsby, S. Howarth, C.-H. Huang, T. Kieser, L. Larke, L. Murphey, K. Oliver, S. O'Neil, E. Rabbinowitsch, M.-A. Rajandream, K. Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B. G. Barrell, J. Parkhill, and D. A. Hopwood.** 2002. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). Nature **417:**141-147.
12. **Biketov, S., V. Potapov, E. Ganina, K. Downing, B. D. Kana, and A. Kaprelyants.** 2007. The role of resuscitation promoting factors in pathogenesis and reactivation of *Mycobacterium tuberculosis* during intra-peritoneal infection in mice. BMC Infect Dis **7:**146.
13. **Boitel, B., M. Ortiz-Lombardia, R. Duran, F. Pompeo, S. T. Cole, C. Cervenansky, and P. M. Alzari.** 2003. PknB kinase activity is regulated by phosphorylation in two Thr residues and dephosphorylation by PstP, the cognate phospho-Ser/Thr phosphatase, in *Mycobacterium tuberculosis*. Mol Microbiol **49:**1493-508.
14. **Bott, M., and A. Niebisch.** 2003. The respiratory chain of *Corynebacterium glutamicum*. J Biotechnol **104:**129-53.

15. **Brewster, J. M., and E. A. Morgan.** 1981. Tn9 and IS1 inserts in a ribosomal ribonucleic acid operon of *Escherichia coli* are incompletely polar. J Bacteriol **148:**897-903.
16. **Burrus, V., J. Marrero, and M. K. Waldor.** 2006. The current ICE age: biology and evolution of SXT-related integrating conjugative elements. Plasmid **55:**173-83.
17. **Carroll, J. D., I. Pastuszak, V. K. Edavana, Y. T. Pan, and A. D. Elbein.** 2007. A novel trehalase from *Mycobacterium smegmatis* - purification, properties, requirements. FEBS J **274:**1701-14.
18. **Cha, J. H., and G. C. Stewart.** 1997. The *divIVA* minicell locus of *Bacillus subtilis*. J Bacteriol **179:**1671-83.
19. **Chandler, M., and J. Mahillon.** 2002. Insertion sequences revisited, p. 305-366. *In* N. L. Craig, R. Craigie, M. Gellert, and A. M. Lambowitz (ed.), Mobile DNA II. ASM Press, Washington, D. C.
20. **Chen, H. H., G. Z. Zhao, D. J. Park, Y. Q. Zhang, L. H. Xu, J. C. Lee, C. J. Kim, and W. J. Li.** 2009. *Micrococcus endophyticus* sp. nov., isolated from surface-sterilized *Aquilaria sinensis* roots. Int J Syst Evol Microbiol **59:**1070-5.
21. **Chopra, P., B. Singh, R. Singh, R. Vohra, A. Koul, L. S. Meena, H. Koduri, M. Ghildiyal, P. Deol, T. K. Das, A. K. Tyagi, and Y. Singh.** 2003. Phosphoprotein phosphatase of *Mycobacterium tuberculosis* dephosphorylates serine-threonine kinases PknA and PknB. Biochem Biophys Res Commun **311:**112-20.
22. **Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornby, K. Jagels, A. Krogh, J. McLean, S. Moule, L. Murphy, K. Oliver, J. Osborne, M. A. Quail, M. A. Rajandream, J. Rogers, S. Rutter, K. Seeger, J. Skelton, R. Squares, S. Squares, J. E. Sulston, K. Taylor, S. Whitehead, and B. G. Barrell.** 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature **393:**537-544.
23. **Cole, S. T., K. Eiglmeier, J. Parkhill, K. D. James, N. R. Thomson, P. R. Wheeler, N. Honore, T. Garnier, C. Churcher, D. Harris, K. Mungall, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. M. Davies, K. Devlin, S. Duthoy, T. Feltwell, A. Fraser, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, C. Lacroix, J. Maclean, S. Moule, L. Murphy, K. Oliver, M. A. Quail, M. A. Rajandream, K. M. Rutherford, S. Rutter, K. Seeger, S. Simon, M. Simmonds, J. Skelton, R. Squares, S. Squares, K. Stevens, K. Taylor, S. Whitehead, J. R. Woodward, and B. G. Barrell.** 2001. Massive gene decay in the leprosy bacillus. Nature **409:**1007-1011.
24. **Daniel, J., C. Deb, V. S. Dubey, T. D. Sirakova, B. Abomoelak, H. R. Moribondi, and P. E. Kolattukudy.** 2004. Induction of a novel class of diacyl glycerol transferases and triacylglycerol accumulation in *Mycobacterium tuberculosis* as it goes into a dormancy-like state in culture. J Bacteriol. **186:**5017-5030.
25. **Deng, L., and J. S. Anderson.** 1997. Biosynthesis of teichuronic acid in the bacterial cell wall. Purification and characterization of the glucosyltransferase of *Micrococcus luteus*. J Biol Chem **272:**479-85.
26. **Doddamani, H. P., and H. Z. Ninnekar.** 2001. Biodegradation of carbaryl by a *Micrococcus* species. Curr Microbiol **43:**69-73.
27. **Downing, K. J., V. V. Mischenko, M. O. Shleeva, D. I. Young, M. Young, A. S. Kaprelyants, A. S. Apt, and V. Mizrahi.** 2005. Mutants of *Mycobacterium tuberculosis* lacking three of the five *rpf*-like genes are defective for growth *in vivo* and for resuscitation *in vitro*. Infection and Immunity **73:**3038-3043.

28. **Erickson, S. K., and G. L. Parker.** 1969. The electron-transport system of *Micrococcus lutea* (*Sarcina lutea*). Biochim Biophys Acta **180:**56-62.
29. **Fleming, A.** 1922. On a remarkable bacteriolytic substance found in secretions and tissues. Proceedings of the Royal Society of London Series B **93:**306-317.
30. **Fleming, A., and V. D. Allison.** 1922. Further observations on a bacteriolytic element found in tissues and secretions. Proceedings of the Royal Society of London Series B **94:**142-151.
31. **Frias, J. A., J. E. Richman, and L. P. Wackett.** 2009. C_{29} olefinic hydrocarbons biosynthesized by *Arthrobacter* species. Appl Environ Microbiol **75:**1774-7.
32. **Gartemann, K. H., B. Abt, T. Bekel, A. Burger, J. Engemann, M. Flugel, L. Gaigalat, A. Goesmann, I. Grafen, J. Kalinowski, O. Kaup, O. Kirchner, L. Krause, B. Linke, A. McHardy, F. Meyer, S. Pohle, C. Ruckert, S. Schneiker, E. M. Zellermann, A. Puhler, R. Eichenlaub, O. Kaiser, and D. Bartels.** 2008. The genome sequence of the tomato-pathogenic actinomycete *Clavibacter michiganensis* subsp. *michiganensis* NCPPB382 reveals a large island involved in pathogenicity. J Bacteriol **190:**2138-49.
33. **Gassner, G. T., J. P. Dickie, D. A. Hamerski, J. K. Magnuson, and J. S. Anderson.** 1990. Teichuronic acid reducing terminal N-acetylglucosamine residue linked by phosphodiester to peptidoglycan of *Micrococcus luteus*. Journal of Bacteriology **172:**2273-2279.
34. **Gel'man, N. S., G. V. Tikhonova, I. M. Simakova, M. A. Lukoyanova, S. D. Taptykova, and H. M. Mikelsaar.** 1970. Fragmentation by detergents of the respiratory chain of *Micrococcus lysodeikticus* membranes. Biochim Biophys Acta **223:**321-31.
35. **Ghuysen, J. M.** 1991. Serine beta-lactamases and penicillin-binding proteins. Annu Rev Microbiol **45:**37-67.
36. **Gibson, K. J., L. Eggeling, W. N. Maughan, K. Krumbach, S. S. Gurcha, J. Nigou, G. Puzo, H. Sahm, and G. S. Besra.** 2003. Disruption of Cg-Ppm1, a polyprenyl monophosphomannose synthase, and the generation of lipoglycan-less mutants in *Corynebacterium glutamicum*. J Biol Chem **278:**40842-50.
37. **Goffin, C., and J. M. Ghuysen.** 1998. Multimodular penicillin-binding proteins: an enigmatic family of orthologs and paralogs. Microbiol Mol Biol Rev **62:**1079-93.
38. **Gomez, J. E., and W. R. Bishai.** 2000. *whmD* is an essential mycobacterial gene required for proper septation and cell division. Proc. Natl. Acad. Sci. USA. **97:**8554-8559.
39. **Greenblatt, C. L., J. Baum, B. Y. Klein, S. Nachshon, V. Koltunov, and R. J. Cano.** 2004. *Micrococcus luteus* - Survival in amber. Microbial Ecology **48:**120-127.
40. **Griffiths-Jones, S., S. Moxon, M. Marshall, A. Khanna, S. R. Eddy, and A. Bateman.** 2005. Rfam: annotating non-coding RNAs in complete genomes. Nucleic Acids Res **33:**D121-4.
41. **Gurcha, S. S., A. R. Baulard, L. Kremer, C. Locht, D. B. Moody, W. Muhlecker, C. E. Costello, D. C. Crick, P. J. Brennan, and G. S. Besra.** 2002. Ppm1, a novel polyprenol monophosphomannose synthase from *Mycobacterium tuberculosis*. Biochem J **365:**441-50.
42. **Hase, S., and Y. Matsushima.** 1972. Structural studies on a glucose-containing polysaccharide obtained from cell walls of *Micrococcus lysodeikticus*. 3. Determination of the structure. J Biochem **72:**1117-28.
43. **Hildebrandt, K. M., and J. S. Anderson.** 1990. Biosynthetic elongation of isolated teichuronic acid polymers via glucosylsyltransferase and N-

acetylmannosaminuronosyltransferases from solubilized cytoplasmic membrane fragments of *Micrococcus luteus*. Journal of Bacteriology **172:**5160-5164.

44. **Höltje, J. V.** 1996. A hypothetical holoenzyme involved in the replication of the murein sacculus of *Escherichia coli*. Microbiology **142:**1911-8.

45. **Hutchings, M. I.** 2007. Unusual two-component signal transduction pathways in the actinobacteria. Adv Appl Microbiol **61:**1-26.

46. **Imhoff, J. F., and F. Rodriguez-Valera.** 1984. Betaine is the main compatible solute of halophilic eubacteria. J Bacteriol **160:**478-9.

47. **Johnson, S. D., K. P. Lacher, and J. S. Anderson.** 1981. Carbon-13 nuclear magnetic resonance spectroscopic study of teichuronic acid from *Micrococcus luteus* cell walls. Comparison of the polysaccharide isolated from cells with that synthesized in vitro. Biochemistry **20:**4781-5.

48. **Joseleau-Petit, D., D. Thevenet, and R. D'Ari.** 1994. ppGpp concentration, growth without PBP2 activity, and growth-rate control in *Escherichia coli*. Mol Microbiol **13:**911-7.

49. **Kana, B. D., B. G. Gordhan, K. J. Downing, N. Sung, G. Vostroktunova, E. E. Machowski, L. Tsenova, M. Young, A. Kaprelyants, G. Kaplan, and V. Mizrahi.** 2008. The resuscitation-promoting factors of *Mycobacterium tuberculosis* are required for virulence and resuscitation from dormancy but are collectively dispensable for growth *in vitro*. Mol Microbiol **67:**672-84.

50. **Kang, C. M., D. W. Abbott, S. T. Park, C. C. Dascher, L. C. Cantley, and R. N. Husson.** 2005. The *Mycobacterium tuberculosis* serine/threonine kinases PknA and PknB: substrate identification and regulation of cell shape. Genes Dev **19:**1692-704.

51. **Kano, A., Y. Andachi, T. Ohama, and S. Osawa.** 1991. Novel anticodon composition of transfer RNAs in *Micrococcus luteus*, a bacterium with a high genomic G + C content. Correlation with codon usage. J Mol Biol **221:**387-401.

52. **Kano, A., T. Ohama, R. Abe, and S. Osawa.** 1993. Unassigned or nonsense codons in *Micrococcus luteus*. J Mol Biol **230:**51-6.

53. **Kaprelyants, A. S., J. C. Gottschal, and D. B. Kell.** 1993. Dormancy in non-sporulating bacteria. FEMS Microbiol Rev **104:**271-286.

54. **Kaprelyants, A. S., and D. B. Kell.** 1993. Dormancy in stationary-phase cultures of *Micrococcus luteus*: flow cytometric analysis of starvation and resuscitation. Appl Environ Microbiol **59:**3187-3196.

55. **Kaprelyants, A. S., G. V. Mukamolova, and D. B. Kell.** 1994. Estimation of dormant *Micrococcus luteus* cells by penicillin lysis and by resuscitation in cell-free spent medium at high dilution. FEMS Microbiol Lett **115:**347-352.

56. **Karakousis, P. C., T. Yoshimatsu, G. Lamichhane, S. C. Woolwine, E. L. Nuermberger, J. Grosset, and W. R. Bishai.** 2004. Dormancy phenotype displayed by extracellular *Mycobacterium tuberculosis* within artificial granulomas in mice. J Exp Med **200:**647-57.

57. **Kato, J., H. Suzuki, and Y. Hirota.** 1985. Dispensability of either penicillin-binding protein-1a or -1b involved in the essential process for cell elongation in *Escherichia coli*. Mol Gen Genet **200:**272-7.

58. **Kim, T. H., J. S. Park, H. J. Kim, Y. Kim, P. Kim, and H. S. Lee.** 2005. The *whcE* gene of *Corynebacterium glutamicum* is important for survival following heat and oxidative stress. Biochem Biophys Res Commun **337:**757-64.

59. **Kocur, M., Z. Paova, and T. Martinec.** 1972. Taxonomic status of *Micrococcus luteus* (Schroeter 1872) Cohn 1872, and designation of the neotype strain. International Journal of Systematic Bacteriology **22:**218-223.

60. **Konstantinidis, K. T., and J. M. Tiedje.** 2004. Trends between gene content and genome size in prokaryotic species with larger genomes. Proc Natl Acad Sci U S A **101:**3160-5.
61. **Kowal, A. K., and J. S. Oliver.** 1997. Exploiting unassigned codons in *Micrococcus luteus* for tRNA-based amino acid mutagenesis. Nucleic Acids Res **25:**4685-9.
62. **Krogh, A., B. Larsson, G. von Heijne, and E. L. Sonnhammer.** 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol **305:**567-80.
63. **Krubasik, P., M. Kobayashi, and G. Sandmann.** 2001. Expression and functional analysis of a gene cluster involved in the synthesis of decaprenoxanthin reveals the mechanisms for C_{50} carotenoid formation. Eur J Biochem **268:**3702-8.
64. **Lennarz, W. J., and B. Talamo.** 1966. The chemical characterization and enzymatic synthesis of mannolipids in *Micrococcus lysodeikticus*. J Biol Chem **241:**2707-19.
65. **Letek, M., E. Ordonez, J. Vaquera, W. Margolin, K. Flardh, L. M. Mateos, and J. A. Gil.** 2008. DivIVA is required for polar growth in the MreB-lacking rod-shaped actinomycete *Corynebacterium glutamicum*. J Bacteriol **190:**3283-92.
66. **Levchenko, L. A., A. P. Sadkov, N. V. Lariontseva, E. M. Koldasheva, A. K. Shilova, and A. E. Shilov.** 2002. Gold helps bacteria to oxidize methane. J Inorg Biochem **88:**251-3.
67. **Levchenko, L. A., A. P. Sadkov, S. A. Marakushev, and N. V. Lariontseva.** 1997. Participation of biological membranes in colloidal gold transformation by *Micrococcus luteus* cells. Membr Cell Biol **11:**131-5.
68. **Liolios, K., K. Mavromatis, N. Tavernarakis, and N. C. Kyrpides.** 2008. The Genomes On Line Database (GOLD) in 2007: status of genomic and metagenomic projects and their associated metadata. Nucleic Acids Res **36:**D475-9.
69. **Liu, H., Y. Xu, Y. Ma, and P. Zhou.** 2000. Characterization of *Micrococcus antarcticus* sp. nov., a psychrophilic bacterium from Antarctica. International Journal of Systematic and Evolutionary Microbiology **50:**715-9.
70. **Liu, X. Y., B. J. Wang, C. Y. Jiang, and S. J. Liu.** 2007. *Micrococcus flavus* sp. nov., isolated from activated sludge in a bioreactor. Int J Syst Evol Microbiol **57:**66-9.
71. **Mahillon, J., and M. Chandler.** 1998. Insertion sequences. Microbiology and Molecular Biology Reviews **62:**725-774.
72. **Markowitz, V. M., K. Mavromatis, N. N. Ivanova, I. M. Chen, K. Chu, and N. C. Kyrpides.** 2009. IMG ER: a system for microbial genome annotation expert review and curation. Bioinformatics **25:**2271-8.
73. **Markowitz, V. M., E. Szeto, K. Palaniappan, Y. Grechkin, K. Chu, I. M. Chen, I. Dubchak, I. Anderson, A. Lykidis, K. Mavromatis, N. N. Ivanova, and N. C. Kyrpides.** 2008. The integrated microbial genomes (IMG) system in 2007: data content and analysis tool extensions. Nucleic Acids Res **36:**D528-33.
74. **McKinney, J. D., K. H. zu Bentrup, E. J. Munoz-Elias, A. Miczak, B. Chen, W. T. Chan, D. Swenson, J. C. Sacchettini, W. R. Jacobs, and D. G. Russell.** 2000. Persistence of *Mycobacterium tuberculosis* in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase. Nature **406:**735-738.
75. **Mingorance, J., J. Tamames, and M. Vicente.** 2004. Genomic channeling in bacterial cell division. J Mol Recognit **17:**481-7.
76. **Mishra, A. K., L. J. Alderwick, D. Rittmann, R. V. Tatituri, J. Nigou, M. Gilleron, L. Eggeling, and G. S. Besra.** 2007. Identification of an alpha(1-->6) mannopyranosyltransferase (MptA), involved in *Corynebacterium glutamicum* lipomanann biosynthesis, and identification of its orthologue in *Mycobacterium tuberculosis*. Mol Microbiol **65:**1503-17.

77. **Mishra, A. K., L. J. Alderwick, D. Rittmann, C. Wang, A. Bhatt, W. R. Jacobs, Jr., K. Takayama, L. Eggeling, and G. S. Besra.** 2008. Identification of a novel alpha(1-->6) mannopyranosyltransferase MptB from *Corynebacterium glutamicum* by deletion of a conserved gene, NCgl1505, affords a lipomannan- and lipoarabinomannan-deficient mutant. Mol Microbiol **68:**1595-613.
78. **Mongodin, E. F., N. Shapir, S. C. Daugherty, R. T. DeBoy, J. B. Emerson, A. Shvartzbeyn, D. Radune, J. Vamathevan, F. Riggs, V. Grinberg, H. Khouri, L. P. Wackett, K. E. Nelson, and M. J. Sadowsky.** 2006. Secrets of soil survival revealed by the genome sequence of *Arthrobacter aurescens* TC1. PLoS Genet **2:**e214.
79. **Morgan, E. A.** 1980. Insertions of Tn*10* into an *E. coli* ribosomal RNA operon are incompletely polar. Cell **21:**257-65.
80. **Morris, R. P., L. Nguyen, J. Gatfield, K. Visconti, K. Nguyen, D. Schnappinger, S. Ehrt, Y. Liu, L. Heifets, J. Pieters, G. Schoolnik, and C. J. Thompson.** 2005. Ancestral antibiotic resistance in *Mycobacterium tuberculosis*. Proc Natl Acad Sci USA **102:**12200-5.
81. **Mukamolova, G. V., A. S. Kaprelyants, D. I. Young, M. Young, and D. B. Kell.** 1998. A bacterial cytokine. Proc Natl Acad Sci USA **95:**8916-8921.
82. **Mukamolova, G. V., S. S. Kormer, D. B. Kell, and A. S. Kaprelyants.** 1999. Stimulation of the multiplication of *Micrococcus luteus* by an autocrine growth factor. Arch Microbiol **172:**9-14.
83. **Mukamolova, G. V., S. S. Kormer, N. D. Yanopolskaya, and A. S. Kaprelyants.** 1995. Properties of dormant cells in stationary-phase cultures of *Micrococcus luteus* during prolonged incubation. Mikrobiologia **64:**284-288.
84. **Mukamolova, G. V., A. G. Murzin, E. G. Salina, G. R. Demina, D. B. Kell, A. S. Kaprelyants, and M. Young.** 2006. Muralytic activity of *Micrococcus luteus* Rpf and its relationship to physiological activity in promoting bacterial growth and resuscitation. Mol Microbiol **59:**84-98.
85. **Mukamolova, G. V., O. A. Turapov, K. Kazaryan, M. Telkov, A. S. Kaprelyants, D. B. Kell, and M. Young.** 2002. The *rpf* gene of *Micrococcus luteus* encodes an essential secreted growth factor. Mol Microbiol **46:**611-621.
86. **Mukamolova, G. V., O. A. Turapov, D. I. Young, A. S. Kaprelyants, D. B. Kell, and M. Young.** 2002. A family of autocrine growth factors in *Mycobacterium tuberculosis*. Mol Microbiol **46:**623-635.
87. **Mukamolova, G. V., N. D. Yanopolskaya, D. B. Kell, and A. S. Kaprelyants.** 1998. On resuscitation from the dormant state of *Micrococcus luteus*. Antonie van Leeuwenhoek **73:**237-243.
88. **Mullany, P., A. P. Roberts, and H. Wang.** 2002. Mechanism of integration and excision in conjugative transposons. Cell Mol Life Sci **59:**2017-22.
89. **Munoz, E., J. H. Freer, D. J. Ellar, and M. R. Salton.** 1968. Membrane-associated ATPase activity from *Micrococcus lysodeikticus*. Biochim Biophys Acta **150:**531-3.
90. **Murayama, O., M. Matsuda, and J. E. Moore.** 2003. Studies on the genomic heterogeneity of *Micrococcus luteus* strains by macro-restriction analysis using pulsed-field gel electrophoresis. J Basic Microbiol **43:**337-40.
91. **Muttucumaru, D. G., G. Roberts, J. Hinds, R. A. Stabler, and T. Parish.** 2004. Gene expression profile of *Mycobacterium tuberculosis* in a non-replicating state. Tuberculosis (Edinb) **84:**239-46.
92. **Nikolaichik, Y. A., and W. D. Donachie.** 2000. Conservation of gene order amongst cell wall and cell division genes in Eubacteria, and ribosomal genes in Eubacteria and Eukaryotic organelles. Genetica **108:**1-7.

93. **Ostrovskii, D. N., N. A. Pereverzev, I. G. Zhukova, S. M. Trutko, and N. S. Gel'man.** 1968. [Some physico-chemical characteristics of the complex of NAD-H2 and malate dehydrogenases in *Micrococcus lysodeikticus* membranes]. Biokhimiia **33:**319-25.
94. **Pakkiri, L. S., and C. J. Waechter.** 2005. Dimannosyldiacylglycerol serves as a lipid anchor precursor in the assembly of the membrane-associated lipomannan in *Micrococcus luteus*. Glycobiology **15:**291-302.
95. **Parkhill, J., and N. L. Brown.** 1990. Site-specific insertion and deletion mutants in the *mer* promoter-operator region of Tn*501*; the nineteen base-pair spacer is essential for normal induction of the promoter by MerR. Nucleic Acids Res **18:**5157-62.
96. **Peirs, P., L. De Wit, M. Braibant, K. Huygen, and J. Content.** 1997. A serine/threonine protein kinase from *Mycobacterium tuberculosis*. Eur J Biochem **244:**604-12.
97. **Pless, D. D., A. S. Schmit, and W. J. Lennarz.** 1975. The characterization of mannan of *Micrococcus lysodeikticus* as an acidic lipopolysaccharide. J Biol Chem **250:**1319-27.
98. **Pucci, M. J., J. A. Thanassi, L. F. Discotto, R. E. Kessler, and T. J. Dougherty.** 1997. Identification and characterization of cell wall-cell division gene clusters in pathogenic gram-positive cocci. J Bacteriol **179:**5632-5.
99. **Ravagnani, A., C. L. Finan, and M. Young.** 2005. A novel firmicute protein family related to the actinobacterial resuscitation-promoting factors by non-orthologous domain displacement. BMC Genomics **6:**39.
100. **Reddy, P. S., A. Raghavan, and D. Chatterji.** 1995. Evidence for a ppGpp-binding site on *Escherichia coli* RNA polymerase: proximity relationship with the rifampicin-binding domain. Mol Microbiol **15:**255-65.
101. **Roberts, D. M., R. P. Liao, G. Wisedchaisri, W. G. Hol, and D. R. Sherman.** 2004. Two sensor kinases contribute to the hypoxic response of *Mycobacterium tuberculosis*. J Biol Chem **279:**23082-7.
102. **Russell-Goldman, E., J. Xu, X. Wang, J. Chan, and J. M. Tufariello.** 2008. A double Rpf knockout *Mycobacterium tuberculosis* strain exhibits profound defects in reactivation from chronic tuberculosis and innate immunity phenotypes. Infect Immun **76:**4269-81.
103. **Rustad, T. R., M. I. Harrell, R. Liao, and D. R. Sherman.** 2008. The enduring hypoxic response of *Mycobacterium tuberculosis*. PLoS ONE **3:**e1502.
104. **Sassetti, C. M., D. H. Boyd, and E. J. Rubin.** 2003. Genes required for mycobacterial growth defined by high density mutagenesis. Mol Microbiol **48:**77-84.
105. **Scherr, N., and L. Nguyen.** 2009. *Mycobacterium* versus *Streptomyces*-we are different, we are the same. Curr Opin Microbiol.
106. **Schleifer, K. H., and O. Kandler.** 1967. *Micrococcus lysodeikticus*: a new type of cross-linkage of the murein. Biochem Biophys Res Commun **28:**965-72.
107. **Schleifer, K. H., and O. Kandler.** 1972. Peptidoglycan types of bacterial cell walls and their taxonomic implications. Bacteriol Rev **36:**407-77.
108. **Schmit, A. S., D. D. Pless, and W. J. Lennarz.** 1974. Some aspects of the chemistry and biochemistry of membranes of gram-positive bacteria. Ann N Y Acad Sci **235:**91-104.
109. **Schnappinger, D., S. Ehrt, M. I. Voskuil, Y. Liu, J. A. Mangan, I. M. Monahan, G. Dolganov, B. Efron, P. D. Butcher, C. Nathan, and G. K. Schoolnik.** 2003. Transcriptional adaptation of *Mycobacterium tuberculosis* within macrophages: insights into the phagosomal environment. J Exp Med **198:**693-704.

110. **Schneider, D., C. J. Bruton, and K. F. Chater.** 2000. Duplicated gene clusters suggest an interplay of glycogen and trehalose metabolism during sequential stages of aerial mycelium development in *Streptomyces coelicolor* A3(2). Mol Gen Genet **263:**543-53.
111. **Shah, I. M., M. H. Laaberki, D. L. Popham, and J. Dworkin.** 2008. A eukaryotic-like Ser/Thr kinase signals bacteria to exit dormancy in response to peptidoglycan fragments. Cell **135:**486-96.
112. **Sharma, V., S. Sharma, K. Hoener zu Bentrup, J. D. McKinney, D. G. Russell, W. R. Jacobs, Jr., and J. C. Sacchettini.** 2000. Structure of isocitrate lyase, a persistence factor of *Mycobacterium tuberculosis*. Nat Struct Biol **7:**663-8.
113. **Shleeva, M. O., K. Bagramyan, M. V. Telkov, G. V. Mukamolova, M. Young, D. B. Kell, and A. S. Kaprelyants.** 2002. Formation and resuscitation of "non-culturable" cells of *Rhodococcus rhodochrous* and *Mycobacterium tuberculosis* in prolonged stationary phase. Microbiology **148:**1581-1591.
114. **Siefert, J. L., and G. E. Fox.** 1998. Phylogenetic mapping of bacterial morphology. Microbiology **144:**2803-8.
115. **Siguier, P., J. Perochon, L. Lestrade, J. Mahillon, and M. Chandler.** 2006. ISfinder: the reference centre for bacterial insertion sequences. Nucleic Acids Res **34:**D32-6.
116. **Sims, G. K., L. E. Sommers, and A. Konopka.** 1986. Degradation of pyridine by *Micrococcus luteus* isolated from soil. Appl Environ Microbiol **51:**963-968.
117. **Skerman, V. B. D., V. McGowan, P. H. A. Sneath, and (editors).** 1980. Approved Lists of Bacterial Names. International Journal of Systematic Bacteriology **30:**225-420.
118. **Smith, M. C., and H. M. Thorpe.** 2002. Diversity in the serine recombinases. Mol Microbiol **44:**299-307.
119. **Soliveri, J. A., J. Gomez, W. R. Bishai, and K. F. Chater.** 2000. Multiple paralogous genes related to the *Streptomyces coelicolor* developmental regulatory gene *whiB* are present in *Streptomyces* and other actinomycetes. Microbiology **146:**333-43.
120. **Spratt, B. G.** 1977. Properties of the penicillin-binding proteins of *Escherichia coli* K12. Eur J Biochem **72:**341-52.
121. **Stackebrandt, E., C. Koch, O. Gvozdiak, and P. Schumann.** 1995. Taxonomic dissection of the genus *Micrococcus*: *Kocuria* gen. nov., *Nesterenkonia* gen. nov., *Kytococcus* gen. nov., *Dermacoccus* gen. nov., and *Micrococcus* Cohn 1872 gen. emend. Int J Syst Bacteriol **45:**682-92.
122. **Stark, N. J., G. N. Levy, T. E. Rohr, and J. S. Anderson.** 1977. Reactions of second stage of biosynthesis of teichuronic acid of *Micrococcus lysodeikticus* cell walls. J Biol Chem **252:**3466-72.
123. **Takarada, H., M. Sekine, H. Kosugi, Y. Matsuo, T. Fujisawa, S. Omata, E. Kishi, A. Shimizu, N. Tsukatani, S. Tanikawa, N. Fujita, and S. Harayama.** 2008. Complete genome sequence of the soil actinomycete *Kocuria rhizophila*. J Bacteriol **190:**4139-46.
124. **Tatituri, R. V., P. A. Illarionov, L. G. Dover, J. Nigou, M. Gilleron, P. Hitchen, K. Krumbach, H. R. Morris, N. Spencer, A. Dell, L. Eggeling, and G. S. Besra.** 2007. Inactivation of *Corynebacterium glutamicum* NCgl0452 and the role of MgtA in the biosynthesis of a novel mannosylated glycolipid involved in lipomannan biosynthesis. J Biol Chem **282:**4561-72.
125. **Telkov, M. V., G. R. Demina, S. A. Voloshin, E. G. Salina, T. V. Dudik, T. N. Stekhanova, G. V. Mukamolova, K. A. Kazaryan, A. V. Goncharenko, M. Young, and A. S. Kaprelyants.** 2006. Proteins of the Rpf (resuscitation promoting factor) family are peptidoglycan hydrolases. Biochemistry (Mosc) **71:**414-22.

126. **Thanky, N. R., D. B. Young, and B. D. Robertson.** 2007. Unusual features of the cell cycle in mycobacteria: polar-restricted growth and the snapping-model of cell division. Tuberculosis (Edinb) **87:**231-6.
127. **Tornabene, T. G., E. Gelpi, and J. Oro.** 1967. Identification of fatty acids and aliphatic hydrocarbons in *Sarcina lutea* by gas chromatography and combined gas chromatography-mass spectrometry. J Bacteriol **94:**333-43.
128. **Trutko, S. M., L. I. Evtushenko, L. V. Dorofeeva, M. G. Shliapnikov, E. Gavrish, N. E. Suzina, and V. K. Akimenko.** 2003. [Terminal oxidases in different genera of the family Microbacteriaceae]. Mikrobiologiia **72:**301-7.
129. **Tufariello, J. M., K. Mi, J. Xu, Y. C. Manabe, A. K. Kesavan, J. Drumm, K. Tanaka, W. R. Jacobs, Jr., and J. Chan.** 2006. Deletion of the *Mycobacterium tuberculosis* resuscitation-promoting factor Rv1009 gene results in delayed reactivation from chronic tuberculosis. Infect Immun **74:**2985-95.
130. **Ventura, M., C. Canchaya, A. Tauch, G. Chandra, G. F. Fitzgerald, K. F. Chater, and D. van Sinderen.** 2007. Genomics of Actinobacteria: tracing the evolutionary history of an ancient phylum. Microbiol Mol Biol Rev **71:**495-548.
131. **Voskuil, M. I., D. Schnappinger, K. C. Visconti, M. I. Harrell, G. M. Dolganov, D. R. Sherman, and G. K. Schoolnik.** 2003. Inhibition of respiration by nitric oxide induces a *Mycobacterium tuberculosis* dormancy program. J Exp Med **198:**705-13.
132. **Voskuil, M. I., K. C. Visconti, and G. K. Schoolnik.** 2004. *Mycobacterium tuberculosis* gene expression during adaptation to stationary phase and low-oxygen dormancy. Tuberculosis (Edinb) **84:**218-27.
133. **Votyakova, T. V., A. S. Kaprelyants, and D. B. Kell.** 1994. Influence of viable cells on the resuscitation of dormant cells in *Micrococcus luteus* cultures held in an extended stationary phase: the population effect. Appl Environ Microbiol **60:**3284-3291.
134. **Wayne, L. G.** 1994. Dormancy of *Mycobacterium tuberculosis* and latency of disease. Eur J Clin Microbiol Infect Dis **13:**908-914.
135. **Wayne, L. G., and C. D. Sohaskey.** 2001. Nonreplicating persistence of *Mycobacterium tuberculosis*. Annu Rev Microbiol **55:**139-63.
136. **Weiss, D. S., K. Pogliano, M. Carson, L. M. Guzman, C. Fraipont, M. Nguyen-Disteche, R. Losick, and J. Beckwith.** 1997. Localization of the *Escherichia coli* cell division protein Ftsl (PBP3) to the division site and cell pole. Mol Microbiol **25:**671-81.
137. **Wiens, G. D., D. D. Rockey, Z. Wu, J. Chang, R. Levy, S. Crane, D. S. Chen, G. R. Capri, J. R. Burnett, P. S. Sudheesh, M. J. Schipma, H. Burd, A. Bhattacharyya, L. D. Rhodes, R. Kaul, and M. S. Strom.** 2008. Genome sequence of the fish pathogen *Renibacterium salmoninarum* suggests reductive evolution away from an environmental *Arthrobacter* ancestor. J Bacteriol **190:**6970-82.
138. **Wieser, M., E. B. Denner, P. Kampfer, P. Schumann, B. Tindall, U. Steiner, D. Vybiral, W. Lubitz, A. M. Maszenan, B. K. Patel, R. J. Seviour, C. Radax, and H. J. Busse.** 2002. Emended descriptions of the genus *Micrococcus*, *Micrococcus luteus* (Cohn 1872) and *Micrococcus lylae* (Kloos et al. 1974). Int J Syst Evol Microbiol **52:**629-37.
139. **Wolf, A., R. Kramer, and S. Morbach.** 2003. Three pathways for trehalose metabolism in *Corynebacterium glutamicum* ATCC13032 and their significance in response to osmotic stress. Mol Microbiol **49:**1119-34.
140. **Wolin, H. L., and H. B. Naylor.** 1957. Basic nutritional requirements of *Micrococcus lysodeikticus*. J Bacteriol **74:**163-7.
141. **Yeats, C., R. D. Finn, and A. Bateman.** 2002. The PASTA domain: a beta-lactam-binding domain. Trends Biochem Sci **27:**438.

1

Table 1 Genome statistics for members of the *Micrococcaceae*.

	Micrococcus luteus	*Kocuria rhizophila*	*Arthrobacter aurescens* TC1	*Arthrobacter* sp. FB24	*Renibacterium salmoninarum*
Genome size (bp)	**2,501,097**	**2,697,540**	**5,226,648**	**5,070,478**	**3,155,210**
Coding region (bp)	**2,296,689 (91.8 %)**	**2,408,673 (89.29 %)**	**4,705,572 (90 %)**	**4,573,776 (90.2%)**	**2,863,187 (90.7%)**
G+C content	**73 %**	**71 %**	**62 %**	**65 %**	**56 %**
Plasmids	0*	0	2	3	0
Total genes	**2,348**	**2,413**	**4,793**	**4,622**	**3,558**
RNA genes	**60 (2.6 %)**	**56 (2.4 %)**	**94 (2.0%)**	**86 (1.9 %)**	**51 (1.4 %)**
Protein-coding genes	**2,288 (97.4 %)**	**2,357 (97.7 %)**	**4,699 (98.0 %)**	**4,536 (98.1 %)**	**3,507 (98.6 %)**
Genes with function	**1,742 (74.2 %)**	**1,478 (61.3 %)**	**3,419 (71.33 %)**	**3,279 (70.9 %)**	**2,679 (75.3 %)**
Genes in ortholog clusters			**4,316 (90.4%)**	**4,216 (91.5%)**	
Genes in paralog clusters	**217 (9.2 %)**	**967 (40.1 %)**	**2,749 (57.6%)**	**2,824 (61.3%)**	
Genes assigned to COGs	**1,717 (73.1 %)**	**1,799 (74.6 %)**	**3,307 (69.0 %)**	**3,361 (72.7 %)**	**2,389 (67.1 %)**
Genes assigned to Pfam	**1,731 (73.7 %)**	**1,822 (75.5 %)**	**3,525 (73.5 %)**	**3,426 (74.1 %)**	**2,478 (69.7 %)**
Genes with signal peptides	**487 (20.7 %)**	**653 (27.1 %)**	**1,442 (30.1 %)**	**1,454 (31.5 %)**	**1,030 (29.0 %)**
Genes with transmembrane helices	**543 (23.1 %)**	**569 (23.6 %)**	**1,187 (24.9 %)**	**1,168 (25.3 %)**	**836 (23.5 %)**
Fused genes	**195 (8.3 %)**	**135 (5.6 %)**	**292 (6.1 %)**	**320 (6.9 %)**	**125 (3.5 %)**

*Although a plasmid denoted pMLU1 has previously been reported from the NCTC 2665 strain of *M. luteus* (81), there was no evidence of it in the DNA provided for the genome sequencing project

Table 2 Genes in selected COG functional categories

Organism	COG Genes	Amino acid Metabolism	Carbohydrate Metabolism	Energy	Ion Transport	Lipid Metabolism	Replication Repair	Secondary Metabolism	Transcription	Translation
Micrococcus luteus (Fleming) NCTC 2665	1768	195 (11.0 %)	97 (5.5 %)	118 (6.7 %)	113 (6.4 %)	92 (5.2 %)	172 (9.7 %)	39 (2.2 %)	106 (6.0 %)	147 (8.3 %)
Kocuria rhizophila DC2201	1799	212 (11.8 %)	125 (7.0 %)	123 (6.8 %)	110 (6.1 %)	101 (5.6 %)	107 (6.0 %)	44 (2.5 %)	129 (7.2 %)	150 (8.3 %)
Arthrobacter aurescens TC1	3307	370 (11.2 %)	441 (13.3 %)	213 (6.4 %)	198 (6.0 %)	153 (4.6 %)	157 (4.8 %)	108 (3.3 %)	364 (11.0 %)	165 (5.0 %)
Arthrobacter sp. FB24	3361	364 (10.8 %)	436 (13.0 %)	239 (7.1 %)	208 (6.2 %)	157 (4.7 %)	164 (4.9 %)	112 (3.3 %)	363 (10.8 %)	162 (4.8 %)
Renibacterium salmoninarum ATCC 33209	2389	288 (12.1 %)	250 (10.5 %)	155 (6.5 %)	142 (5.9 %)	141 (5.9 %)	183 (7.7 %)	84 (3.5 %)	238 (10.0 %)	158 (6.6 %)

Table 3 Distribution of *M. luteus* IS elements among the different families

Family	Chemistry[a]	Distinct elements	Total no of copies [Partial]
IS*3*	DDE	**8**	**8 [7]**
IS*5*	DDE	**4**	**12 [0]**
IS*21*	DDE	**1**	**1 [0]**
IS*30*	DDE	**1**	**2 [0]**
IS*110*	DDE?	**1**	**2 [0]**
IS*256*	DDE	**5**	**24 [5]**
IS*481*	DDE	**6**	**7 [4]**
Total		**31**	**73 [19]**

The various IS families have been described and documented by Mahillon and Chandler (19,71). IS*1*, IS*4*, IS*6*, IS*66*, IS*91*, IS*200*/IS*605* IS*607* IS*630*, IS*701*, IS*982*, IS*1380*, IS*1634*, ISAs*1*, ISH*3*, ISL*3* and Tn*3* family elements were not found.

[a]Details of transposase chemistry are given on the ISfinder website (http://www-is.biotoul.fr/).

Table 4 *M. luteus* genes concerned with production of polyprenyl lipid-linked peptidoglycan monomer precursors

Gene	Mlut identifier	Product function
murA	Mlut_08760	UDP-GlcNAc carboxyvinyltransferase
murB	Mlut_17500	UDP-MurNAc dehydrogenase
murC	Mlut_13590	UDP-MurNAc-L-Ala ligase
murD	Mlut_13620	UDP-MurNAc-L-Ala-D-glutamate ligase
murE	Mlut_13650	UDP-MurNAc-L-Ala-D-glu-L-Lys ligase
alr	Mlut_13550	Alanine racemase
ddl	Mlut_08790	D-Ala-D-Ala ligase
murF	Mlut_13640	UDP-MurNAc-tripeptide D-Ala-D-Ala ligase
mraY	Mlut_13630	Phospho-MurNac-pentapeptide transferase
murG	Mlut_13600	Polyprenyl diphospho-MurNAc-pentapeptide GlcNAc transfera
murI	?	UDP-MurNAc penatapeptide(D-Glu) glycinyltransferase

Table 5 Genes involved in carotenoid production

Mlut	aa	Annotation	Comment	EC/COG
21270	143	Thioredoxin	Ubiquitous oxidoreductase	1.8.7.1.
21260	209	Isopentenyl diphosphate delta isomerase	3-isopentenyl pyrophosphate → dimethylallyl pyrophosphate	2.5.1.1 (?)
21250	182	hypothetical	No matches	
21240		Geranylgeranyl pyrophosphate synthase	trans, trans-farnesyl diphosphate + isopentenyl diphosphate → diphosphate + geranylgeranyl diphosphate	2.5.1.29
21230	298	squalene/phytoene synthetase	Probably phytoene synthetase 2 geranylgeranyl diphosphate → diphosphate + prephytoene diphosphate	2.5.1.32
21220	566	Phytoene desaturase	Carotene desaturation, a step in carotenoid biosynthesis	COG1233
21210	294	4-hydroxybenzoate polyprenyltransferase & related prenyltransferases	*crtEB** Lycopene elongation	2.5.1.39
21200	129	putative C50 carotenoid epsilon cyclase	*crtYe** Ring closure	
21190	117	hypothetical	Matches short segments of lycopene e-cyclase isoprenoid and putative C50 carotenoid epsilon cyclase (*crtYf**)	

* Notation for genes in *Corynebacterium glutamicum* (63)

Table 6 Distribution of genes involved in carotenoid synthesis

Gene number	Mlut_21190	Mlut_21200	Mlut_21210	Mlut_21220	Mlut_21230	Mlut_21240	Mlut_21250	Mlut_21260
Annotation	Partial carotenoid cyclase	C50 carotenoid epsiolon cyclase	Polyprenyl transferase	Phytoene desaturase	Phytoene synthetase	Geranylgeranyl synthase	Hypothetical	Isopentenyl isomerase
Actinobacteria	+	+	+	+	+	+		+
α proteobacteria			+	+				+
β proteobacteria						+		
γ proteobacteria				+				
δ proteobacteria	+		+	+		+		
Firmicutes			+	+				
Archaea	+		+	+	+	+		
Green non-sulfur				+	+			
Green sulfur			+		+			
Cyanobacteria			+		+			
Planctomycetes			+	+	+			
Basidiomycetes				+				
Ascomycetes				+				
Verrumicrobia			+					

Table 7 Comparison of the non-mevalonate pathway for isoprenoid biosynthesis in *M. luteus* and *My. tuberculosis*

Enzyme	***My. tuberculosis* gene**	***M. luteus* homolog**	**Percent identity**
1-deoxy-D-xylulose-5-phosphate synthase [EC:2.2.1.7]	Rv2682c *dxs1*	Mlut_13030	57.5
1-deoxy-D-xylulose 5-phosphate reductoisomerase [EC:1.1.1.267]	Rv2870c *dxr*	Mlut_06920	56.3
2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase [EC:2.7.7.60]	Rv3582c *ispD*	Mlut_03780*	42.6
4-diphosphocytidyl-2-C-methyl-D-erythritol kinase [EC:2.7.1.148]	Rv1011 *ispE*	Mlut_05400	44.7
2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase [EC:4.6.1.12]	Rv3581c *ispF*	M.lut_03780*	60.5
4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase [EC:1.17.4.3]	Rv2868c *ispG* (*gcpE*)	Mlut_06940	77.23
4-hydroxy-3-methylbut-2-enyl diphosphate reductase [EC:1.17.1.2]	Rv1110 *ispH*, *lytB*	Mlut_16300	62.7
isopentenyl-diphosphate delta-isomerase [EC:5.3.3.2]	Rv1745c *idi*	Mlut_21260	46.4

*Mlut_03780 encodes a bifunctional enzyme as also occurs in some other actinobacteria (see text)

Table 8. Number of genes up-regulated (more than 3-fold induction) in different models of *M. tuberculosis* dormancy and their *M. luteus* homologs[a]

	All genes	Seen in >2 models	Seen in all 5 models	Non-replicating Persistence (NRP1, 8 d) (132)	Non-replicating Persistence (NRP2, 20 d) (132)	Chemostat model (7)	Seen in both NRP1 and NRP2 (91)	Macrophages activated for 48 h (109)
Total	247	72	17	55	94	52	90	111
Homolog in *M. luteus*	137	36	9	24	53	30	44	61
% with *M. luteus* homolog	55.5	50.0	53	43.6	56.4	57.7	48.9	55.0

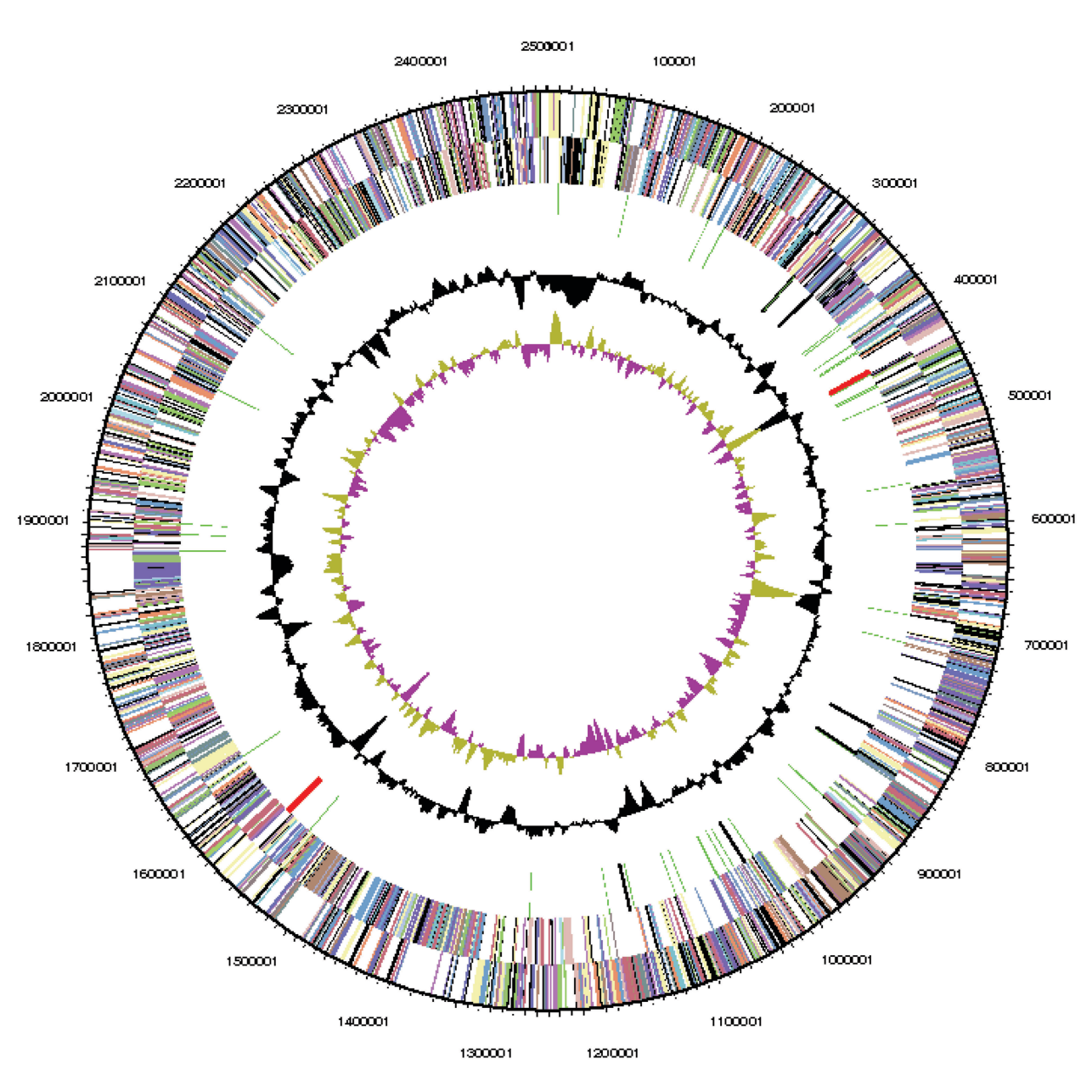
[a] Minimum 20% identity; maximum e value $1e^{-2}$. A low stringency was employed to identify as many genes as possible in *M. luteus* that might be homologs of the dormancy-related genes of *M. tuberculosis*. Even so, comparatively few candidates emerged from the analysis. Despite a significant reduction in the number of "dormancy-related" genes *M. luteus* can readily adopt a dormant state.

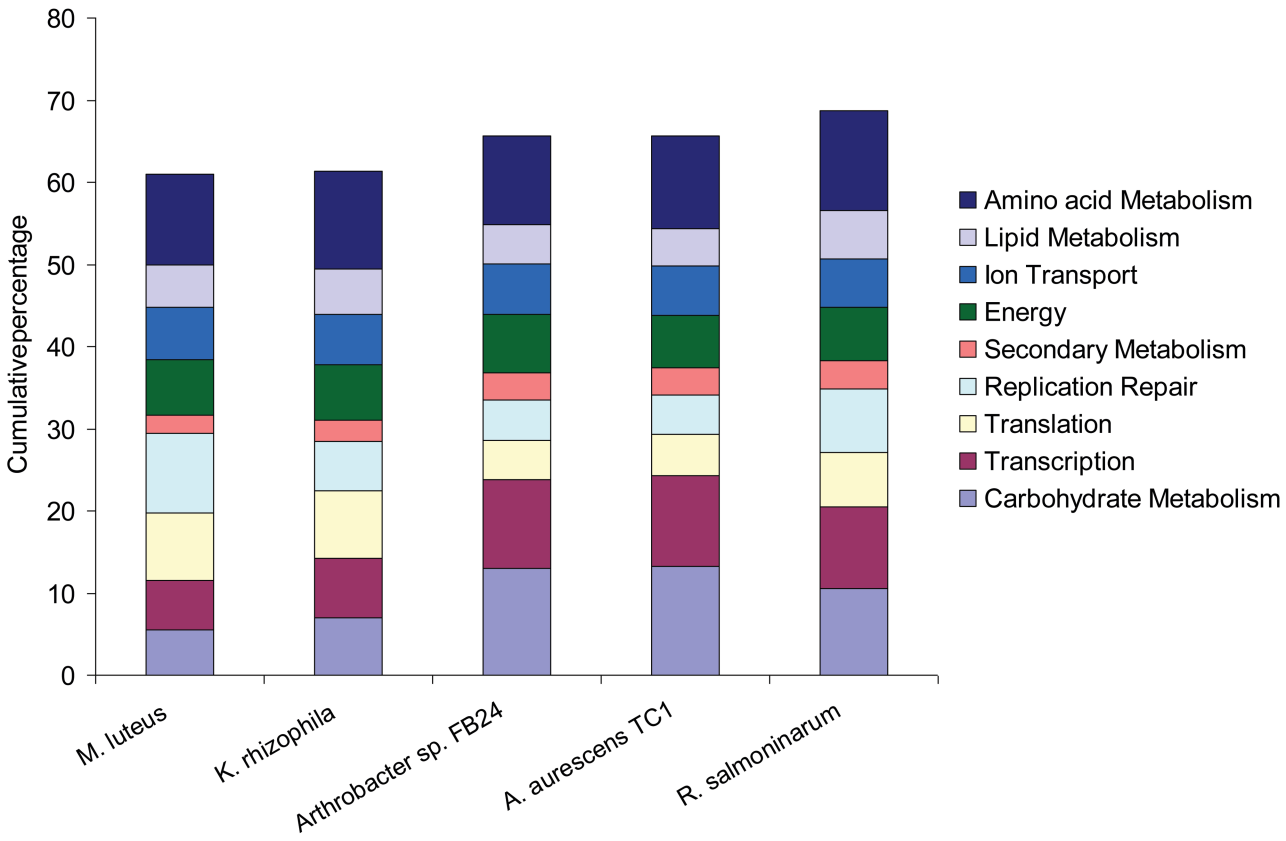
Table 9. Many of the genes up-regulated in all 5 *My. tuberculosis* dormancy models have *M. luteus* homologs

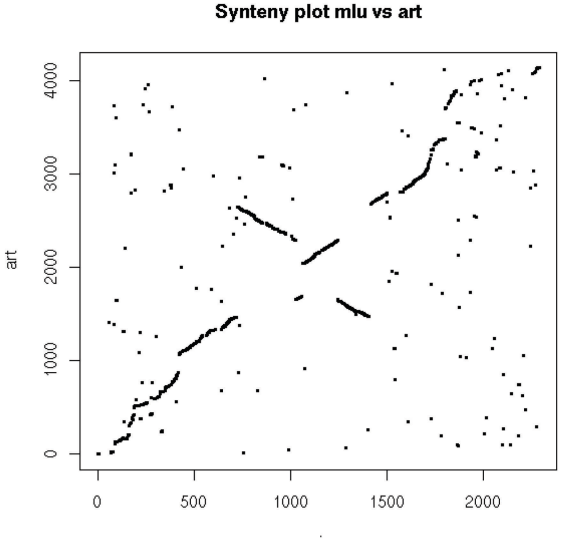
My. tuberculosis locus tag	Product name / assignment	*M. luteus* homolog	Product name / assignment	E-value	Percent identity / Bit Score
Rv0079	hypothetical protein				
Rv0080	conserved hypothetical protein	Mlut_04380	Predicted flavin-nucleotide-binding protein	3e-08	30.4 / 50
Rv1733c*	possible membrane protein				
Rv1738	conserved hypothetical protein				
Rv1996	COG0589 - Universal stress protein UspA and related nucleotide-binding proteins	Mlut_01830	Universal stress protein UspA and related nucleotide-binding proteins	9e-29	31.6 / 120
Rv2007c	ferredoxin	Mlut_15510	Ferredoxin	5e-36	58.0 / 142
Rv2030c	COG2312 - Erythromycin esterase homolog COG1926 – Predicted phosphoribosyltransferases	Mlut_18600[$]	amidophosphoribosyltransferase (EC 2.4.2.14)	1e-03	27.9 / 38
Rv2031c	14kD antigen, heat shock protein Hsp20 family	Mlut_16210	heat shock protein Hsp20 (IMGterm)	6e-09	36.9 / 52
Rv2032	Conserved hypothetical protein Acg				
Rv2623	COG0589 - Universal stress protein UspA and related nucleotide-binding proteins	Mlut_01830[$$]	Universal stress protein UspA and related nucleotide-binding proteins	4e-34	36.8 / 137
Rv2624c	COG0589 - Universal stress protein UspA and related nucleotide-binding proteins	Mlut_01830	Universal stress protein UspA and related nucleotide-binding proteins	7e-22	32.4 / 97
Rv2625c	COG1994 - Zn-dependent proteases (probable conserved transmembrane alanine and leucine rich protein)	Mlut_11840	Zn-dependent proteases	7e-31	30.1 / 127
Rv2626c	conserved hypothetical protein				
Rv2627c	conserved hypothetical protein				
Rv3127	conserved hypothetical protein				
Rv3130c	IGR02946 acyltransferase, WS/DGAT/MGAT				

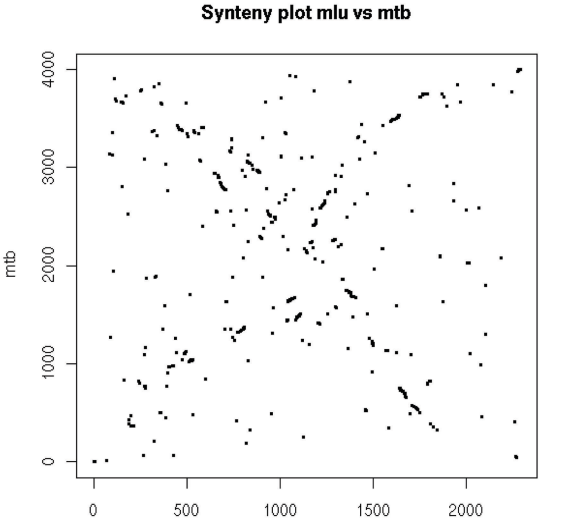
Rv3134c	COG0589 - Universal stress protein UspA and related nucleotide-binding proteins	Mlut_01830	Universal stress protein UspA and related nucleotide-binding proteins	2e-16	36.4 / 78

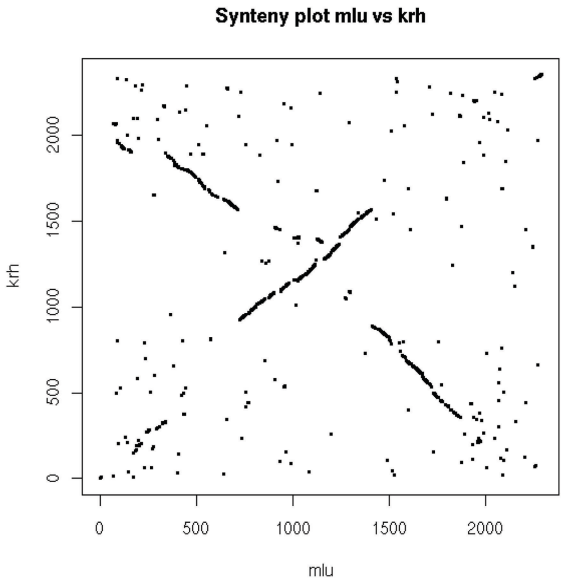
*All genes except Rv1733c belong to the DOS regulon
[$]3 more homologs
[$$]1 more homolog

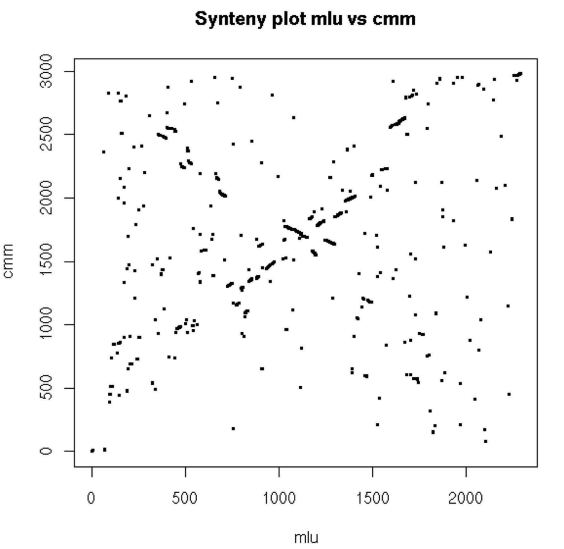

2500001
100001
200001
300001
400001
500001
600001
700001
800001
900001
1000001
1100001
1200001
1300001
1400001
1500001
1600001
1700001
1800001
1900001
2000001
2100001
2200001
2300001
2400001

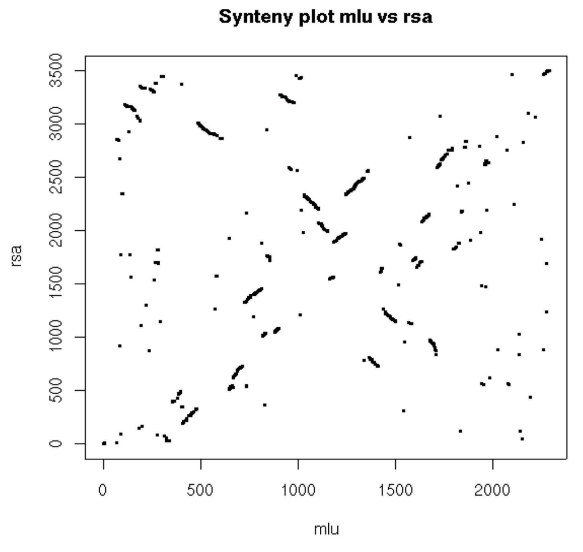

80
70
60
50
40
30
20
10
0
Cumulativepercentage
Amino acid Metabolism
Lipid Metabolism
Ion Transport
Energy
Secondary Metabolism
Replication Repair
Translation
Transcription
Carbohydrate Metabolism
M. luteus
K. rhizophila
Arthrobacter sp. FB24
A. aurescens TC1
R. salmoninarum

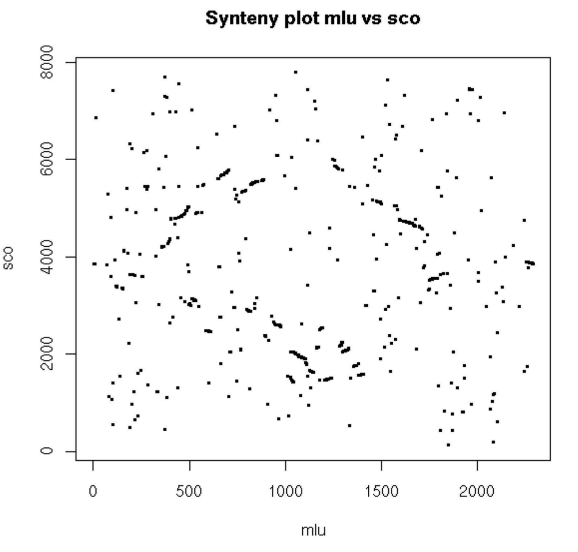

Synteny plot mlu vs art
art
0
1000
2000
3000
4000
0
500
1000
1500
2000

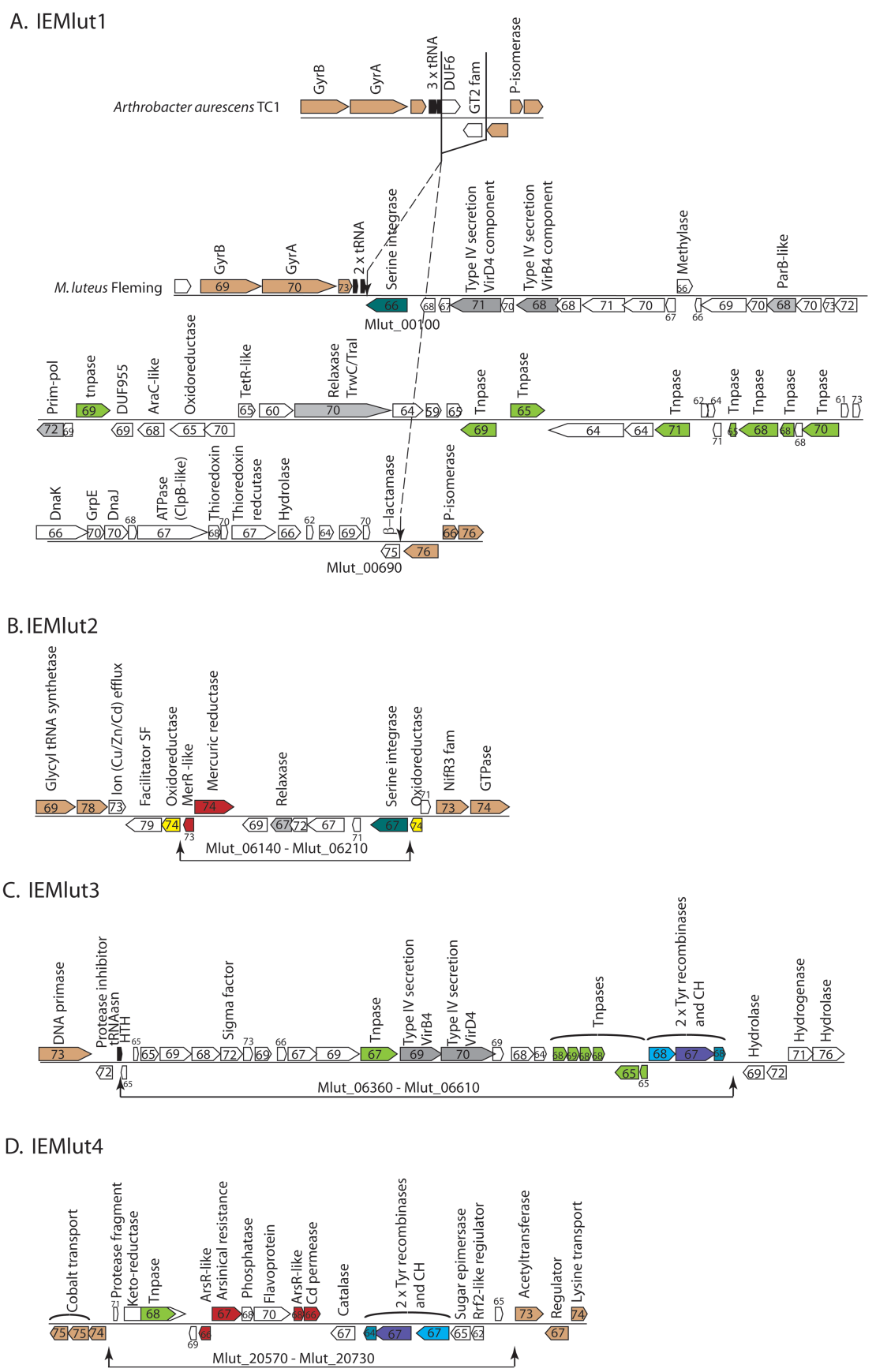

Synteny plot mlu vs mtb
mtb
0
1000
2000
3000
4000
0
500
1000
1500
2000

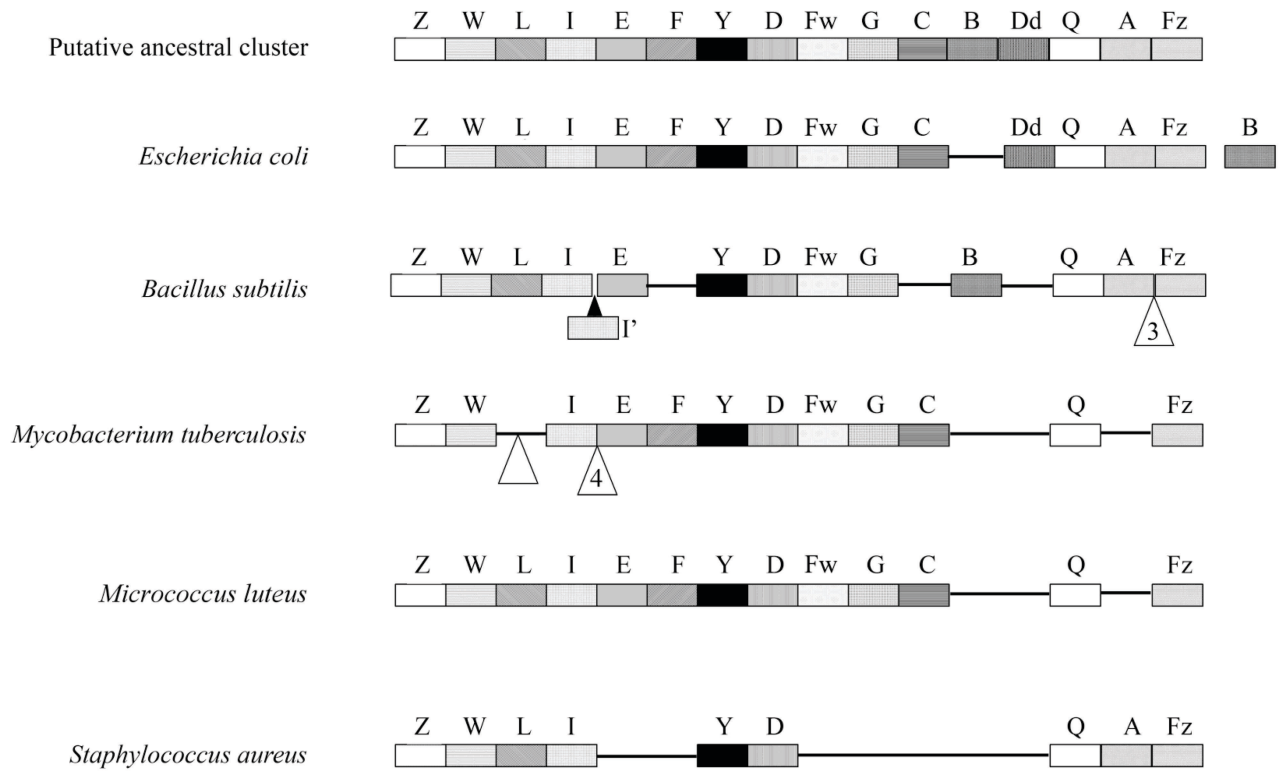

Synteny plot mlu vs krh
krh
0
500
1000
1500
2000
0
500
1000
1500
2000
mlu


Synteny plot mlu vs cmm
cmm
0
500
1000
1500
2000
2500
3000
0
500
1000
1500
2000
mlu


Synteny plot mlu vs rsa
rsa
0
500
1000
1500
2000
2500
3000
3500
0
500
1000
1500
2000
mlu


Synteny plot mlu vs sco
sco
0
2000
4000
6000
8000
0
500
1000
1500
2000
mlu


A. IEMlut1
Arthrobacter aurescens TC1
GyrB
GyrA
3 x tRNA
DUF6
GT2 fam
P-isomerase
M. luteus Fleming
2 x tRNA
Serine integrase
Type IV secretion VirD4 component
Type IV secretion VirB4 component
Methylase
ParB-like
Mlut_00100
Prim-pol
tnpase
DUF955
AraC-like
Oxidoreductase
TetR-like
Relaxase TrwC/TraI
Tnpase
DnaK
GrpE
DnaJ
ATPase (ClpB-like)
Thioredoxin
Thioredoxin reductase
Hydrolase
β-lactamase
P-isomerase
Mlut_00690
B. IEMlut2
Glycyl tRNA synthetase
Ion (Cu/Zn/Cd) efflux
Facilitator SF
Oxidoreductase MerR -like
Mercuric reductase
Relaxase
Serine integrase
Oxidoreductase
NifR3 fam
GTPase
Mlut_06140 - Mlut_06210
C. IEMlut3
DNA primase
Protease inhibitor tRNAasn HTH
Sigma factor
Tnpase
Type IV secretion VirB4
Type IV secretion VirD4
Tnpases
2 x Tyr recombinases and CH
Hydrolase
Hydrogenase
Hydrolase
Mlut_06360 - Mlut_06610
D. IEMlut4
Cobalt transport
Protease fragment
Keto-reductase
Tnpase
ArsR-like
Arsinical resistance
Phosphatase
Flavoprotein
ArsR-like
Cd permease
Catalase
2 x Tyr recombinases and CH
Sugar epimersase
Rrf2-like regiulator
Acetyltransferase
Regulator
Lysine transport
Mlut_20570 - Mlut_20730


Putative ancestral cluster
Z W L I E F Y D Fw G C B Dd Q A Fz
Escherichia coli
Z W L I E F Y D Fw G C Dd Q A Fz B
Bacillus subtilis
Z W L I E Y D Fw G B Q A Fz
I'
3
Mycobacterium tuberculosis
Z W I E F Y D Fw G C Q Fz
4
Micrococcus luteus
Z W L I E F Y D Fw G C Q Fz
Staphylococcus aureus
Z W L I Y D Q A Fz

Stenotrophomonas maltophilia R551-3

Arthrobacter sp. FB24

Arthrobacter aurescens TC1

Kocuria rhizophila

Micrococcus luteus