Antonie van Leeuwenhoek **83:** 265–274, 2003.
© 2003 *Kluwer Academic Publishers. Printed in the Netherlands.*

Evidence for phylogenetic inheritance in pathogenicity of *Mycobacterium*

M.C. Leclerc*, F. Thomas and J.F. Guégan
*Centre d'Etudes sur le Polymorphisme des Micro-Organismes CEPM/UMR CNRS-IRD 9926, IRD, 911 Avenue Agropolis, B.P. 5045, 34032 Montpellier Cedex 1, France; *Author for correspondence (phone: (33) 04 67 41 62 32; fax: (33) 04 67 41 62 99; secrétariat: (33) 04 67 41 61 97)*

Received 5 September 2001; accepted in revised form 26 July 2002

Key words: Comparative analysis, *Mycobacterium*, Pathogenicity, Phylogenetic tree

Abstract

In this study, we attempt to highlight part of the adaptive and phylogenetic constraints in mycobacterial pathogenicity. For this purpose, we first provide a phylogeny of *Mycobacteria* based on cladistic analyses of 64 different taxa. We then performed a comparative analysis, taking into account both ecological factors and phylogenetic relationships. The GLIM modelling analysis showed that different ecological and phylogenetic factors might be invoked to explain the variation in pathogenicity levels. Interestingly, the most harmful species were shown to be connected with the most diversified habitats. However, the independent contrast analysis revealed that once phylogeny was taken into account, none of the relationships between ecological factors and pathogenicity remained significant, and the pathogenicity appeared to be phylogenetically inherited among mycobacteria. The most pathogen were found in the slow-growing/long helix 18 group, and within this group in the most derived taxa.

Introduction

Increasingly, research into infectious diseases acknowledges the need to consider evolutionary processes to understand virulence polymorphisms among and within groups of pathogens (Bowman and Taylor 1993; Bowman et al. 1996; Bidochka et al. 1999; Spratt and Maiden 1999; Picard et al. 1999; Stearns 1999; Poulin and Combes 1999; Maynard Smith et al. 2000). Like any other quantitative trait, virulence can be adjusted by natural selection in a way that maximises the pathogen's fitness. Virulence may thus theoretically arise independently in several different lineages, depending on the selective context experienced by the pathogens (i.e., ecological factors) (Ewald 1983, 1994; Agnew and Koella 1997; Lipsitch and Moxon 1997). The occurrence of virulence within taxa needs, however, to be examined properly within a phylogenetic context, since it can also be inherited from a common ancestor (Poulin 1998). Molecular phylogenies have in this respect many implications for understanding the evolution of pathogen variation since they, for instance, allow one to determine whether virulence in a given taxon is adaptive (i.e., independent event) and (or) an ancestral legacy, i.e., genetically based.

The genus *Mycobacterium* presently includes 94 species. Despite this, *M. tuberculosis* is responsible for some 2 million deaths annually (WHO reference), whereas most other species have not been isolated from clinical material and are regarded as non-pathogens, or at least as pathogens only under some extremely unusual conditions (Goodfellow and Magee 1998). The problem of pathogenicity among mycobacteria remains, however, complex to assess and currently the differentiation of opportunistic pathogens from non-pathogens is not always clear-cut (Tsukamura 1984). If it is likely that most infections do not lead to disease, it is even possible that non-pathogenic species *sensu stricto* do not exist (Tsukamura 1984). For instance, following environmental disturbances, taxa rarely described as pathogens can emerge as aggressive pathogens (Portaels 1987; Johnson et al. 1999). In addition, the fact that one species

is considered to be an opportunistic pathogen compared to another one may be biased, if the first species occurs more frequently in the human environment than the second one (i.e., differential ecological niche overlap).

Despite these difficulties, evidence suggests that some correlation could exist between the ecology, the phylogenetic relationships and the pathogenicity of mycobacteria. Since the first classification of Runyon (1959), which put mycobacteria into four groups according to their growth rate and their pigmentation state, the sequencing of SSU rDNA (small subunit ribosomal DNA) allowed the separation of the genus into three groups according to the growth rate, slow or fast, were associated with nucleotide structure helix 18, which can be short or long (Stahl and Urbance 1990; Rogall et al. 1990). The first group of fast-growing mycobacteria with a short helix 18 (group IV of Runyon) includes a great number of soil-isolated species, with very few pathogenic species. The second group is slow growing, with a long helix 18 that includes many pathogenic and opportunistic species to animal and human and also non-pathogenic species. The third group is slow growing, with a short helix 18 and is a comparatively small group of relatively recent origin and with opportunistic pathogen behaviour.

The aim of this study was to improve our knowledge concerning the possible relationships between the ecology, the phylogenetic relationships, and the pathogenicity of *Mycobacteria*. We first provided cladistic trees including 64 *Mycobacterium* taxa. Based on this information and on ecological data, we then performed comparative analyses so as to minimise phylogenetic effects, because species values in cross-species comparisons are often not independent (Felsenstein 1985; Harvey and Pagel 1991). In doing so, we attempted to determine whether pathogenicity among *Mycobacteria* was under adaptive, i.e., due to environmental pressures and (or) phylogenetic constraints, i.e., due to inheritance. We show how such comparative approaches can aid studies to tease apart the biological or ecological factors that determine pathogen virulence across different microbial taxa.

Materials and methods

Phylogenetic reconstruction

We constructed a phylogeny for 64 *Mycobacterium* species based on information derived from extensive studies of SSU rDNA sequences from GeneBank. Sequence editing and management were carried out using the MUST package (Philippe 1993), and sequence alignment was manually improved according to criteria proposed by Barriel (1994): minimise the mutation number, preferentially choose a nucleotide substitution by transition rather than transversion and reduce the phylogenetic effects. The most parsimonious trees (cladistic analysis) were calculated using the heuristic search of PAUP (Swofford 1993), with simple stepwise addition and the following options: TBR (tree bisection-reconnection) branch swapping, mulpars, branches having maximum length zero collapsed to yield polytomies**.** *Nocardia asteroides* and *Rhodococcus sp* (family *Nocardiaceae*) *Tsukamurella sp*, (family *Tsukamurellaceae*) and *Gordonia* sp. (family Gordionaceae) were chosen as outgroups. Their GeneBank accession numbers are respectively X84850, AJ007003, AF150494, AF150493 and X82242.

Variables

We included all variables known to us that could possibly cause interspecific differences in pathogenicity levels of *Mycobacterium* species (Table 1). The variables used were: (*i*) pathogenicity, we first scored in four levels, 1) non-pathogen, 2) rare pathogen, 3) potential pathogen, and 4) pathogen, and then we grouped non-pathogen and rare pathogen cases into a single level, since there may be some difficulties in distinguishing between these two categories; (*ii*) the date of species description (year of the description), since it can theoretically influence the extent of the knowledge for a given species (e.g., number of habitats, see below) and this may represent a sampling artefact in analysis; (*iii*) growth rate, slow or fast, associated with a SSU rDNA nucleotide structure, helix 18, which can be short or long; three groups were considered, 1) a slow-growing group with a short helix 18, 2) a slow-growing group with a long helix 18, 3) a fast-growing group with a short helix 18; (*iv*) habitat diversity, which corresponds to the number of living and non-living reservoirs where a given species has been isolated, e.g., soils, vegetation, fresh water, salt water, bovine host, caprine host, human host; (*v*) pigmentation, which was coded into four categories: 1) non pigmented, 2) photochromogen (pigmentation reaction to luminosity), 3) scotochromogen (pigmentation reaction to darkness), and 4) unknown. We collected data for a total subset of 64 *Mycobacterium* species for which information

was available. *M. avium* subsp *paratuberculosis*, *M. farcinogenes* and *M. senegalanse* were removed from the statistical analysis since they are known to be animal pathogen species but their human pathogenicity has not been proven.

Statistical analysis

The statistical methods used in the present study were both GLIM modelling, which provides a major improvement upon conventional linear regression analyses (Wilson and Grenfell 1997), and standardised linear phylogenetic contrasts (Felsenstein 1985). To overcome the problems caused by resemblance among *Mycobacterium* species due to shared common ancestry (Felsenstein 1985; Harvey and Pagel 1991), comparative analysis was carried out in two ways. In order to analyse the evolution of pathogenicity across different species of *Mycobacterium*, we first used GLIM modelling with the quasi-likelihood method, to simultaneously assess which explanatory variables and/or interaction terms better explained the interspecific differences in pathogenicity levels across *Mycobacterium* species. Quasi-likelihood methods allow one to estimate regression relationships without fully knowing the error distribution of the response variable (McCullagh and Nelder 1989). For this purpose, we used a GLIM model with a normal link function, which represents the most appropriate statistical tool for analysing our pathogenicity data.

We also compared the results obtained with the GLIM approach to more traditional methods. Here, we used the S-plus statistical package (Venables and Ripley 1994), which permits one to estimate the parameter values associated with the terms in the model. The scaled deviances for terms in the model were compared using F-tests. When data suggested no linear trends, explanatory variables were transformed and fitted again, trying to improve their contribution to the models. The variable year of description was used to control for possible time biases in *Mycobacterium* pathogenicity estimates. Second, to control for similarity in variables of interest due to common ancestry, we used statistically independent standardised linear contrasts (Felsenstein 1985; Harvey and Pagel 1991; Pagel 1992). Generally, statistical methods treat species (or any kind of taxa) values as statistically independent points, but this is simply not valid, since closely related taxa tend to share many characters through common descent rather than through independent evolution. The independent contrast method avoids spurious correlations due to non-independence of taxa by considering independent evolutionary events. The phylogenetically independent contrasts method has been the most popular approach over the past ten years (see Harvey and Pagel (1991), Martins (1996)). Readers are invited to refer to the original references for further details on the method (Garland et al. 1992; Purvis and Rambaut 1995). Here, we used two different techniques of comparative tests, i.e., GLIM and independent contrasts, to circumvent the problem of non-independence between *Mycobacterium* species. This was entirely effective since the independent contrast method considers *a priori* the importance of history on trait variation in extracting the component due to common ancestry, whereas GLIM modelling does not (Ricklefs 1996; Martins 1996). Many authors have argued for the use of different comparative tests, since in this way it is less likely to have to seek the results in a single comparison (e.g., Ricklefs (1996), Martins (1996), Guégan and Morand (1996)). Two procedures were used to calculate branch lengths in the cladistic tree. First, all branches have the same length, which is equivalent to a punctuational model of evolution. Then, the ages of taxa are proportional to the number of species they contain, which is similar to a gradual model of evolution (Purvis and Rambaut 1995).

The results generated from the two independent contrast models were similar, and we here illustrate (Figure 2) the results for the models based on branch lengths proportional to the number of *Mycobacterium* species that they contain. All regressions were forced through the origin as suggested by Garland et al. (1992). The standardisation of independent contrast values was checked by examination of absolute values of standardised contrasts *versus* their standard deviations (Garland 1992). Since some predictor variables were dependent on the year of description, this effect was withdrawn by regressing contrasts on time. The relationships between pathogenicity and the independent contrasts were tested using multiple linear regression analysis. The year of description variable was $\log_{10}$-transformed to normalise its distribution (Zar 1996).

Results

Sequence alignment and variation

The total length of the aligned sequence for the 64 *Mycobacterium* taxa was 1165 nucleotides. Two hundred and fifty-one variable sites (including gaps) were

Table 1. The *Mycobacterium* species-descriptive factors used in the present analysis. The data were generally extracted from Collins et al. (1984), Grosset et al. (1991), Falkinham (1996, 1998), Goodfellow and Magee (1998), Thorel et al. (1997), Vincent Levy-Frebault (1991), Wayne and Sramek (1992). Habitat diversity number corresponds to the living and non living reservoirs where the species could be isolated, for example fresh water, salt water, bovine, caprine, and human. Human pathogenicity was coded: 1 = non pathogen, 2 = rare pathogen, 3 = potential pathogen, 4 = pathogen, nc = not coded. Pigmentation was coded: 1 = Non pigmented, 2 = photochromogen 3 = scotochromogen, 4 = unknown.

Mycobacterium sp	Growth and helix 18 type	First description date	Habitat number	Human pathogenicity	Pigmentation	GeneBank accession number	Strain reference[8]
Slow growth with short helix 18							
M. genavense	1	1993	3	3	1	X60070	
M. heidelbergense[1]	1	1997	1	3	1	AJ000684	
M. intermedium	1	1993	1	3	2	X67847	
M. interjectum	1	1993	1	3	3	X70961	
M. lentiflavum[2]	1	1996	1	3	2	X80770	
M. simiae	1	1965	5	3	2	X52931	ATCC 25275[T]
M. triplex[3]	1	1996	1	2	1	U57632	
M. triviale	1	1970	2	2	1	X88924	ATCC 23292[T]
Slow growth with long helix 18							
M. asiaticum	2	1971	2	2	2	X55604	ATCC 25276[T]
M. avium subsp avium	2	1901	10	3	1	X52918	DSM 43212
M. avium subsp paratuberculosis	2	1923	4	nc	1	X52934	ATCC 19698[T]
M. bovis	2	1970	3	4	1	X55589	
M. celatum	2	1993	1	3	1	L08170	ATCC 51130
M. conspicuum	2	1995	1	3	3	X88922	DSM 44146
M. cookii	2	1990	2	1	3	X53896	ATCC 49103[T]
M. gastri	2	1966	3	2	1	X52919	ATCC 15754[T]
M. gordonae	2	1962	5	2	3	X52923	ATCC 14470[T]
M. haemophilum	2	1978	1	3	1	U06638	
M. hibernae	2	1993	2	1	3	X67096	ATCC 9874
M. intracellulare	2	1949	4	3	1	X52927	ATCC 15985
M. kansasii	2	1955	3	3	2	X15916	DSM 43224
M. leprae	2	1880	1	4	4	X55022	
M. malmoense	2	1977	3	3	1	X52930	ATCC 29571[T]
M. marinum	2	1926	4	3	2	X52920	
M. nonchromogenicum	2	1965	4	2	1	X52928	ATCC 19530[T]
M. scrofulaceum	2	1956	5	3	3	X52924	ATCC 19981[T]
M. shimodei	2	1982	1	2	1	X82459	ATCC 27962[T]
M. szulgai	2	1972	3	3	3	X52926	ATCC 25799
M. terrae	2	1966	6	2	1	X52925	ATCC 15755[T]
M. tuberculosis	2	1883	1	4	1	Z83862	
M. ulcerans	2	1950	2	3	1	X88926	
M. xenopi	2	1959	3	3	1	X52929	ATCC 19250[T]
Fast growth with short helix 18							
M. aichiense	3	1981	1	1	3	X55598	ATCC 27280[T]
M. alvei	3	1992	3	1	1	AF023664	CIP 103464
M. aurum	3	1966	2	1	3	X55595	ATCC 23366[T]
M. austroafricanum	3	1983	2	1	3	X93182	ATCC 33464[T]
M. chelonae	3	1923	7	3	1	M29559	
M. chitae	3	1967	1	1	1	X55603	ATCC 19627
M. chlorophenolicum	3	1984	1	1	3	X81926	NCIMB 12325[T]
M. chubuense	3	1981	1	1	3	X55596	ATCC 27278[T]
M. confluentis	3	1992	1	1	1	X63608	
M. diernhoferi	3	1983	1	1	1	X55593	ATCC 19340[T]
M. duvalii	3	1971	1	1	3	U94745	ATCC 43910[T]
M. fallax	3	1983	2	1	1	M29562	
M. farcinogenes	3	1973	1	nc	3	AF055333	ATCC 35753[T]
M. flavescens	3	1962	3	2	3	X52932	ATCC 14474[T]
M. fortuitum	3	1938	7	3	1	X52921	ATCC 14472
M. gadium	3	1974	1	2	3	X55594	ATCC 27726[T]

Table 1. (*continued*)

Mycobacterium sp	Growth and helix 18 type	First description date	Habitat number	Human pathogenicity	Pigmentation	GeneBank accession number	Strain reference[8]
M. gilvum	3	1971	1	1	3	X81996	ATCC 43909[T]
M. komossense	3	1979	1	1	3	X55591	ATCC 33013[T]
M. madagascariense	3	1992	1	1	3	X55600	
M. neoaurum	3	1972	2	2	3	M29564	
M. novocastrense	3	1997	1	3	2	U96747	
M. obuense	3	1981	1	1	3	X55597	ATCC 27023[T]
M. parafortuitum	3	1966	1	1	2	X93183	DSM 43528
M. perigrinum	3	1992	1	2	1	AF058712	ATCC 14467[T]
M. phlei	3	1899	4	1	3	M29566	
M. senegalense	3	1973	1	nc	3	M29567	
M. smegmatis	3	1889	4	2	1	X52922	ATCC 14468
M. sphagni	3	1980	1	1	3	X55590	ATCC 33026
M. thermoresistibile	3	1966	1	1	3	X55602	ATCC 19527[T]
M. tusciae[4]	3[6]	1999	1	3	3	AF055333	ATCC 35753[T]
M. vaccae	3	1964	2	1	2	X55601	ATCC 15483[T]
M. wolinskyi[5]	3	1999	1	2	1	Y12871	ATCC 700009

For [1]*M. heidelbergense*,[2]*M. lentiflavum*,[3]*M. triplex*,[4]*M. tusciae*,[5]*M. wolinskyi* the data respectively come from Haas et al. (1997), Springer et al. (1996), Floyd et al. (1996), Tortoli et al. (1999), Brown et al. (1999). [6]Although *M. tusciae* was previously described as a slow-growing species (Tortoli et al. 1999), it was coded 3 in the present study considering its short helix 18 and its fast growing phylogenetic cluster based on 16S rDNA sequence. [7]The GeneBank accession numbers of SSU rDNA sequences. [8]ATCC, American Type Culture Collection, Manassas,VA., USA; CIP, Collection Institut Pasteur, Paris, France; DSM, Deutsche Sammlung von Mikroorganismen, Braunschweig, Germany; NCIMB, National Collection of Industrial and Marine Bacteria, Aberdeen, Scotland,UK.

present, and among them 174 were informative for a cladistic analysis. The PAUP parsimony analysis yielded 390 most parsimonious trees (cladistic analysis) 1097 steps in length, with a consistency index of 0.356 and a retention index of 0.668. The strict consensus tree of these most parsimonious trees is illustrated in Figure 1. For clarity, we do not illustrate the neighbour-joining trees, since they yielded a similar tree structure to the cladistic approach (see above).

Phylogenetic analysis

In the cladistic tree (Figure 1), the ancestry of a group was not resolved. The group with slow-growing species/short and long helix 18 appeared as a sister group to the fast-growing species group. In the tree (Figure 1), the nodes of these two groups were sustained respectively by 6/10 and 6/9 synapomorphic characters. We observe the partitioning of the slow-growing species (i.e., with a short or long helix 18) into two sister groups. *M. triviale* was positioned outside its conventional group of slow-growing/short helix 18 and has 16/22 autapomorphic characters. It can be remarked that the slow-growing/long helix 18 group was a relatively strong group supported by 15/19 synapomorphic characters. The other group of species with many synapomorphies corresponded to the already known closely-related species or complex of species (also found by Goodfellow and Magee (1998)), as for example the complex *M. terrae-M. hiberniae-M. nonchromogenicum*, the complex *M. avium-M. intracellulare-M. paratuberculosis,* the two necrotic species *M. marinum-M. ulcerans* and the two pathogenic species *M. tuberculosis-M. bovis*. Moreover, these two last species shared 5/6 characters with *M. marinum-M. ulcerans*.

Comparative analysis

In a total number of 61 *Mycobacterium* species analysed in this study, we found that a mix of different variables could possibly cause interspecific differences in pathogenicity. The GLIM modelling (Table 2) demonstrated that a strong trend was associated with growth rate. Individual species with a growth rate coded 2 were more pathogenic (28,5%) than those coded 1 (7,6%) or 3 (13,1%). In addition, *Mycobacterium* species harbouring diversified niches were more pathogenic than specialists (Table 2). Thus, *Mycobacterium* species harbouring pigmentation coded 3 were slightly more pathogenic (4,9%) than those with pigmentation type 4 (2,7%), type 1 (1,2%) and type 2 (0,3%). The direct effect of the year of description was not significant (Table 2). Two

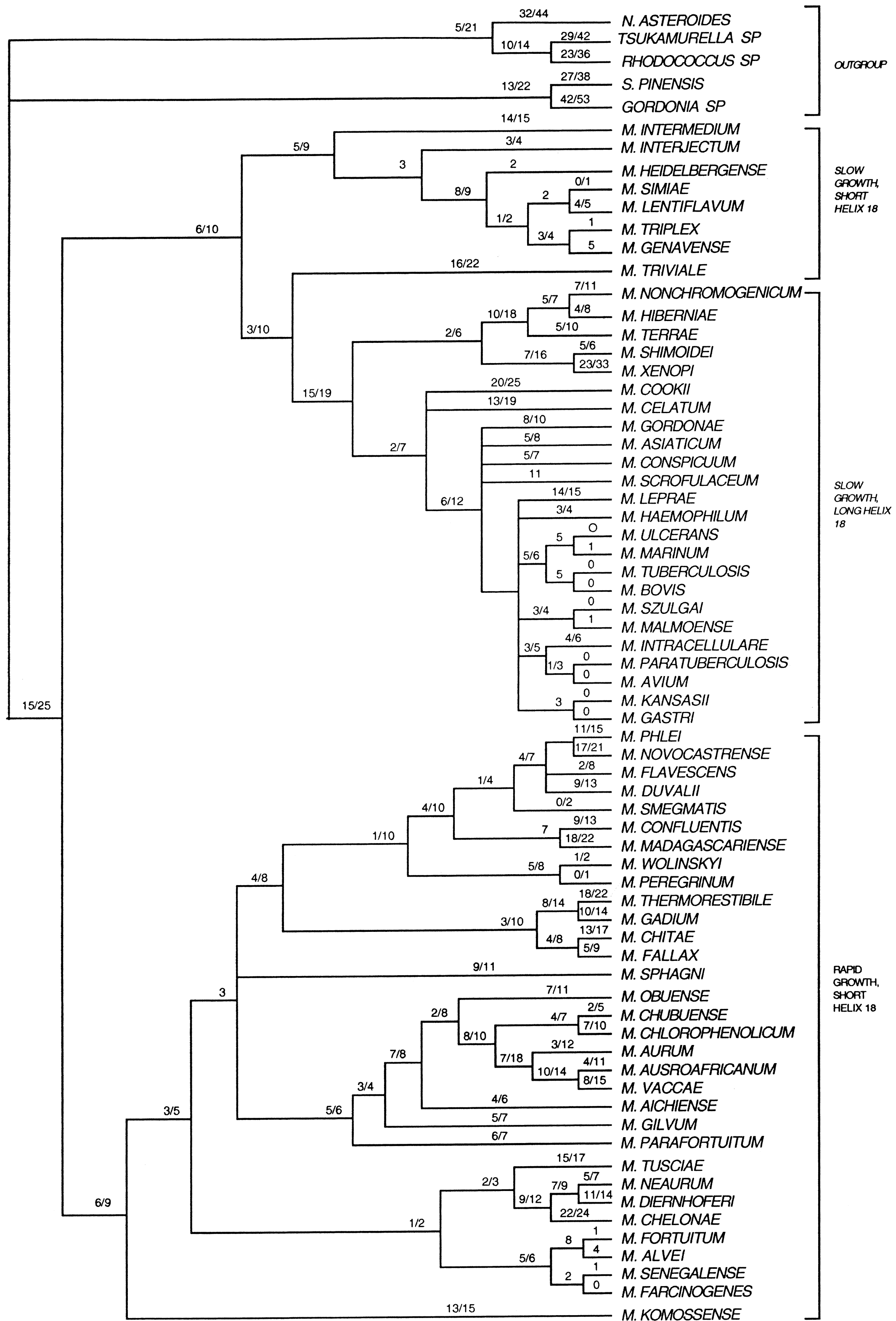

Figure 1. Strict consensus of the 390 most parsimonious trees (cladistic trees) using PAUP 3.1.1, heuristic search, gaps = fifth state) from the 1165 aligned positions of SSU rDNA of 64 *Mycobacterium* spp. Numbers on the branches indicate the minimal/maximal number of synapomorphies and autapomorphies according to the optimisation of informative sites.

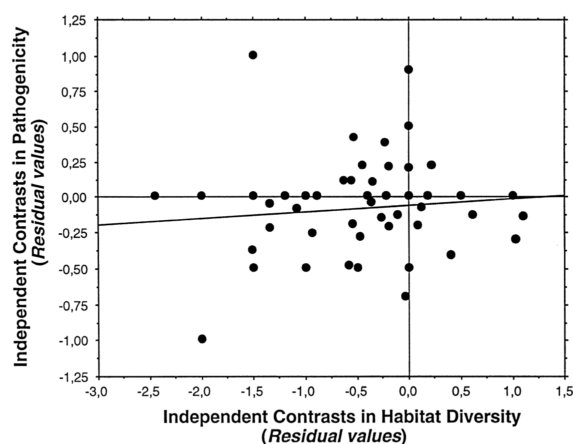

Figure 2. Pathogenicity levels in *Mycobacterium* species in relation to their habitat diversity expressed as the number of potential living or non-living reservoirs available to them. Values are statistically independent contrasts for four levels of pathogenicity (n = 61, r = 0.025, F = 0.480, p = 0.566) with the hypothesis that the ages of taxa are proportional to the number of species they contain. Data based on the assumption that all branches have the same length gave statistically similar results (see Materials and Methods and Table 2).

interaction terms were significant, indicating an effect of the year of description on the type of growth rate (Growth rate × year of description), and a difference in characterisation of pigmentation type with both the year of description and the growth rate (pigmentation × growth rate × year). An analysis considering three levels of pathogenicity instead of four yielded similar results, with the same group of variables explaining the variation of pathogenicity level.

An analysis using independent linear contrasts made on a parsimony tree (see Figure 1) with five outgroup species, demonstrated no significant relationship (n = 61 contrasts, R^2 = .035, F-ratio = .695, p = .559) between both the year of description and the habitat variables (and their interaction) and pathogenicity level in *Mycobacterium* species (see Table 3). Then, when considering the categorical variables (growth rate and pigmentation) as dependent parameters, any of the relationships corresponding to a change at nodes in the phylogeny from one categorical state to another were significant, tending to demonstrate that we could not reject the hypothesis that both growth rate and pigmentation characteristics were confounded by phylogeny. For clarity, we only illustrated the single relationship between independent contrasts of pathogenicity and independent contrasts of habitat diversity (Figure 2). An analysis considering only three levels of pathogenicity also showed a strong correlation between those traits and phylogeny (n = 61 contrasts, R^2 = .028, F-ratio = .553, p = .648).

Discussion

The main results of our study showed significant relationships between *Mycobacterium* pathogenicity

Table 2. Analysis of Deviance table for testing for effects of different factors plus all the different interaction terms on the level of disease pathogenicity. In the quasi-likelihood model, the dispersion parameter (ϕ) equals 0.528. Also given are the residual deviance, the residual degree of freedom in analysis (Res. df), the F-statistics (F) and the associated probabilities (p). A stepwise procedure yielded similar results, i.e., the same block of significant variables was retained.

	Residual deviance	Res. *df*	*F*	*p*
Independent variables				
Intercept	56.210	61	.	.
Habitat	51.443	60	12.659	0.0013
Growth rate	32.589	58	25.038	0.0000
Pigmentation	28.384	55	3.723	0.0218
Year of description (log)	26.941	54	3.833	0.0596
Growth rate × Habitat	25.378	52	2.075	ns
Pigmentation × Habitat	24.981	50	0.528	ns
Growth rate × Pigmentation	24.093	45	0.471	ns
Year of description × Habitat	24.028	44	0.173	ns
Growth rate × Year of description	20.038	42	5.298	0.0107
Pigmentation × Year of description	18.636	40	1.862	ns
Pigmentation × Growth rate × Habitat	16.292	37	2.075	ns
Year of description × Growth rate × Habitat	16.285	35	0.010	ns
Year of description × Pigmentation × Habitat	16.039	33	0.327	ns
Growth rate × Pigmentation × Year of description	11.538	31	5.977	0.0065
4-way interaction	11.296	30	0.644	ns
Residuals	11.296	30		

Table 3. Summary statistics of the independent linear phylogenetic contrasts of level of disease pathogenicity *versus* the independent linear phylogenetic contrasts of explanatory variables (n = 61 contrasts, R^2 = .035, F-ratio = .695, p = .559) using the parsimony tree as illustrated in Figure 1. Slope coefficients, standardised partial regression coefficients (Std. b), partial-F statistics (F) and associated probabilities (p) are given for each independent explanatory variable and their two-way interaction term. A step-down backward elimination procedure yielded similar results of no effect of selected predictors on disease pathogenicity contrasts.

	Slope coeff.	Std. b	F	p
Independent Phylogenetic contrasts				
Intercept	.030			
Time (log)	.002	.069	.216	.644
Habitat	−.137	−.316	2.242	.140
Time × Habitat	.005	.279	1.455	.233

and (i) growth rate, (ii) habitat diversity and (iii) the type of species pigmentation. However, none of these relationships remained significant when phylogenetic relationships were taken into account, suggesting that the pathogenic behaviour of mycobacteria was mainly phylogenetically inherited. This finding used comparative analysis to support the idea that an evolutionary change in pathogenicity goes hand-in-hand with a corresponding change in the number of habitats occupied by taxa, and that these correlated life-history traits are driven by common descent. Despite this, it remains possible that the adaptive value of inherited traits was maintained in derived groups if they still increased the pathogen fitness. In the *Verticillium* genus (Bidochka et al. 1999), plant pathogens having pectinase activity formed a separate clade from the insect, nematode and fungi pathogen species with chitinase activity. In addition, although virulent genes are undoubtedly partly shared by the taxa (synapomorphic characters), some virulence components are probably autapomorphic, given the important variability of pathologies between closely-related taxa (for example between the species *M. tuberculosis*/*M. bovis*, the species *M. ulcerans* /*M. marinum*, and the species *M.avium-M. paratuberculosis).*

The level of pathogenicity across species was reflected in the kind of growth rate, with species harbouring a growth rate type 2 (i.e., slow-growing/long helix 18 group) being more pathogenic than the other forms. However, the three growth rate categories exactly mirrored the three taxonomic groups accepted in the literature, so that it seemed possible that the significant relationship between the level of pathogenicity and the growth rate simply reflected the phylogenetic influence.

Habitat diversity played a significant role in the evolution of pathogenicity across taxa, with specific taxa being, on average, less pathogenic than their relatives, with a wider range of potential hosts (but see notable exceptions such as *M. tuberculosis* and *M. leprae*, Table 1). This was in accordance with the evolutionary idea that any specific virulent species having a strong impact on host fitness could not take refuge in another host, and thus would become extinct (Poulin 1998). Our results also suggested that within the slow-growing/long helix group of *Mycobacterium* species there was strong reason to regard the most derived taxa as the most pathogenic (e.g., *M. leprae*, *M. avium*, *M. bovis*, *M. tuberculosis . . .*).

Finally, the kind of pigmentation had an additional, but moderate, effect on the pathogenic response variable. Melanin is a potent free radical scavenger, and in the fungus *Cryptococcus neoformans* for instance, there is now good evidence that the production of this pigmentation is a significant virulence determinant (Hamilton and Holdom 1999). However, there is as yet no evidence that pigmentary components are involved in the pathogenicity of *Mycobacterium* species.

One limit to our study was that pathogenicity was measured qualitatively rather than quantitatively. In practice, the parasite-induced mortality or morbidity rate was not easy to estimate and, concerning the *Mycobacterium* genus, quantitative data were not available. Another reason why virulence measured as pathology could be misleading was that pathology was the result not only of parasite actions but also of its frequence and of host responses (Grange 1999; Poulin and Combes 1999). Although we could not precisely identify the virulent factors involved in the different *Mycobacterium* lineages (Legrand et al. 2000), it was clearly shown that the interdependence between the phylogenetic relationships of taxa and a genetic background was enough for the emergence of virulence factors. In addition, environmental conditions could be involved in the evolution of virulence, with the most harmful species being connected to the most diversified habitats. However, our study failed to demonstrate the effect of extrinsic factors due to *i*) a strong correlation between the number of possible habitats for a given *Mycobacterium* species and its position in the phylogenetic tree and *ii*) the strong conservativeness of the comparative method. It is

possible that an association between pathogenicity and phylogeny arose as a by-product of ecological conditions. Care has to be taken in interpreting the correlation between pathogenicity, number of occupied habitats and phylogeny in a cause-effect framework, because a comparative analysis does not usually allow us to determine the direction of causality. Thus, we have to understand the patterns obtained in this study as associations among a suite of traits, e.g., pathogenicity, niche breadth, and phylogenetic inheritance.

In summary, we are inclined to suggest that a strongly significant relationship does exist between pathogenicity, habitat diversity and phylogeny across *Mycobacterium* species, and that phylogenetic heritage and host habitat diversity may be important factors, both driving the evolution of pathogenicity in this group of pathogens. Verification of this needs other methods of analysis and possibly different datasets to test our hypothesis.

References

Agnew P. and Koella J.C. 1997. Virulence, parasite mode of transmission, and host fluctuating asymmetry. Proc. Roy. Soc. Lond.B 264: 9–15.

Barriel V. 1994. Phylogénies moléculaires et insertions-délétions de nucléotides. *C.R.A.S.* Paris, Sciences Vie/Life Sciences 317: 693–701.

Bidochka M.J., St. Leger R.J., Stuart A. and Gowanlock K. 1999. Nuclear rDNA phylogeny in the fungal genus *Verticillium* and its relationship to insect and plant virulence, extracellular proteases and carboxilases. Microbiol. 145: 955–963.

Bowman B.H. and Taylor J.W. 1993. Molecular phylogeny of pathogenic and non pathogenic originals. In: Reynolds D.R. and Taylor J.W. (eds), The fungal Holomorph: Mitotic, Meiotic and Pleomorphic speciation in fungal systematics. CAB international, Wallingford, pp. 169–178.

Bowman B.H., White T.J. and Taylor J.W. 1996. Human pathogenic fungi and their close nonpathogenic relatives. Mol. Phyl. Evol. 6: 89–96.

Brown B.A., Springer B., Steingrube V.A., Wilson R.W., Pfyffer G.E., Garcia M.J. et al. 1999. *Mycobacterium wolinskyi* sp. nov. and *Mycobacterium goodii* sp. nov. Two rapidly growing species related to *Mycobacterium smegmatis* and associated with human wound infections: a cooperative study from the International Working Group on Mycobacterial Taxonomy. Int. J. Syst. Bacteriol. 49: 1493–1511.

Collins C.H., Grange J.M. and Yates M.D. 1984. Mycobacteria in water. J. Applied Bacteriol. 57: 193–211.

Ebert D. 1994. Virulence and local adaptation of a horizontally transmitted parasite. Science 265: 1084–1086.

Ebert D. and Herre E.A. 1996. The evolution of parasitic diseases. Parasitol. Today 12: 96–101.

Ewald P.W. 1983. Host-parasite relations, vectors, and the evolution of disease severity. An. Rev. Ecol. Syst. 14: 465–485.

Ewald P.W. 1994. Evolution of Infectious Disease. Oxford University Press, Oxford, UK.

Falkinham J.O. III 1996. Epidemiology of infection by non tuberculosous mycobacteria. Clin. Microbiol. Rev. 9: 177–215.

Falkinham J.O. III 1998. Transmission of Mycobacteria. In: Gangadharam P.R.J. and Jenkins P.A. (eds), *Mycobacteria* I Basic Aspects. International Thomson Publishing, pp. 178–209.

Felsenstein J. 1985. Phylogenies and the comparative method. Am. Nat. 125: 1–15.

Floyd M.M., Guthertz L.S., Silcox V.A., Duffey P.S., Jang Y., Desmond E.P. et al. 1996. Characterization of an SAV organism and proposal of *Mycobacterium triplex* sp. nov. J. Clin. Microbiol. 34: 2963–2967.

Garland T. Jr 1992. Rate tests for phenotypic evolution using phylogenetically independent contrasts. Am. Nat. 140: 509–519.

Garland T. Jr, Harvey P.H. and Ives A.R. 1992. Procedures for the analysis of comparative data using phylogenetically independent contrast. Syst. Biol. 41: 18–32.

Goodfellow M. and Magee J.G. 1998. Taxonomy of mycobacteria. In: Gangadharam P.R.J. and Jenkins P.A. (eds), Mycobacteria I Basic Aspects. International Thomson Publishing, pp. 1–71.

Grange J.M. 1999. Pathogenesis of mycobacterial disease. In: Gangadharam P.R.J. and Jenkins P.A. (eds), *Mycobacteria* I Basic Aspects. International Thomson Publishing, pp. 145–177.

Grosset J., Truffot-Pernot Ch., Boisvert H. and Lalande V. 1991. Qu'est ce que les mycobactéries atypiques? Méd. Mal. Infect 21 Sp.: 7–15.

Guégan J.F. and Morand S. 1996. Polyploid hosts: strange attractors for parasites? Oikos 77: 366–370.

Haas W.H., Butler W.R., Kirschner P., Plikaytis B.B., Coyle M.B., Amthor B. et al. 1997. A new agent of mycobacterial lymphadenitis in children: *Mycobacterium heidelbergense* sp. nov. J. Clin. Microbiol. 35: 3203–3209.

Hamilton A.J. and Holdom M.D. 1999. Antioxidant systems in the pathogenic fungi of man and their role in virulence. Med. Mycol. 37: 375–389.

Harvey P.H. and Pagel M.D. 1991. The Comparative Method in Evolutionary Biology. Oxford University Press, Oxford, UK.

Johnson P.D.R., Stinear T.P. and Hayman J.A. 1999. *Mycobacterium ulcerans*-a mini review. J. Med. Microbiol. 48: 511–513.

Kuhnert P., Boerlin P. and Frey J. 2000. Target genes for virulence assessment of *Escherichia coli* isolates from water, food and the environment. FEMS Microbiol. Rev. 24: 107–117.

Legrand E., Sola C. and Rastogi N. 2000. Le complexe *Mycobacterium avium-intracellulare*: marqueurs phénotypiques et génotypiques et les bases moléculaires de la transmission interespèces. Bull. Soc. Pathol. Exot. 93: 182–192.

Lipsitch M. and Moxon E.R. 1997. Virulence and transmissibility of pathogens: what is the relationship? Trends in Microbiol. 5: 31–37.

Martins E.P. 1996. Phylogenies and the Comparative Method in Animal Behavior. Oxford University Press, Oxford, UK.

Maynard Smith J., Feil E.J. and Sith N.H. 2000. Population structure and evolutionary dynamics of pathogenic bacteria. BioEssays 22: 115–1122.

McCullagh P. and Nelder J.A. 1989. Generalized Linear Models. 2nd edn. Chapman & Hall Ltd., London, UK.

Pagel M.D. 1992. A method for the analysis of comparative data. J. Theor. Biol. 156: 431–442.

Philippe H. 1993. MUST: a computer package of Management Utilities for Sequences and Trees. Nucleic Acids Res. 21: 5264–5272.

Picard B., Garcia J.S., Gouriou S., Duriez P., Brahimi N., Bingen E. et al. 1999. The link between phylogeny and virulence in *Escherichia coli* extra-intestinal infection. Infect. Immun. 67: 546–553.

Portaels F. 1987. Le sida et les mycobactérioses atypiques. Ann. Soc. Belge Méd. Trop. 67: 93–113.

Poulin R. 1998. Evolutionary Ecology of Parasites. From Individuals to Communities. Oxford University Press, Oxford, UK.

Poulin R. and Combes C. 1999. The concept of virulence: interpretations and implications. Parasitol. Today 15: 474–475.

Purvis A. and Rambaut A. 1995. Comparative analysis of independent contrasts (CAIC): an Apple Macintosh application for analysing comparative data. Comput. Applied Biosc. 11: 247–251.

Ricklefs R.E. 1996. Applications of phylogenetically independent contrasts: a mixed progress report. Oikos 77: 167–172.

Rogall T., Wolters J., Thomas F. and Böttger E.C. 1990. Towards a phylogeny and definition of species at the molecular level within the genus *Mycobacterium*. Int. J. Syst. Bacteriol. 40: 323–330.

Runyon E.H. 1959. Anonymous mycobacteria in pulmonary disease. Med. Clin. North Am. 43: 273–290.

Spratt B.G. and Maiden M.C. 1999. Bacterial population genetics, evolution and epidemiology. Phil. Trans. Roy. Soc. Lond. B. Biol. Sc. 354: 701–710.

Springer B., Wu W.K., Bodmer T., Haase G., Pfyffer G.E., Kroppenstedt R.M. et al. 1996. Isolation and characterization of a unique group of slowly growing mycobacteria: description of *Mycobacterium lentiflavum sp.* nov. J. Clin. Microbiol. 34: 1100–1107.

Stahl D.A. and Urbance J.W. 1990. The division between fast and slow growing species corresponds to natural relationships among the *Mycobacteria*. J. Bacteriol. 172: 116–124.

Stearns S.C. 1999. Evolution in Health and Disease. Oxford University Press, UK.

Swofford D.L. 1993. PAUP: Phylogenetic Analysis Using Parsimony, Version 3.1, Computer Program distributed by the Illinois Natural History Survey. Champaign, Illinois.

Thorel M.F., Huchzermeyer H., Weiss R. and Fontaine J.J. 1997. *Mycobacterium avium* infections in animals. Lit. Rev., Vet. Res. 28: 439–447.

Tortoli E., Kroppenstedt R.M., Bartoloni A., Caroli G., Jan I., Pawlowski J. et al. 1999. *Mycobacterium tusciae* sp. nov. Int. J. Syst. Bacteriol. 49: 1839–1844.

Tsukamura M. 1984. The nonpathogenic species of *Mycobacteria*. Their distribution and ecology in non-living reservoirs. In: Kubica G.P. and Wayne L.G. (eds), The Mycobacteria, a Sourcebook Part B. Marcel Drekker INC, New York and Basel, pp. 1339–1359.

Venables W.N. and Ripley B.D. 1994. Modern Applied Statistics with S-Plus. Springer Verlag, Berlin.

Vincent Levy-Frebault V. 1991. Ecologie des mycobactéries et mode de contamination humaine. Med. Mal. Infect 21 Sp.: 16–25.

Wayne L.G. and Sramek H.A. 1992. Agents of newly recognized or infrequently encountered mycobacterial diseases. Clin. Microbiol. Rev. 5: 1–25.

Wilson K. and Grenfell B.T. 1997. Generalized linear modelling for parasitologists. Parasitol. Today 13: 33–38.

Zar J.H. 1996. Biostatistical Analysis. 3rd edn. Prentice Hall INC, Upper Saddle River, NJ.