HHS Public Access

Author manuscript

Nat Genet. Author manuscript; available in PMC 2012 July 01.

Published in final edited form as:

Nat Genet. ; 44(1): 106–110. doi:10.1038/ng.1038.

Whole-genome sequencing of rifampicin-resistant *M. tuberculosis* strains identifies compensatory mutations in RNA polymerase

Iñaki Comas[1,a], **Sonia Borrell**[2,3], **Andreas Roetzer**[4], **Graham Rose**[1], **Bijaya Malla**[2,3], **Midori Kato-Maeda**[5], **James Galagan**[6,7], **Stefan Niemann**[4], and **Sebastien Gagneux**[2,3,*]

[1]Medical Research Council, National Institute for Medical Research, London, UK [2]Swiss Tropical and Public Health Institute, Basel, Switzerland [3]University of Basel, Basel, Switzerland [4]Research Centre Borstel, Molecular Mycobacteriology, Borstel, Germany [5]University of California San Francisco, San Francisco, USA [6]Broad Institute, Cambridge, USA [7]Boston University, Boston, USA

Abstract

Drug-resistant bacteria are emerging worldwide, despite frequently being less fit than drug-susceptible strains[1]. Data from model systems suggest the fitness cost of antimicrobial resistance can be mitigated by compensatory mutations[2]. However, current evidence that compensatory evolution plays any significant role in the success of drug-resistant bacteria in human populations is weak[3–6]. Here we describe a set of novel compensatory mutations in the RNA polymerase of rifampicin-resistant *Mycobacterium tuberculosis*, the etiologic agent of human tuberculosis (TB). *M. tuberculosis* strains harbouring these compensatory mutations exhibited a high competitive fitness *in vitro*. Moreover, these mutations were associated with high *in vivo* fitness as determined by their relative clinical frequency across patient populations. Importantly, in countries with the world's highest incidence of multidrug-resistant (MDR) TB[7], more than 30% of MDR clinical isolates had such a mutation. Our findings support a role for compensatory evolution in the global epidemics of MDR-TB[8].

*Corresponding author: Sebastien GAGNEUX, Ph.D., Department of Medical Parasitology & Infection Biology, Swiss Tropical & Public Health Institute, Socinstrasse 57, 4002 Basel, Switzerland, Phone: +41 -61-284-8369, Fax: +41 -61-284-8101 sebastien.gagneux@unibas.ch.

[a]Current affiliation: Genomics and Health Unit, Centre for Public Health Research, Valencia, Spain

Database accession numbers

SRP001097

Author contributions

I.C., S.B. and S.G. planned the experiments. I.C., S.B., A.R., B.M., G.R., M.K., J.G. and S.G. performed the experiments. I.C., S.B., A.R., G.R., S.N. and S.G. analysed the data. I.C., S.B. and S.G. wrote the manuscript. All authors critically reviewed the manuscript.

Conflict of interest statement

The authors declare that they have no competing financial interests.

Cited URLs

http://www.sanger.ac.uk/resources/downloads/bacteria/mycobacterium.html

The worldwide emergence of multidrug-resistant (MDR) strains of *Mycobacterium tuberculosis* is threatening to make one of humankind's most important infectious diseases incurable[8]. MDR strains of *M. tuberculosis* are resistant to isoniazid and rifampicin, the two most important anti-TB drugs. Theory predicts that one of the key factors driving the current epidemics of MDR-TB is the relative fitness of drug-resistant strains compared to drug-susceptible forms[9,10]. Experimental work has shown that drug resistance in bacteria is often associated with a fitness deficit[1,2,11–13], but some drug resistance-conferring mutations cause little- or no loss of fitness[12,14,15]. Furthermore, fitness cost linked to drug resistance can be reduced by compensatory evolution[2,14]. However, little data exist on the clinical relevance of this phenomenon[3,6]. In *M. tuberculosis*, compensatory mechanisms were identified for fitness defects related to isoniazid and aminoglycoside resistance[16,17]. However, the corresponding compensatory mutations are rare in clinical strains[18], suggesting they play a minor role in the epidemiology of MDR-TB. While some work on the compensatory evolution in resistance to rifampicin has been reported for *Escherichia coli*[14], nothing is known with respect to compensatory evolution in rifampicin-resistant *M. tuberculosis*.

Rifampicin binds to the beta-subunit of the RNA polymerase encoded by *rpoB* and inhibits transcription. More than 95% of *M. tuberculosis* clinical strains resistant to rifampicin harbour a mutation in an 81-basepair region of *rpoB* known as the rifampicin resistance determining region (RRDR)[18]. These mutations convey high-level resistance. We have previously shown that all laboratory-generated mutants of *M. tuberculosis* harbouring a rifampicin resistance-conferring mutation in RRDR suffer a significant fitness defect compared to their drug-susceptible ancestors when grown in the absence of rifampicin[12]. By contrast, some *M. tuberculosis* clinical strains isolated from TB patients who developed rifampicin resistance during treatment exhibited no fitness cost compared to their rifampicin-susceptible counterparts, despite carrying the same *rpoB* mutation as some of the laboratory-derived strains[12]. At the time, we hypothesized that these clinical strains might have acquired compensatory mutations during patient treatment.

Here we tested this hypothesis by comparing the genome sequences of 10 paired clinical rifampicin-resistant isolates to the genomes of the corresponding rifampicin-susceptible isolates recovered from the same patient (Supplementary Table 1)[12]. We extracted all non-synonymous and intergenic mutations found only in the rifampicin-resistant genomes (Supplementary Table 2). In addition, we experimentally evolved six laboratory-derived rifampicin-resistant mutants and their rifampicin-susceptible ancestors[12] during 45 weeks of serial sub-culture in the absence of rifampicin (Supplementary Table 3). Comparison of the whole-genome sequences of these *in vitro* evolved strains to their respective un-evolved rifampicin-susceptible ancestors allowed us to identify putative compensatory mutations, as well as mutations likely to represent adaptations to growth in the laboratory (Supplementary Table 4). Of note, all of the *in vitro* evolved rifampicin-resistant strains maintained their original *rpoB* mutation, which is consistent with a higher number of mutational targets for compensation compared to reversion[2,19].

After combining our clinical and *in vitro* data, and excluding mutations representing laboratory adaptations or phylogenetic markers (Supplementary Tables 4 and 5), we

identified 54 putative compensatory mutations in 38 genes and 10 intergenic regions (Supplementary Table 6). *RpoA* and *rpoC* stood out by harbouring multiple mutations in the evolved strains (1 strain) or the paired clinical strains (4 strains; Table 1, Fig. 1). These genes encode the α- and β′-subunits of the RNA polymerase, respectively. Based on the known interactions between the RpoA, RpoB, and RpoC subunits[20], we reasoned that non-synonymous changes in *rpoA* and *rpoC* occurring only in rifampicin-resistant genomes were likely compensatory. Mapping these mutations onto the 3D-structure of the *E. coli* RNA polymerase[20] showed them localized at the interface between the α- and β′-subunits (Fig. 2), indicating a potential impact on the interaction between these subunits.

In addition to these plausible effects on RNA polymerase structure, we expect compensatory mutations in rifampicin-resistant *M. tuberculosis* i) to occur frequently in MDR clinical isolates, ii) not to occur in rifampicin-susceptible isolates, iii) to be associated with mutations in RRDR, and iv) not to occur in rifampicin-resistant strains without *rpoB* mutations. Because *M. tuberculosis* is genetically monomorphic[21], with no ongoing horizontal gene transfer[22,23], rates of convergent evolution in this microbe are extremely low[24,25]. By contrast, drug resistance-conferring mutations exhibit convergent evolution, as drug pressure selects for the same mutations across the different phylogenetic lineages of *M. tuberculosis*[26]. Following this rational, we expect compensatory mutations also to show convergent evolution. Furthermore, because *M. tuberculosis* is genetically homogeneous, more than two amino acid variants at the same codon position are rare[27], except in the context of drug resistance[18]. Thus, we also expect particular codon positions involved in compensation to harbour multiple alleles, as several alternative amino acid substitutions might have similar compensatory effects.

To test these predictions, we screened four complementary panels of clinical *M. tuberculosis* strains for non-synonymous changes in *rpoA* and *rpoC*. The first panel comprised 117 MDR strains from global sources, representing five of the six major lineages of human-adapted *M. tuberculosis*[28] (Supplementary Table 7). The second panel served as a control and included 131 rifampicin-susceptible strains representing the global diversity of *M. tuberculosis*[29,30]. The third panel consisted of 212 MDR clinical isolates from Abkhazia/Georgia, Uzbekistan and Kazakhstan (Supplementary Table 7)[31–33]; these countries are among the regions with the highest MDR-TB incidence in the world[8]. The fourth panel comprised 40 pan-susceptible isolates from Uzbekistan (Supplementary Table 8). All of the 329 MDR strains included in panel 1 and 3 had phenotypically confirmed rifampicin resistance, and 321/332 (99.7%) of them harboured at least one non-synonymous mutation in RRDR.

Our results showed that after exclusion of phylogenetic markers and mutations likely due to laboratory adaptation, 89/329 (27.1%) of all MDR strains harboured a non-synonymous mutation in *rpoA* or *rpoC*. In addition to the mutations already observed in our clinically paired or experimentally evolved strains, we found 28 additional non-synonymous changes in these genes (Table 1). By contrast, none of the 171 rifampicin-susceptible control strains harboured any of these mutations. Furthermore, all MDR strains harbouring an *rpoA* or *rpoC* mutation also had a mutation in RRDR, whereas none of the 11 rifampicin-resistant strains without *rpoB* mutation had any mutation in *rpoA* or *rpoC*.

When combining all our data, we found that 11 codon positions in *rpoA* or *rpoC* had the same putative compensatory mutations in more than one phylogenetic lineage of *M. tuberculosis*, and 8 codon positions were found to have more than one amino acid change (Fig. 1). As discussed above, these phenomena are rarely observed in *M. tuberculosis* outside of drug resistance, and positions exhibiting both phenomena by chance are particularly unlikely. Hence, we focused the remaining of our investigation on the mutations falling in codon positions that satisfied both of these criteria (Table 1).

We computationally predicted the effect of these high-probability compensatory mutations (HCMs) on protein function, by comparing the degree of evolutionary conservation of the orthologous protein positions in other bacteria using SIFT scores [34]. As a comparison, we used 13 publicly available mycobacterial genomes not belonging to the *M. tuberculosis* complex. As a proof of concept, we first tested whether we could correctly predict that mutations in *rpoB* known to convey rifampicin resistance (Supplementary Table 7) were more likely to be functional than phylogenetic markers in the same gene (Supplementary Table 5); we found this to be the case (Mann-Whitney U Test $p < 0.01$). When testing the HCMs in *rpoA* and *rpoC* (Table 1), we found that these mutations were also predicted to be more functional compared to phylogenetic markers found in the same genes (Supplementary Table 5, Mann-Whitney U Test $p < 0.01$).

To test whether the predicted functional effects of HCMs correlated with strain fitness, we combined our new data on the occurrence of these mutations with our older data on the relative fitness of rifampicin-resistant *M. tuberculosis*[12]. We found that three out of four clinical MDR strains with no competitive fitness defect harboured an HCM (Fig. 3A). By contrast, none of the six clinical strains with a statistically significantly reduced fitness harboured any HCM (Fischer's exact test $p < 0.05$). Furthermore, when calculating the difference in fitness between the laboratory-derived rifampicin-resistant strains and the clinical strains harbouring the same rifampicin resistance-conferring mutation and belonging to the same phylogenetic lineage, we found that this difference was always towards increased fitness in the clinical strains harbouring an HCM, and larger compared to the median difference in fitness among the other clinical strains (Fig. 3B; Mann-Whitney U test $p < 0.05$). Finally, we found that the median time between the isolation of the susceptible patient isolate and the follow-up resistant isolate was longer for strains harbouring an HCM than for strains carrying only an RRDR mutation (20 months versus 6 months; Fig. 3C; Mann-Whitney U test $p < 0.05$). Taken together, these data show that HCMs in *rpoA* and *rpoC* are associated with a high *in vitro* fitness of MDR clinical strains of *M. tuberculosis*, and that the emergence of HCMs is time-dependent.

One could argue that because the three clinical strains harbouring HCMs have accumulated additional mutations (Supplementary Table 2), the high fitness of these strains cannot directly be attributed to HCMs. While not discarding possible alternative compensatory mechanisms, at least for clinical Pair 3 (Fig. 3a), several observations support a causal relationship between the HCM observed in *rpoC* and increased fitness. Specifically, this strain contains only one additional non-synonymous mutation compared to its rifampicin-susceptible ancestor (Supplementary Table 2). However, this additional mutation occurs in *aroG*, a gene which does not belong to the transcription functional class and for which no

interactions with RNA polymerase subunits are known according to the latest version of the STRING database[35]. Moreover, when performing a SIFT analysis as outlined above[34], the *aroG* change M311T was predicted to have no functional consequence (SIFT score=1.00). Of note, several environmental mycobacteria show the same amino acid substitution, providing additional evidence against a putative role of this *aroG* mutation in compensatory evolution. In sum, the HCM appears solely responsible for the high fitness observed in this strain.

We and others have shown that for rifampicin resistance, *in vitro* competitive fitness in *M. tuberculosis* correlates with *in vivo* fitness as measured by the frequency of different rifampicin resistance-conferring mutations in clinical settings[12,13]. In other words, rifampicin resistance-conferring mutations associated with no- or low fitness cost *in vitro* are the most frequent in clinical strains. Hence, in the context of rifampicin resistance, clinical frequency of mutations can be used as a proxy for *in vivo* fitness of drug-resistant *M. tuberculosis* among different patient populations[4,15]. When we determined the proportion of HCMs in MDR clinical strains, we found that 12.0% of our global panel of MDR isolates carried such a mutation (Fig. 3D). This proportion increased to 21.3% in our strain panel from countries with a high MDR-TB burden (Chi2=4.5, $p < 0.05$). When we relaxed our selection criteria and repeated this analysis with all putative compensatory mutations in *rpoA* and *rpoC* (Table 1), we found that 19.7% of the global MDR strains carried such a mutation, compared to 31.3% of MDR strains from high MDR-TB burden countries (Chi2=5.14, $p < 0.05$). The high proportion of compensatory mutations in strains from Abkhazia/Georgia, Uzbekistan and Kazakhstan is consistent with the success of MDR strains in these countries, where up to 50% of TB patients are estimated to carry MDR strains, compared to a global average of only 3%[7].

In conclusion, our results suggest that the emergence over time of particular mutations in *rpoA* or *rpoC* of rifampicin-resistant *M. tuberculosis* lead to MDR strains with high fitness. Furthermore, our data show that these mutations occur at high frequencies in clinical settings, particularly in hotspot regions of MDR-TB[9]. Future studies will tell whether MDR strains of *M. tuberculosis* harbouring mutations in *rpoA* or *rpoC* are particularly transmissible, and how these mutations contribute to the success of these strains. Moreover, targeted genotyping will enable TB control programmes to focus on the most transmissible MDR strains. Our findings also suggest that mathematical models aiming at predicting the future of the global MDR-TB epidemic should consider the effects of compensatory mutations, as well as the time necessary for such mutations to emerge.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank P. Small and C. Davis Long for stimulating discussions at the onset of this project, D. Young and M. Coscolla for reviewing the manuscript, A. Candel for advice in the interpretation of the RNA polymerase molecular structure, and T. Van for technical support. We acknowledge the Wellcome Trust Sanger Institute for making available unpublished DNA sequence data (see URL). We would like to thank T. Ubben, I. Razio and the other members of the German National Reference Center for Mycobacteria (Borstel, Germany) for technical assistance

and all partners that have contributed to previous studies in high MDR-TB incidence regions and collected part of the strains analyzed here. This project has been funded in whole or in part with Federal funds from the National Institute of Allergy and Infectious Disease, National Institutes of Health, Department of Health and Human Services, under Contract No. HHSN266200400001C. This work was further supported by the Medical Research Council, UK (MRC_U117588500), the Swiss National Science Foundation (PP00A-119205), the National Institutes of Health (AI090928 and AI034238), and the European Community LONG-DRUG (QLK-CT-2002-01612) and TB-PAN-NET (FP7-223681) projects.

References

1. Andersson DI, Levin BR. The biological cost of antibiotic resistance. Curr Opin Microbiol. 1999; 2:489–493. [PubMed: 10508723]
2. Andersson DI, Hughes D. Antibiotic resistance and its cost: is it possible to reverse resistance? Nat Rev Microbiol. 2010; 8:260–271. [PubMed: 20208551]
3. Bjorkholm B, et al. Mutation frequency and biological cost of antibiotic resistance in *Helicobacter pylori*. Proc Natl Acad Sci U S A. 2001; 98:14607–14612. [PubMed: 11717398]
4. Bottger EC, Springer B, Pletschette M, Sander P. Fitness of antibiotic-resistant microorganisms and compensatory mutations. Nat Med. 1998; 4:1343–1344. [PubMed: 9846553]
5. Borrell S, Gagneux S. Infectiousness, reproductive fitness and evolution of drug-resistant *Mycobacterium tuberculosis*. Int J Tuberc Lung Dis. 2009; 13:1456–1466. [PubMed: 19919762]
6. Nagaev I, Bjorkman J, Andersson DI, Hughes D. Biological cost and compensatory evolution in fusidic acid-resistant *Staphylococcus aureus*. Mol Microbiol. 2001; 40:433–439. [PubMed: 11309125]
7. World Health Organization. Global tuberculosis control - surveillance, planning, financing. Geneva, Switzerland: WHO; 2010.
8. Gandhi NR, et al. Multidrug-resistant and extensively drug-resistant tuberculosis: a threat to global control of tuberculosis. Lancet. 2010; 375:1830–1843. [PubMed: 20488523]
9. Blower SM, Chou T. Modeling the emergence of the 'hot zones': tuberculosis and the amplification dynamics of drug resistance. Nat Med. 2004; 10:1111–1116. [PubMed: 15378053]
10. Cohen T, Murray M. Modeling epidemics of multidrug-resistant *M. tuberculosis* of heterogeneous fitness. Nat Med. 2004; 10:1117–1121. [PubMed: 15378056]
11. Mariam DH, Mengistu Y, Hoffner SE, Andersson DI. Effect of *rpoB* mutations conferring rifampin resistance on fitness of *Mycobacterium tuberculosis*. Antimicrob Agents Chemother. 2004; 48:1289–1294. [PubMed: 15047531]
12. Gagneux S, et al. The competitive cost of antibiotic resistance in *Mycobacterium tuberculosis*. Science. 2006; 312:1944–1946. [PubMed: 16809538]
13. Billington OJ, McHugh TD, Gillespie SH. Physiological cost of rifampin resistance induced in vitro in *Mycobacterium tuberculosis*. Antimicrob Agents Chemother. 1999; 43:1866–1869. [PubMed: 10428904]
14. Reynolds MG. Compensatory evolution in rifampin-resistant *Escherichia coli*. Genetics. 2000; 156:1471–1481. [PubMed: 11102350]
15. Sander P, et al. Fitness cost of chromosomal drug resistance-conferring mutations. Antimicrob Agents Chemother. 2002; 46:1204–1211. [PubMed: 11959546]
16. Sherman DR, et al. Compensatory *ahpC* gene expression in isoniazid-resistant *Mycobacterium tuberculosis*. Science. 1996; 272:1641–1643. [PubMed: 8658136]
17. Shcherbakov D, et al. Directed mutagenesis of *Mycobacterium smegmatis* 16S rRNA to reconstruct the in-vivo evolution of aminoglycoside resistance in *Mycobacterium tuberculosis*. Mol Microbiol. 2010; 77:830–840. [PubMed: 20545852]
18. Ramaswamy S, Musser JM. Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*: 1998 update. Tuber Lung Dis. 1998; 79:3–29. [PubMed: 10645439]
19. Levin BR, Perrot V, Walker N. Compensatory mutations, antibiotic resistance and the population genetics of adaptive evolution in bacteria. Genetics. 2000; 154:985–997. [PubMed: 10757748]
20. Opalka N, et al. Complete structural model of *Escherichia coli* RNA polymerase from a hybrid approach. PLoS Biology. 2010; 8:e1000483. [PubMed: 20856905]

21. Achtman M. Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens. Annu Rev Microbiol. 2008; 62:53–70. [PubMed: 18785837]
22. Hirsh AE, Tsolaki AG, DeRiemer K, Feldman MW, Small PM. Stable association between strains of *Mycobacterium tuberculosis* and their human host populations. Proc Natl Acad Sci U S A. 2004; 101:4871–4876. [PubMed: 15041743]
23. Supply P, et al. Linkage disequilibrium between minisatellite loci supports clonal evolution of *Mycobacterium tuberculosis* in a high tuberculosis incidence area. Mol Microbiol. 2003; 47:529–538. [PubMed: 12519202]
24. Comas I, Homolka S, Niemann S, Gagneux S. Genotyping of genetically monomorphic bacteria: DNA sequencing in *Mycobacterium tuberculosis* highlights the limitations of current methodologies. PLoS ONE. 2009; 4:e7815. [PubMed: 19915672]
25. Hershberg R, et al. High functional diversity in *Mycobacterium tuberculosis* driven by genetic drift and human demography. PLoS Biology. 2008; 6:e311. [PubMed: 19090620]
26. Hazbon MH, et al. Convergent evolutionary analysis identifies significant mutations in drug resistance targets of *Mycobacterium tuberculosis*. Antimicrob Agents Chemother. 2008; 52:3369–3376. [PubMed: 18591265]
27. Comas I, et al. Human T cell epitopes of *Mycobacterium tuberculosis* are evolutionarily hyperconserved. Nat Genet. 2010; 42:498–503. [PubMed: 20495566]
28. Gagneux S, Small PM. Global phylogeography of *Mycobacterium tuberculosis* and implications for tuberculosis product development. Lancet Infect Dis. 2007; 7:328–337. [PubMed: 17448936]
29. Gagneux S, et al. Variable host-pathogen compatibility in *Mycobacterium tuberculosis*. Proc Natl Acad Sci U S A. 2006; 103:2869–2873. [PubMed: 16477032]
30. Wirth T, et al. Origin, spread and demography of the *Mycobacterium tuberculosis* complex. PLoS Pathog. 2008; 4:e1000160. [PubMed: 18802459]
31. Kubica T, et al. The Beijing genotype is a major cause of drug-resistant tuberculosis in Kazakhstan. Int J Tuberc Lung Dis. 2005; 9:646–653. [PubMed: 15971392]
32. Cox HS, et al. The Beijing genotype and drug resistant tuberculosis in the Aral Sea region of Central Asia. Respir Res. 2005; 6:134. [PubMed: 16277659]
33. Pardini M, et al. Characteristics of drug-resistant tuberculosis in Abkhazia (Georgia), a high-prevalence area in Eastern Europe. Tuberculosis (Edinb). 2009; 89:317–324. [PubMed: 19539531]
34. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res. 2003; 31:3812–3814. [PubMed: 12824425]
35. Szklarczyk D, et al. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. Nucleic Acids Res. 2011; 39:D561–D568. [PubMed: 21045058]
36. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010; 38:e164. [PubMed: 20601685]
37. Guerois R, Nielsen JE, Serrano L. Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. J Mol Biol. 2002; 320:369–387. [PubMed: 12079393]
38. Bonfield JK, Smith K, Staden R. A new DNA sequence assembly program. Nucleic Acids Res. 1995; 23:4992–4999. [PubMed: 8559656]
39. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. Mol Biol Evol. 2007; 24:1596–1599. [PubMed: 17488738]

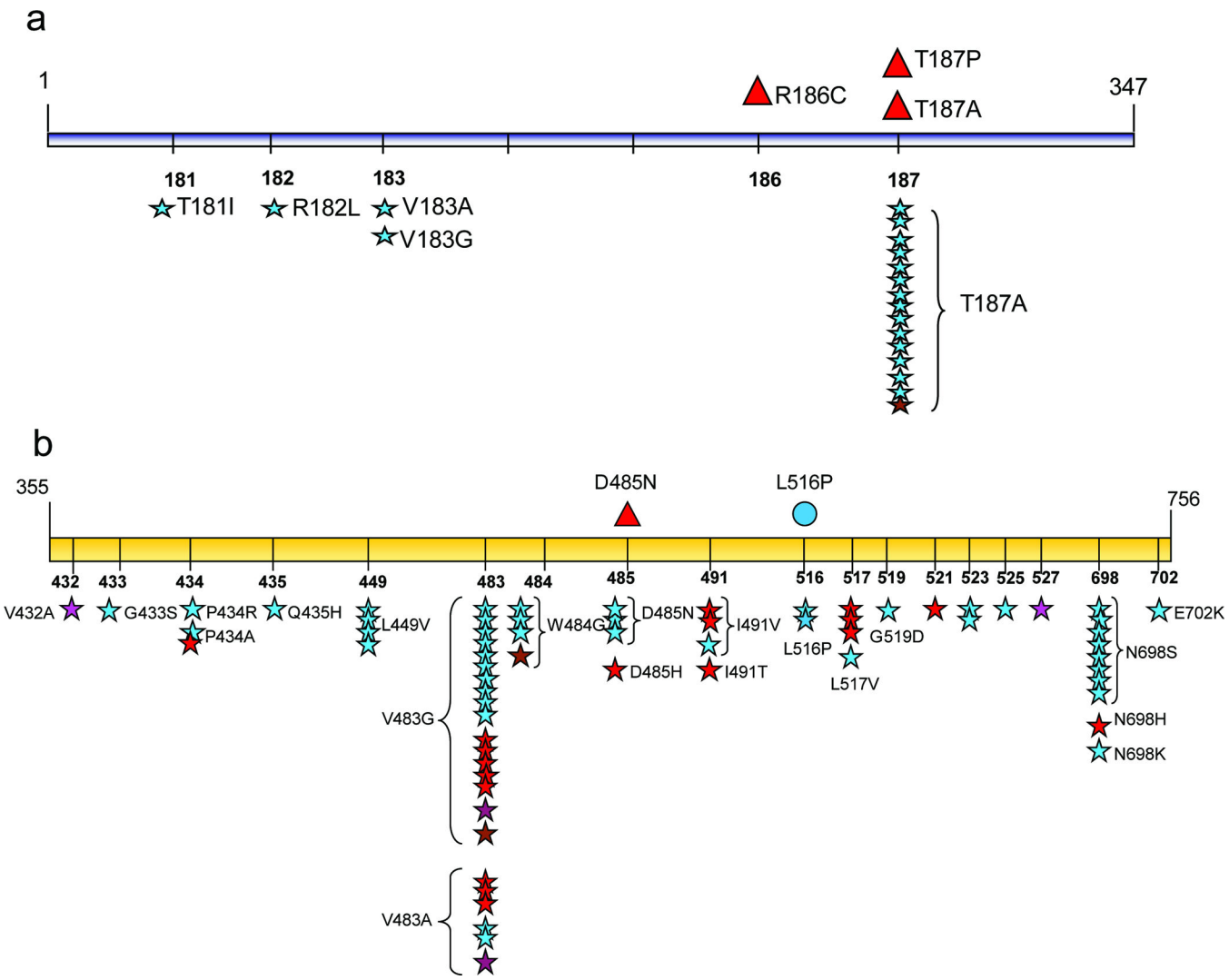

Figure 1. Putative compensatory mutations in a) *rpoA* and b) *rpoC* of *Mycobacterium tuberculosis*

Mutations identified after genome sequencing of experimentally evolved strains or paired clinical isolates are indicated above the gene by a circle and a triangle, respectively. Mutations identified by screening a global and a high-burden collection of MDR strains are indicated by stars below the gene. Colours indicate the respective strain lineage (blue - lineage 2, red - lineage 4, brown - lineage 5, pink - lineage 1). Some of these mutations occurred in multiple lineages and/or affect the same codon position.

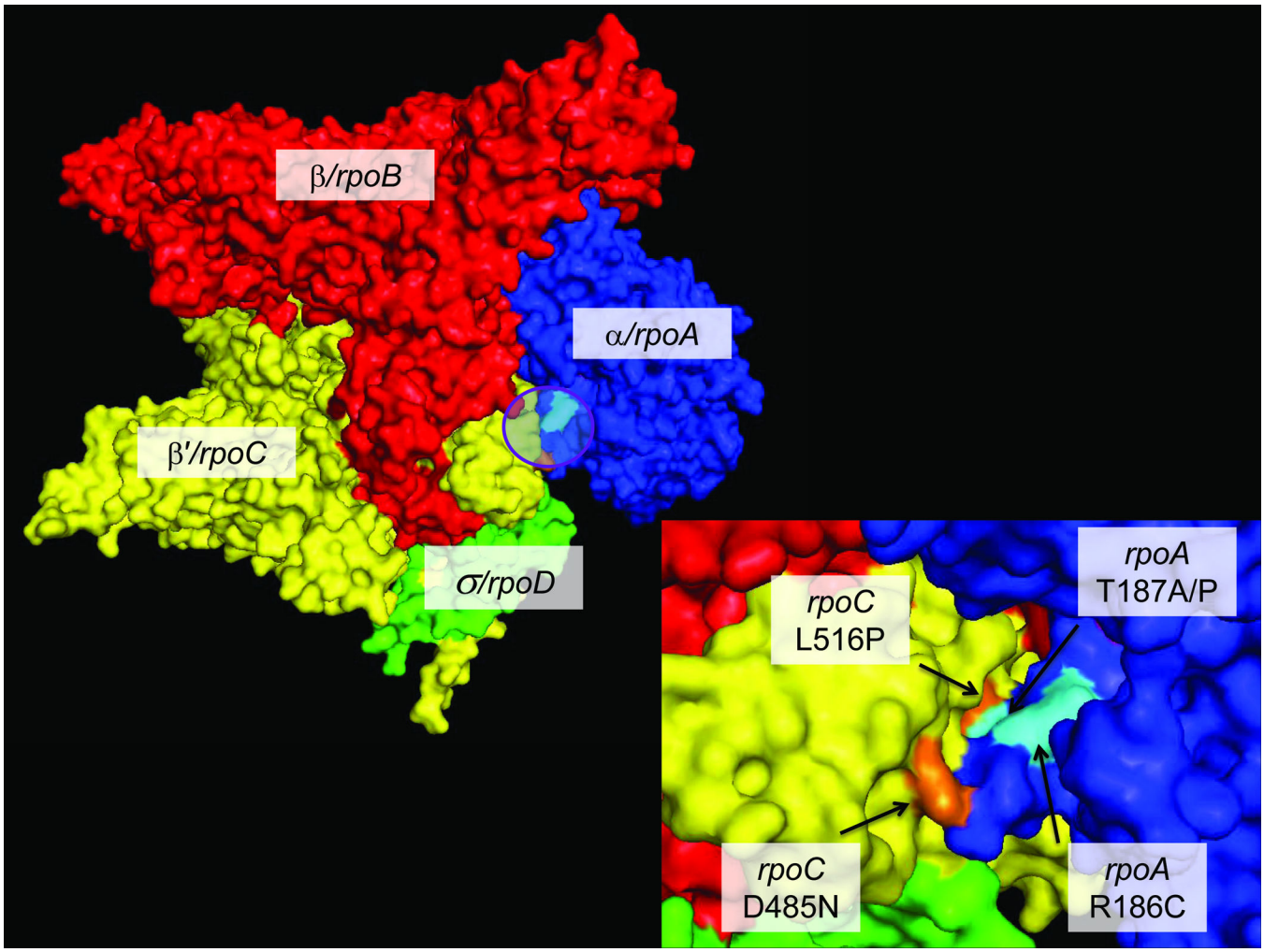

Figure 2. Putative compensatory mutations in *rpoA* and *rpoC* fall at the interface of RNA polymerase subunits

Mutations identified in rifampicin-resistant experimentally evolved isolates and paired clinical isolates were mapped onto the RNA polymerase of *Escherichia coli*. The mutations fall in predicted interacting residues of RpoA (light blue) and RpoC (orange) of the *E. coli* RNA polymerase. Residue numbers are indicated according to *M. tuberculosis* coordinates. Colour code: RpoA (α-subunit) - blue, RpoB (β-subunit) - red, RpoC (β'-subunit) - yellow, RpoD (σ-subunit) - green.

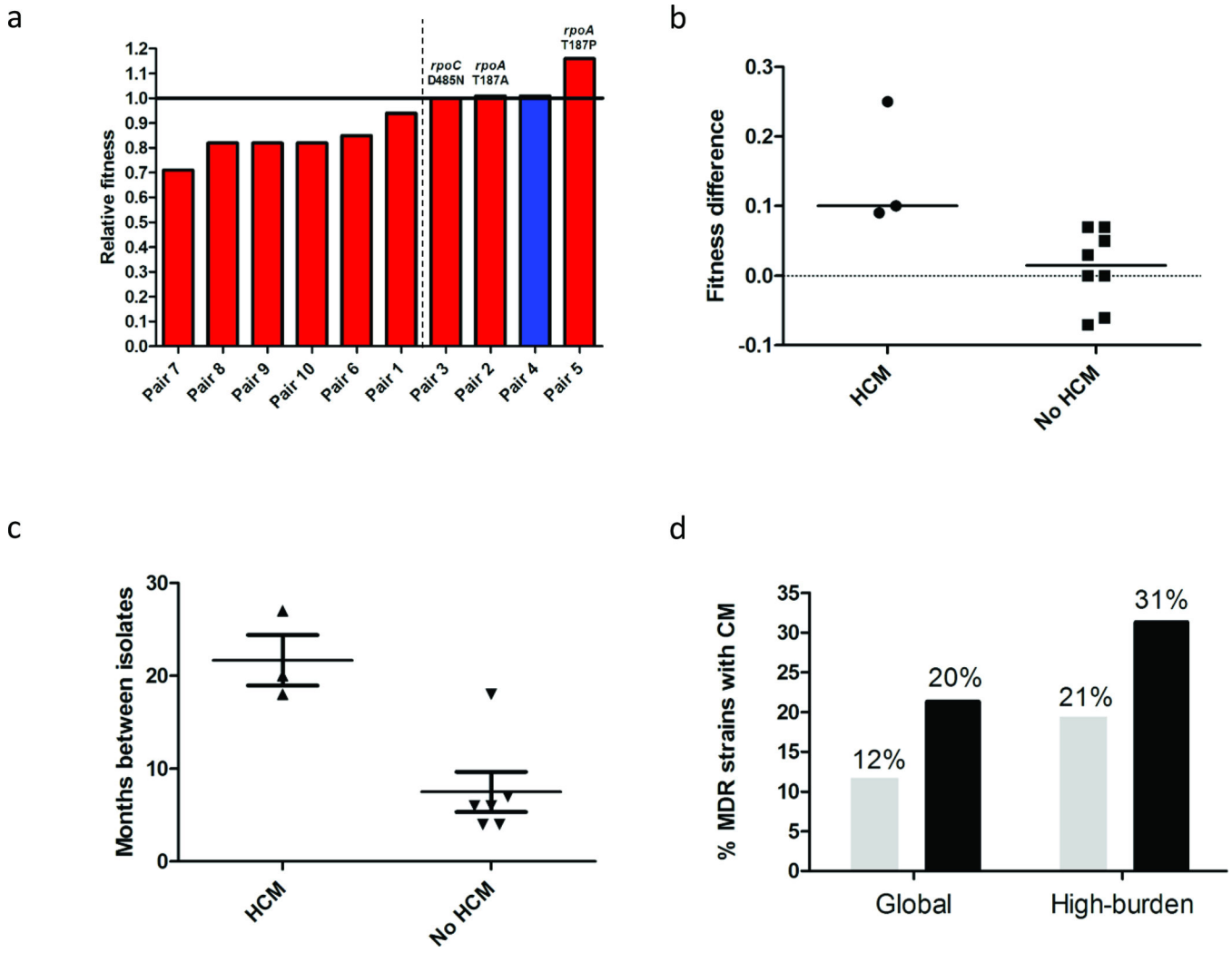

Figure 3. Experimental and clinical relevance of putative compensatory mutations

a) Experimental competitive fitness of 10 clinical isolates that acquired rifampicin resistance during treatment compared to their susceptible counterparts. High-probability compensatory mutations (HCM) are indicated in the pair in which they were identified. Bar colours indicate strain lineage (blue - lineage 2, red - lineage 4). **b**) Difference in relative fitness between 10 rifampicin-resistant paired clinical isolates compared to laboratory-generated mutants carrying the same rifampicin resistance-conferring mutation and genetic background as defined by strain lineage. Data are shown for clinical strains harbouring (or not) an HCM. **c**) Time in months between the first and the second isolate of each clinical pair. The horizontal line indicates the median time interval. **d**) Percentage of MDR strains carrying putative compensatory mutations in *rpoA* or *rpoC*. Grey bars refer to the percentage of strains carrying HCMs and black bars refer to strains carrying any putative compensatory mutation. Data for a global collection of strains and for the high MDR-TB burden regions of Abkhazia/Georgia, Uzbekistan and Kazakhstan are shown.

Table 1

Putative compensatory mutations in rifampicin-resistant *Mycobacterium tuberculosis*.

Gene	Amino acid change	Experimentally evolved strains	Clinical paired isolates	Global MDR strains	High Burden MDR strains	Total strains	High-probability compensatory mutation	SIFT score
rpoA	R182L	0	0	1	0	1	No	0.00
rpoA	R186C	0	1	0	0	0	No	0.00
rpoA	T187P	0	1	0	0	0	Yes	0.00
rpoA	T181I	0	0	0	1	1	No	0.00
rpoA	T187A	0	1	5	10	15	Yes	0.00
rpoA	V183A	0	0	1	0	1	No	0.00
rpoA	V183G	0	0	1	0	1	No	0.00
rpoC	A521D	0	0	1	0	1	No	0.00
rpoC	D485H	0	0	0	1	1	Yes	0.00
rpoC	D485N	0	1	0	3	3	Yes	0.00
rpoC	E702K	0	0	0	1	1	No	0.00
rpoC	G433S	0	0	0	1	1	No	0.00
rpoC	G519D	0	0	0	1	1	No	0.00
rpoC	H525N	0	0	0	1	1	No	0.00
rpoC	I491T	0	0	0	1	1	Yes	0.00
rpoC	I491V	0	0	0	3	3	Yes	0.00
rpoC	L449V	0	0	0	5	5	No	0.00
rpoC	L516P	1	0	0	2	2	No	0.00
rpoC	N698H	0	0	0	1	1	Yes	0.00
rpoC	N698K	0	0	0	1	1	Yes	0.00
rpoC	N698S	0	0	1	7	8	Yes	0.00
rpoC	P434A	0	0	0	2	2	No	0.00
rpoC	P434R	0	0	1	0	1	Yes	0.00
rpoC	Q435H	0	0	0	1	1	No	0.00
rpoC	Q523E	0	0	0	1	1	No	0.00
rpoC	Q523K	0	0	0	1	1	No	0.00
rpoC	G433S	0	0	1	0	1	No	0.00

Gene	Amino acid change	Experimentally evolved strains	Clinical paired isolates	Global MDR strains	High Burden MDR strains	Total strains	High-probability compensatory mutation	SIFT score
rpoC	V432A	0	0	1	0	1	No	0.00
rpoC	V483A	0	0	1	5	6	Yes	0.00
rpoC	V483G	0	0	6	11	17	Yes	0.00
rpoC	V517L	0	0	1	3	4	No	0.00
rpoC	L527V	0	0	1	0	1	No	0.00
rpoC	W484G	0	0	1	3	4	No	0.00