From Marinum to Ulcerans: a Mycobacterial Human Pathogen Emerges

Lateral gene transfers and genome reduction apparently remodeled a fish pathogen into the agent that causes Buruli ulcer in humans

Tim Stinear and Paul D. R. Johnson

Have you heard of Buruli ulcer? Probably not, unless you come from equatorial Africa, Papua New Guinea, the Daintree region in tropical Northern Australia, or seaside towns in temperate South Eastern Australia. Other than sporadic cases among visitors to these endemic zones, Buruli ulcer does not occur in North American or European countries. However, the surge and severity of this disease in Africa led the 57th General Assembly of the World Health Organization (WHO) in 2004 to recognize Buruli ulcer as a global threat to public health. WHO officials called for "research to develop tools to diagnose, treat and prevent the disease, as well as to integrate Buruli ulcer into the national disease surveillance system" (http://www.who.int/topics/mycobacterium_ulcerans/en/). Growing numbers of researchers and clinicians are committed to controlling the spread of Buruli ulcer disease. Under the leadership of Kingsley Asiedu at WHO in Geneva, the Global Buruli Ulcer Initiative continues to expand. WHO now hosts an annual meeting, inviting those involved in Buruli ulcer treatment, research, and disease control, and their combined effort is lifting the veil on this mysterious disease.

WHO officials identified six priorities for research, including mode of transmission; development of diagnostic tests; drug treatments and new treatment modalities; development of vaccines; social and economic studies; and studies to determine the incidence and prevailing mode of transmission. We also recommend investigating the environmental determinants that led to emergence of Buruli ulcer. Despite progress, there is still much to be done before those who live in endemic areas are freed from fear of this disease.

Summary

- Buruli ulcer, which is absent from North America and Europe, is found in Papua New Guinea and in tropical Northern as well as temperate South Eastern Australia, but is surging in equatorial Africa.
- Buruli ulcer arises through infections with *Mycobacterium ulcerans* from the environment that eventually produce deep skin ulcers.
- Comparing the genome sequences of *M. ulcerans* and *Mycobacterium marinum* suggests that the former emerged from the latter by acquiring genes for producing immunosuppressant mycolactones.
- *M. ulcerans* apparently is adapted to a dark and aerobic environment where its reduced antigenicity, slow growth, and production of mycolactones provide advantages for survival.

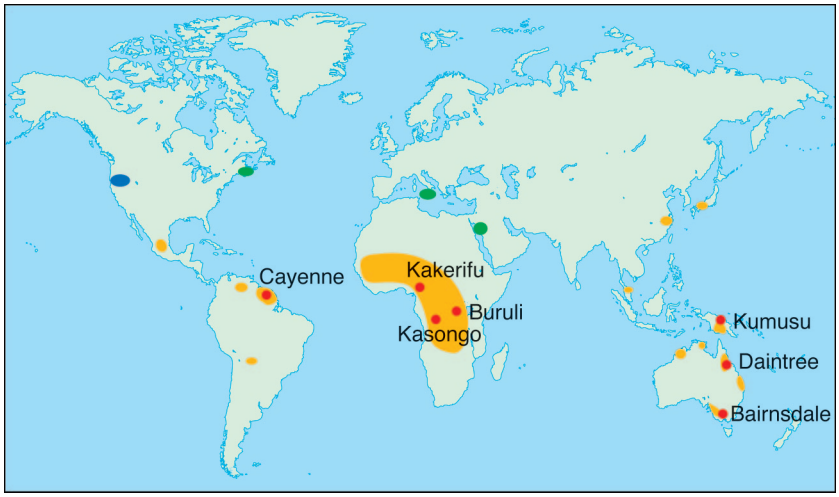
Buruli Ulcer in Humans

What is Buruli ulcer, and what do we know about the causative agent, *Mycobacterium ulcerans*? Since the late 1980s, cases of Buruli ulcer have steadily increased throughout West and Central Africa, including in Ivory Coast, Ghana, Benin, and Cameroon. Moreover, the incidence of Buruli ulcer now exceeds leprosy and, in some regions, tuberculosis, that are caused by two other infamous mycobacterial pathogens.

Like leprosy, Buruli ulcer is also a skin

Tim Stinear is a Research Fellow in the Department of Microbiology, Monash University, Clayton, Victoria, Australia, and Paul D. R. Johnson is the Deputy Director of the Department of Infectious Diseases, Austin Hospital, Heidelberg, Victoria, Australia

FIGURE 1



Regions of the world where *M. ulcerans* and other mycolactone-producing mycobacteria have been reported. Yellow shading indicates regions where human cases of *M. ulcerans* disease (Buruli ulcer) have been reported. Red dots indicate some of the regional names assigned to *M. ulcerans* disease. Green and blue shading indicates regions where *M. ulcerans*-like mycolactone-producing mycobacteria have been recovered from diseased fish and frogs respectively.

disease, but there the similarity ends. Buruli ulcer typically starts with a small mobile lump beneath the skin or a smaller raised lump attached to the skin. How *M. ulcerans* is introduced into the skin of humans remains unknown, but it is acquired directly or indirectly from the environment, not from contact with other patients. The disease progresses slowly over weeks to months, with little pain and none of the usual telltale signs of an infection such as fever or loss of weight and appetite.

Eventually, the affected area breaks down to produce a deep skin ulcer, usually on a limb, but any part of the body can be involved. If left untreated, these ulcers may become massive. Although host immunity can arrest the process, allowing scars to form, the aftermath can be devastating because scarring can cause major disabilities, including loss of limb or destruction of eyes or skin appendages, including the genitalia.

At first, treatment typically depended primarily on surgical excision and reconstructive plastic surgery. However, more recently, Buruli ulcer patients are being treated early with potent antimycobacterial drugs, such as rifampicin combined with streptomycin. In some situations, particularly when it is diagnosed early, some patients with Buruli ulcer may avoid surgery altogether.

The highly focal epidemiology and association with swamps and slow-flowing water are hallmarks of Buruli ulcer and have given rise to a plethora of regional names such as Kasongo ulcer, Kakerifu ulcer, Cayenne ulcer, Kumusu ulcer, Daintree ulcer, and Bairnsdale ulcer (Fig. 1). Bairnsdale is a small provincial town, 200 miles east of Melbourne in temperate South East Australia. An investigation of unusual skin ulcers among people from this district in the 1930s led Australian researchers to discover *M. ulcerans*. It was named Buruli ulcer later from an African epithet for a now-defunct county in Uganda (renamed the Nakasongola District) where many hundreds of cases were studied during the 1950s and 1960s.

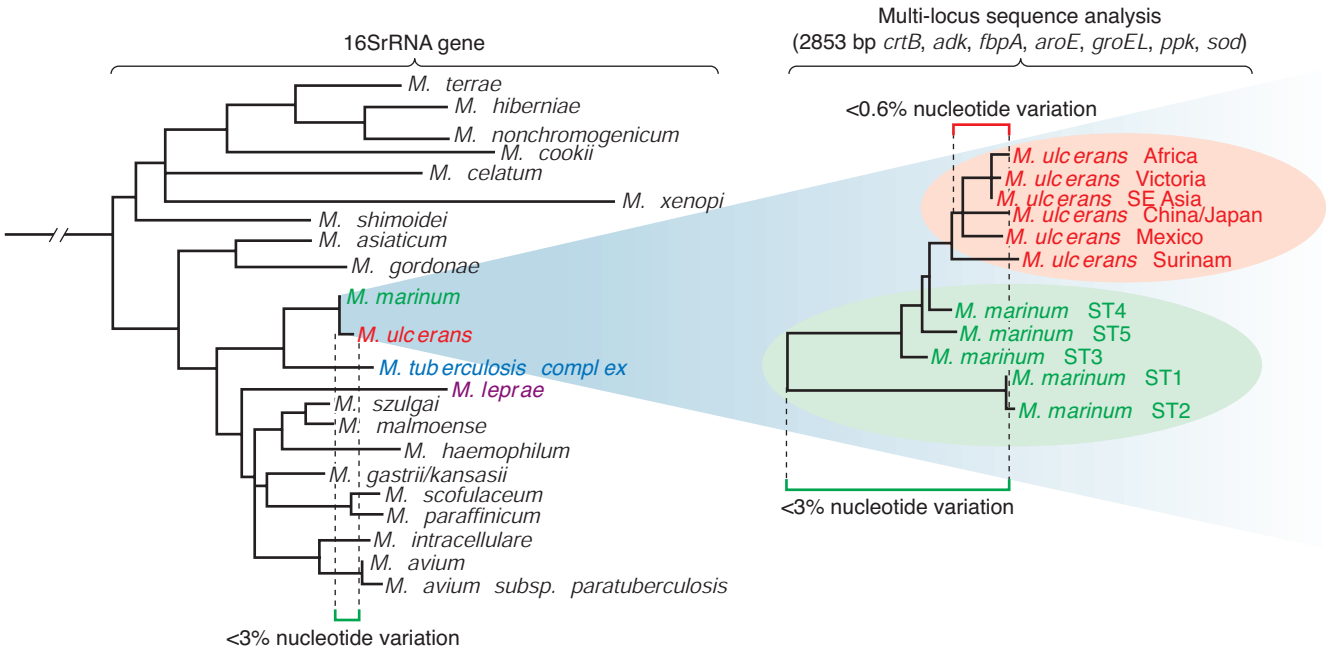
In July 1998 WHO officials, alarmed at the dramatic rise of Buruli ulcer across sub-Saharan Africa, launched the Global Buruli Ulcer Initiative (GBUI) and also made the name Buruli ulcer official for the disease caused by *M. ulcerans*. It remains a neglected disease of the rural poor of Africa, despite the efforts of dedicated researchers.

Key questions such as identifying the environmental reservoir of this bacterium, explaining focal outbreaks, and establishing how *M. ulcerans* is transmitted remain unanswered. However, the GBUI is helping to address these shortcomings. For instance, the genome sequence of *M. ulcerans* was recently completed, which is leading to detailed comparisons with the closely related *Mycobacterium marinum* and could help to explain how evolutionary forces shaped this enigmatic pathogen.

M. marinum and *M. ulcerans* Are Phenotypically Distinct

In the 1990s, when mycobacteriologists began compiling and comparing DNA sequences of the 16S rRNA gene from different members of the genus, they discovered that a few phenotypically distinct pairs of mycobacteria had identical se-

FIGURE 2



Phylogenetic relationship between: (left-side) some slow growing mycobacteria as revealed by analysis of the 16SrRNA gene and (right-side) *Mycobacterium ulcerans* and *Mycobacterium marinum* as revealed my multi-locus sequence analysis.

quences. One of these pairs was *M. marinum* and *M. ulcerans* (Fig. 2).

Their phenotypes are strikingly different. *M. marinum* grows relatively rapidly, doubling within 6–11 hours, produces photochromogenic pigments, and causes an intracellular, granulomatous disease in fish, frogs, and other ectotherms such as reptiles. It also occasionally causes self-limiting skin lesions in humans, usually on the cooler extremities of the body such as the hands. Meanwhile, *M. ulcerans* grows slowly, taking more than 48 hours to double, is nonpigmented, and causes a progressive, ulcerative disease in humans, in which extracellular bacteria collect in pools of necrotic subcutaneous fat. This pathology is largely attributed to mycolactone, a polyketide produced by *M. ulcerans* that has cytotoxic and immunosuppressive properties. *M. marinum* does not make mycolactone.

The relationship between contrasting phenotype and identical 16S rRNA gene sequences piqued the curiosity of Tone Tonjum from the Institute of Microbiology at the University of Oslo in Norway. By hybridizing DNA of different strains of *M. ulcerans* with those of *M. marinum*, she learned that the intraspecies relative binding ratio (RBR) is more than 85%, whereas the interspecies RBR is less than 40%. These findings suggest that, despite the apparent identity on the basis of 16SrRNA sequences, there are other major genetic differences between these two species that can help to explain their very different phenotypes.

Our scheme for multilocus sequence analysis (MLSA) provides another means to probe the genetic relationship between these species. Sequences for seven particular genes from 18 strains of *M. ulcerans* and 22 strains of *M. marinum* show more than 97% nucleotide identity, which is about the same as the 16SrRNA gene comparisons.

However, when we infer phylogeny by aligning these sequences, we see that all the *M. ulcerans* strains fall into a single cluster that could be further separated into unique genotypes based on one or two nucleotide polymorphisms that correlate with the geographic origin of a strain (Fig. 2). This geographic clonality is consistent with what Francoise Portaels of the Tropical

The Discovery of Mycolactone

Because bacterial toxins typically are proteins, it was big news in 1998 when Katie George, Pamela Small, and colleagues at the National Institutes of Health (NIH) Rocky Mountain Labs in Hamilton, Mont., discovered that the *M. ulcerans* toxin is a cytotoxic lipid of polyketide origin that they called mycolactone (Fig. 3). Polyketides are a class of pharmacologically important small molecules that include antibiotics such as erythromycin and immunosuppressants such as rapamycin and FK506. Mycolactone, too, has immunosuppressive properties and, in high concentrations, can provoke the appearance of ulcers. Mycolactone-deficient *M. ulcerans* mutants do not cause Buruli ulcer in animals that are susceptible to this disease. Strains of *M. ulcerans* produce variants of the molecule that are designated mycolactones A – F.

Medicine Institute in Antwerp, Belgium, and her collaborators reported in 1996. For example, *M. ulcerans* isolates from African countries share a particular genotype, whereas isolates from southeast Australia share another. Curiously, the *M. ulcerans* isolates from southeast Asia and Australia are more closely related to the African isolates than to isolates from elsewhere in the world. The lack of sequence diversity among the African isolates suggests that this *M. ulcerans* clone spread recently throughout that continent.

From these observations and discovery of two high-copy-number insertion sequences IS*2404* and IS*2606* specific to *M. ulcerans*, we propose that *M. ulcerans* strains diverged from a common *M. marinum* progenitor by acquiring DNA from the environment. In 2004 the unusual nature of that foreign DNA became apparent from the sequencing of the *M. ulcerans* genome.

Acquiring Genes To Produce Mycolactone Made *M. marinum* into *M. ulcerans*

One of the primary objectives of the *M. ulcerans* genome project was to uncover the genetic basis for the production of mycolactone, the primary determinant of virulence for this pathogen (see box, above). Structural analysis already revealed mycolactone is a polyketide that likely is the product of type I modular polyketide synthases (PKS). These are giant, multifunctional enzymes that produce carbon-chain compounds in an assembly-line manner, condensing organic acids such as acetate or propionate with enzymes that act much like those involved in fatty acid biosynthesis.

When Grant Jenkin, a doctoral student from our research group, conducted DNA subtractive hybridization experiments between *M. ulcerans* and *M. marinum*, he identified fragments of *M. ulcerans*-specific type I PKS genes. These PKS fragments were then used as probes to identify positive clones in a *M. ulcerans* whole-genome BAC library. On the basis of sequencing analysis, those clones span a large 174-kbp circular plasmid. This plasmid, pMUM001, contains a 100-kbp locus that encodes three large type I PKS genes named *mlsA1* (51 kb), *mlsA2* (7 kb), and *mlsB* (42 kb). Transposon mutagenesis confirms that the *mls* genes encode the synthases that are needed to produce mycolactone.

Meanwhile, phylogenetic trees produced by comparing plasmid gene sequences from many strains of *M. ulcerans* have the same branching topology as trees constructed by comparing chromosomal sequences of those same strains. These observations are consistent with plasmid acquisition being the key speciation event for *M. ulcerans*.

Mycolactone Biosynthesis Is Workable but Not Pretty

The unwieldy and counterintuitive trio of *mls* genes provide an example of how evolutionary forces can yield functions that are low in design aesthetics. What makes the mycolactone locus so unusual is its highly repetitive DNA sequence. The three *mls* genes are composed of 18 consecutive direct sequence repeats, where each repeat is approximately 6 kb and spans one PKS module.

Each PKS module has a discrete set of three to six enzymatic domains that together promote one round of polyketide chain extension by adding an acetate or propionate unit to a growing carbon chain. For example, every module has a ketosynthase (KS) domain that helps to form the critical C-C bond. Among type I PKS enzymes from different species, the domains of the same function commonly share up to 70% amino acid (aa) identity, as is the case for the 14 extension modules of the rapamycin PKS cluster.

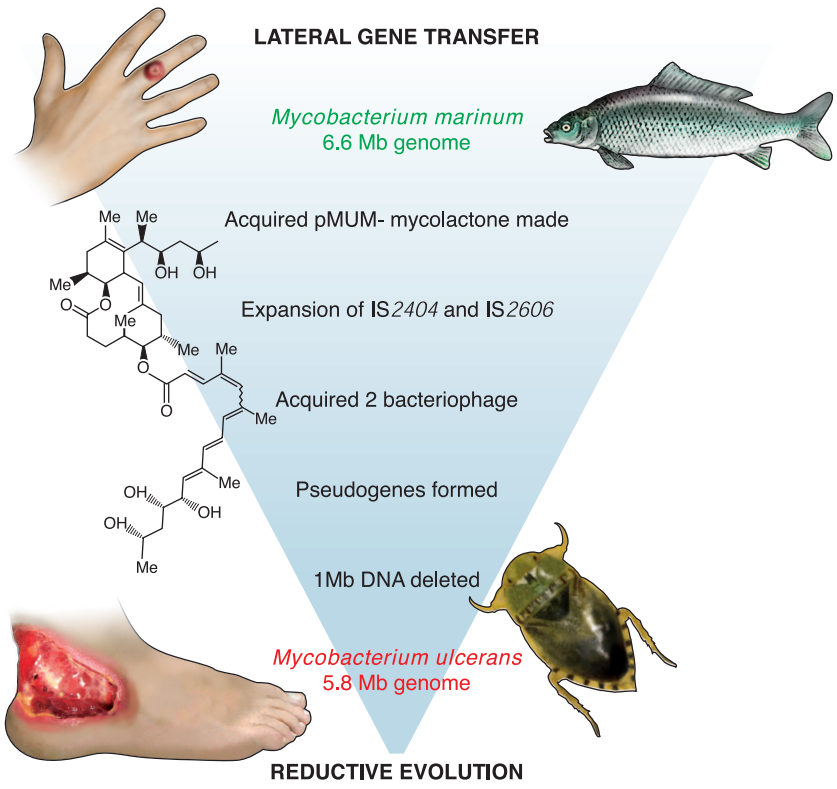
What makes the *mls* genes stand out is that their products have unusually high aa sequence

identity among such domains. For example, the 16 extension modules of the Mls locus have an intra-KS domain aa identity of 97%, and the scores are even higher among the other functional domains within the modules, ranging from 98.7–100%. At the nucleotide level, the entire 100-kb locus, encompassing *mlsA1, A2, and mlsB*, contains only 9.5 kb of unique DNA sequence.

These data suggest that this genetic locus arose recently and probably through a series of recombinations and duplications. The high level of sequence identity also suggests that the locus should be prone to losing its capacity to produce mycolactone, which occurs when clinical isolates of *M. ulcerans* are passaged in the lab. Nevertheless, strains isolated from African countries over a 40-year period or isolated separately from Africa and Malaysia all produce the same mycolactone A/B, suggesting that there is very strong selective force acting on *M. ulcerans* populations to preserve those bacteria that produce mycolactone. However, the vital role that mycolactone appears to play for *M. ulcerans* remains to be discovered.

Another consequence of the high sequence identity between modules occurs when they recombine. Thus, such modules may swap domains and begin producing different mycolactones. An example of such swapping occurred for a *M. ulcerans* strain from China; in this strain, a propionate-selecting domain has replaced an acetate-selecting domain in its acyltransferase, meaning it now produces mycolactones with an extra methyl group.

FIGURE 3



Summary of the evolutionary forces and events that have remodelled *Mycobacterium marinum*, a pathogen of fish and cause of a relatively minor granulomatous infection in humans into *Mycobacterium ulcerans,* the cause of severe, necrotic ulceration of subcutaneous tissue, as uncovered by comparative genomics. The molecule depicted is mycolactone A, the immunosuppressant and cytotoxin produced by *M. ulcerans*.

Snapshot of the *M. ulcerans* Genome

The *M. ulcerans* genome consists of two circular replicons, a 5,631,606-bp chromosome and a 174,155-bp plasmid designated pMUM001. The plasmid contains 81 coding domain sequences (CDS) with an average G+C content of 62.5%, and includes 4 copies of IS*2404* and 8 copies of IS*2606*. The chromosome has an average G+C content of 65.72% and contains 4,281 CDS and 727 pseudogenes. It harbors two prophages and 302 insertion sequence elements (ISE) that include 209 copies of IS*2404* and 83 copies of IS*2606*.

In contrast, *M. marinum* strain M has no plasmids, a 6,636,827-bp chromosome with 5,450 CDS, but few pseudogenes and ISE. Genome analysis confirmes earlier MLSA studies, showing that at the nucleotide level *M. ulcerans* is nearly indistinguishable from *M. marinum*, sharing 90% syntenic orthologs with an average nucleotide identity between each ortholog of 98.3%. However the comparisons with *M. marinum* also reveal substantial differences and show that *M. ulcerans* acquired a plasmid, expanded IS*2404* and IS*2606*, accumulated pseudogenes, and experienced extensive genome rearrangements and deletions (Fig. 3).

M. ulcerans also acquired 11 chromosomal CDS not present in *M. marinum*, some present on the two bacteriophage and some as small

insertions of less than 5 kbp. These CDS are detected among a wide range of *M. ulcerans* strains, making them appear to be "*M. ulcerans* specific." They might, in conjunction with mycolactone, contribute to the pathology associated with Buruli ulcer. They also might be used for developing diagnostic tests.

Rapid and Dramatic Genome Reduction

The high DNA identity between *M. ulcerans* and *M. marinum* suggests that they diverged recently. However, the *M. ulcerans* genome was drastically remodeled, losing more than 1,100 kbp compared with *M. marinum* through more than 150 deletions. The multiple copies of IS*2404* and IS*2606* facilitated many of these deletions and rearrangements. Several classes of genes are overrepresented among those lost from *M. ulcerans*, including the enigmatic PE/PPE genes and also CDS involved in secondary metabolism, such as those required for producing polyketides and other small molecules. *M. ulcerans* resembles *Mycobacterium leprae*, whose genome shows evidence of significant loss of genetic redundancy and metabolic streamlining.

The ESX loci is another deletion that may affect pathogenesis. Although the *M. marinum* genome contains five complete ESX systems, only three remain in *M. ulcerans*. In *M. tuberculosis* and *M. marinum*, the ESX-1 locus encodes a protein secretion apparatus that contributes to virulence, intercellular spread, and immunogenicity. Two deletions disrupt ESX-1 in *M. ulcerans*. Because these systems promote granuloma formation, their absence may in part explain why *M. ulcerans* remains extracellular in tissues that it infects.

M. ulcerans Has the Genomic Signature of Being Niche-Adapted

When bacteria encounter a new niche environment and one develops a mutation that confers a survival advantage, that cell will outcompete the original population by passing through an "evolutionary bottleneck." Its progeny may then expand clonally and its genome may remodel in response to the new environment. In this way a "generalist" bacterial species may become a "specialist."

Thus, the characteristics of recently emerged bacterial specialists include a high degree of clonality, foreign genes (often via plasmids or bacteriophage), ISE, accumulated pseudogenes, chromosomal rearrangements, and evidence of genome downsizing. Genomics analyses are uncovering examples of this phenomenon. For instance, *Yersinia pestis* apparently diverged recently from *Yersinia pseudotuberculosis* to occupy the midgut of the flea, while *Bordetella pertussis* derived from *Bordetella bronchiseptica* to become an obligate human pathogen.

Similarly, *M. ulcerans* incorporated many of the genomic features of a specialist bacterium (Fig. 3), having derived from the generalist *M. marinum*. However, a key piece of this puzzle is missing. What is the environmental niche of *M. ulcerans*? The types of genes that it lost give some clues as to its habitat. For example, phytoene dehydrogenase, encoded by *crtI*, is an essential enzyme for *M. marinum* to make light-inducible carotenoid pigments. These pigments protect the bacteria from UV-induced damage following exposure to sunlight. Although *M. ulcerans* has a pigment locus identical to that of *M. marinum*, the *crtI* gene in *M. ulcerans* is disrupted by a single nucleotide substitution. This change inserts a premature stop codon, indicating that the pigments are not required for the survival of *M. ulcerans*, presumably because it is not exposed to sunlight.

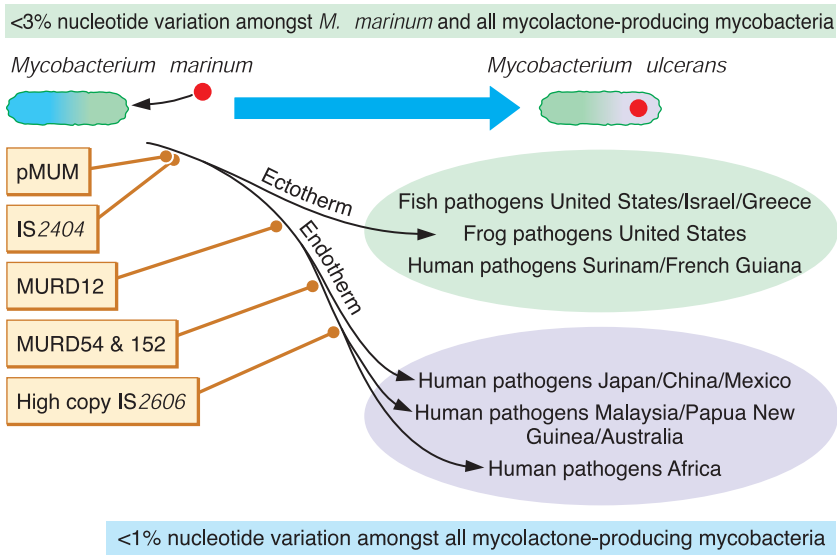
The *M. ulcerans* genome is missing genes for anaerobic respiration, indicating that the bacterium likely occupies an aerobic environment. The many DNA rearrangements, the significant depletion of paralogous gene families, and the metabolic streamlining all likely contribute to the slow growth rate of *M. ulcerans*. Put another way, this bacterium appears to have adapted to a dark and aerobic environment where its reduced antigenicity, slow growth, and production of immunosuppressant mycolactones provide an advantage for survival.

The hunt for an environmental reservoir for *M. ulcerans* has proved frustrating. It is severely hampered by our inability to culture the organism from environmental samples, even though it is relatively straightforward to do so from human ulcers. Its slow growth, which results in rapid overgrowth by other species in samples from swamps and rivers, further complicates such efforts.

The discovery of IS*2404* and PCR screening

of environmental samples led to evidence for *M. ulcerans* in predicted environmental locations. Since then, we learned that *M. ulcerans* forms biofilms on aquatic plants. It also is found in the salivary glands of carnivorous aquatic bugs (*Naucoridae*) in endemic areas in Africa. Mycolactones apparently help and may be required for *M. ulcerans* to colonize such bugs, perhaps explaining how *M. ulcerans* uses mycolactone to exploit this microenvironment. The role of naucorids in transmitting *M. ulcerans* to humans is being investigated. One possibility is that children become infected from bites when they swim in contaminated endemic areas. However, these insects do not seek humans as prey, and few Buruli ulcer patients in Australia report being bitten by them.

FIGURE 4



A proposed scenario for the evolution of mycolactone-producing mycobacteria from a common *Mycobacterium marinum* progenitor. MURD refers to *M. ulcerans* Region of Difference. These are DNA fragments > 1 kb deleted from *M. ulcerans* and given ID numbers from MURD1 – MURD157.

Fish and Frogs May Have Buruli Ulcers from Strains Derived from a Common Progenitor

These environmental questions took an interesting twist in 2006 with discovery of related mycobacteria that contain the pMUM001 plasmid, produce variant mycolactones, and cause disease in fish and frogs. Detecting these mycobacteria, variously called *M. marinum, Mycobacterium pseudoshottsii,* or *Mycobacterium liflandii,* raises the possibility that pMUM transfers among diverse mycobacteria.

Although mycolactone-producing mycobacteria (MPM) are spread globally, only some clones cause disease in humans. We recently showed that known MPM apparently derive from a common *M. marinum* progenitor to form a genetically cohesive group among a more diverse assemblage of *M. marinum* strains (Fig. 4).

Like *M. ulcerans*, the fish and frog MPM have multiple copies of the insertion sequence IS*2404*. However, patterns of DNA deletion and pseudogene accumulation among MPM are distinct, with only those strains that cause Buruli ulcer exhibiting significant evidence of genome downsizing. Further, by comparing pMUM and chromosomal gene sequences, we conclude that pMUM plasmid acquisition and mycolactone-producing capacity were probably key drivers of speciation. As MPM radiated globally, ongoing evolution produced at least two genetically distinct ecotypes that can be broadly divided into those typically causing disease in ectotherms such as fish and frogs and those causing disease in endotherms such as humans (Fig. 4).

SUGGESTED READING

Coutanceau, E., L. Marsollier, R. Brosch, E. Perret, P. Goossens, M. Tanguy, S. T. Cole, P. L. Small, and C. Demangel. 2005. Modulation of the host immune response by a transient intracellular stage of *Mycobacterium ulcerans*: the contribution of endogenous mycolactone toxin. Cell Microbiol. **7:**1187–1196.

George, K. M., D. Chatterjee, G. Gunawardana, D. Welty, J. Hayman, R. Lee, and P. L. Small. 1999. Mycolactone: a polyketide toxin from *Mycobacterium ulcerans* required for virulence. Science **283:**854–857.

Johnson, P. D., T. Stinear, P. L. Small, G. Pluschke, R. W. Merritt, F. Portaels, K. Huygen, J. A. Hayman, and K. Asiedu. 2005. Buruli ulcer (*M. ulcerans* infection): new insights, new hope for disease control. PLoS Med **2:**e108.

MacCallum, P., J. Tolhurst, G. Buckle, and S. HA. 1948. A new mycobacterial infection in man. J. Pathol. Bacteriol. **60:**93–122.

Marsollier, L., R. Robert, J. Aubry, J. P. Saint Andre, H. Kouakou, P. Legras, A. L. Manceau, C. Mahaza, and B. Carbonnelle. 2002. Aquatic insects as a vector for *Mycobacterium ulcerans*. Appl. Environ. Microbiol. **68:**4623–4628.

Portaels, F., P. Elsen, A. Guimares-Peres, P. A. Fonteyne, and W. M. Meyers. 1999. Insects in the transmission of *Mycobacterium ulcerans* infection [letter]. Lancet **353:**986.

Ross, B. C., P. D. Johnson, F. Oppedisano, L. Marino, A. Sievers, T. Stinear, J. A. Hayman, M. G. Veitch, and R. M. Robins-Browne. 1997. Detection of *Mycobacterium ulcerans* in environmental samples during an outbreak of ulcerative disease. Appl. Environ. Microbiol. **63:**4135–4138.

Stinear, T. P., A. Mve-Obiang, P. L. Small, W. Frigui, M. J. Pryor, R. Brosch, G. A. Jenkin, P. D. Johnson, J. K. Davies, R. E. Lee, S. Adusumilli, T. Garnier, S. F. Haydock, P. F. Leadlay, and S. T. Cole. 2004. Giant plasmid-encoded polyketide synthases produce the macrolide toxin of *Mycobacterium ulcerans*. Proc. Natl. Acad. Sci. USA **101:**1345–1349.

Stinear, T. P., T. Seemann, S. Pidot, W. Frigui, G. Reysset, T. Garnier, G. Meurice, D. Simon, C. Bouchier, L. Ma, M. Tichit, J. L. Porter, J. Ryan, P. D. R. Johnson, J. K. Davies, G. A. Jenkin, P. L. C. Small, L. M. Jones, T. F., F. Laval, M. Daffé, J. Parkhill, and S. T. Cole. 2006. Reductive evolution and niche adaptation inferred from the genome of *Mycobacterium ulcerans*, the causative agent of Buruli ulcer. Genome Res. Doi: 10.1101/gr.5942807.

Yip, M. J., J. L. Porter, J. A. M. Fyfe, C. J. Lavender, F. Portaels, M. Rhodes, H. Kator, A. Colorni, G. A. Jenkin, and T. P. Stinear. 2006. Evolution of *Mycobacterium ulcerans* and other mycolactone-producing mycobacteria from a common *Mycobacterium marinum* progenitor. J. Bacteriol. doi:10.1128.