Tuberculosis vaccine strain *Mycobacterium bovis* BCG Russia is a natural *recA* mutant

Peter M Keller[1], Erik C Boettger[1,2] & Peter Sander[1,2*]

[1]*Institut für Medizinische Mikrobiologie, Universität Zürich, Gloriastrasse 30/32, CH-8006 Zürich, Switzerland.* [2]*Nationales Zentrum für Mykobakterien, Gloriastrasse 30, CH-8006 Zürich, Switzerland. Correspondence should be addressed to P. S. (psander@immv.uzh.ch).*

The current tuberculosis vaccine is a live vaccine derived from *Mycobacterium bovis* and attenuated by serial *in vitro* passaging. All vaccine substrains in use stem from one source, strain Bacille Calmette-Guérin. However, they differ in regions of genomic deletions, antigen expression levels, immunogenicity, and protective efficacy[1]. As a RecA phenotype increases genetic stability and may contribute restricting the ongoing evolution of the various BCG substrains, we aimed to inactivate *recA* by allelic replacement in BCG vaccine strains representing different phylogenetic lineages (Pasteur, Frappier, Denmark, Russia). Homologous gene replacement was successful in three out of four strains. However, only illegitimate recombination was observed in BCG substrain Russia. Sequence analyses of *recA* revealed that a single nucleotide insertion in the 5' part of *recA* led to a translational frameshift with an early stop codon making BCG Russia a natural *recA* mutant. At the protein level BCG Russia failed to express RecA. According to phylogenetic analyses BCG Russia is an ancient vaccine strain most closely related to the parental *M. bovis*[1]. Our data suggest that *recA* inactivation in BCG Russia occurred early and is in part responsible for its high degree of genomic stability, resulting in a substrain that has less genetic alterations than other vaccine substrains with respect to *M. bovis* AF2122/97 wild-type.

Tuberculosis is an infectious disease of enormous global importance. It is estimated that about one third of the human population is latently infected with *Mycobacterium tuberculosis* with 1.6 million people dying annually from the disease[2]. The currently employed tuberculosis vaccine, *Mycobacterium bovis* Bacille Calmette-Guérin (BCG) was originally derived from a virulent strain of *M. bovis* back to 1921, by repeated passages on potato slices soaked in glycerol-ox byle. The primary attenuation is attributed to loss of RD1 locus, which affects a protein secretion pathway[3-6]. Subsequent propagation of BCG in several laboratories around the world resulted in further *in vitro* evolution, which is still ongoing. Genetic alterations are mainly due to deletions and duplications[1], although single nucleotide polymorphisms (e.g. in *sigK*) have also been described[7]. The various genetic alterations – some presumably involving RecA-mediated recombination - affect the antigenic, protective, and metabolic properties of BCG. Thus, the term BCG does not refer to an entity but comprises a set of different substrains. Since the early sixties, freeze dried and lyophilized secondary seed lots are used as source of commercially available vaccine strains. WHO vaccine production guidelines[8] call for fresh cultures from secondary seed lots and propose not to exceed twelve passages from the primary seed lot. These recommendations have been modified recently, suggesting not to exceed four passages as phenotypic alterations were observed in vaccine batches undergoing as little as twelve subcultivations[9].

To establish BCG as a vector for stable heterologous antigen expression we intended to stabilize the genome of four BCG substrains representing different phylogenetic lineages by deletion of *recA*. A *recA* phenotype is desirable for live mycobacterial vaccines as homologous recombination is an important contributor of genomic evolution[10,11]. We constructed a replacement vector (**Suppl. Fig. 1** online) carrying an unmarked, inactive deletion of the *recA* allele. In this vector, a hygromycin resistance cassette and the counterselectable marker *sacB* flank *recA*. Transformation of this kind of vector allows generation of allelic replacement mutants

by two subsequent selection steps. In a first step, transformants that have integrated the knock-out plasmid into the bacterial chromosome are selected on hygromycin B containing media. Integration of the vector by homologous recombination, i. e. 3' or 5' of the genomic *recA*, is revealed by Southern blot analyses. In a second step, sucrose-selection is used to select for mutants that have undergone a second intramolecular recombination resulting either in inactivation of the *recA* gene or in reversion to wild-type.

Most of the transformants obtained in substrains Pasteur, Frappier, and Denmark resulted from homologous integration of the targeting vector (3'or 5') at the *recA* locus (**Suppl. Table 1** online). Subsequent counter-selection of clones arising from homologous recombination events readily resulted in *recA* deletion mutants as demonstrated by Southern blot analyses (**Suppl. Fig. 2** online). In striking contrast, all transformants (16/16 analyzed) of substrain Russia resulted from illegitimate recombination (**Suppl. Table 1** and **Suppl. Fig. 2** online).

The absence of marker integration by homologous recombination prompted us to examine the genomic *recA* sequence in BCG Russia in comparison to the annotated sequence of BCG Pasteur 1173P2[1]. We found a single insertional mutation 413 bp from the *recA* start (**Fig. 1**), which leads to a translational frameshift and a premature stop codon at amino acid position 140 in the major central RecA domain (residues 31-269). The truncated RecA has a length of 139 instead of 790 amino acids and lacks the complete C-terminal part of the protein including loops L1 (residues 156-165) and L2 (195-210) implicated in DNA binding[12]. BCG Frappier, Denmark, and Pasteur *recA* were identical with the annotated reference sequence encoding a functional RecA.

Bacteria respond to DNA-damage by coordinated expression of a multitude of genes involved in repair and control of cell division – the SOS response[13]. The SOS

response was induced by addition of the DNA-damaging agent mitomycin C (0.2 µg mL^{-1}) to a BCG culture, followed by incubation for an additional 24 hours[14]. Western blot analysis using a RecA antibody demonstrated induction of RecA in strain BCG Pasteur, but absence and failure of induction of RecA in BCG Russia (**Fig. 2**). A genetically engineered BCG Pasteur *recA* mutant (this study) was used as a negative control.

Replication errors, transpositions and recombination events contribute to genetic alterations and drive genome evolution. For BCG various differences in morphology, growth rate, protein expression and genetic make-up have been noted among commercially available substrains[15]. This is presumably a result of the numerous serial passages on natural (potato slices trenched in ox bile) and artificial media, which have led to the acquisition of genomic alterations and further attenuation. Of note, meta-analyses of BCG vaccination trials have indicated protective efficacies ranging from 0-80 %. A correlation between the number of serial *in vitro* passages and the decrease in protective efficacy has been observed[16]. Several reasons have been put forward to explain the varying protective efficacies of BCG, among others genetic differences between vaccine substrains as well as within an individual substrain[1]. More recently it has been suggested that the protective efficacy of ancient vaccine strains (**Fig. 3**) may be superior to that of the later ones that are more widely used[1].

Comparative genomics indicate that recombination events are a major driving force of bacterial evolution[11]. There is extensive evidence for large-scale rearrangements, duplications and deletions resulting from homologous recombination in *M. leprae*[17], *M. tuberculosis*[18] and *M. bovis* BCG[19]. Half of the proteins present in the tubercle bacillus originate from gene duplications. Tandem duplications are generally caused by unequal crossover between homologous sequences or by recombination of short DNA homologies. Homologous recombination between

similar sequences may invert or delete genes. Several deletions in the *M. tuberculosis* H37Rv genome resulted from recombination between adjacent repeats of IS6110 elements[18,20]. Sometimes, the molecular mechanisms underlying alterations at particular loci remain obscure and subsequent alterations may mask initial events, e.g. tandem duplication of the DU2 region in *M. bovis* BCG Pasteur arose from duplication of a 100 kb genomic segment that subsequently incurred an internal deletion of 64 kb[19]. Comparative genome and transcriptome analyses indicate that BCG Russia is an ancient BCG strain most closely related to the original strain attenuated by Calmette and Guérin[1]. However, a key piece of the puzzle was missing. What was the molecular mechanism underlying the high degree of genome conservation in BCG Russia? Here, analysis of the *recA* locus provides a possible clue. RecA, which is the key element of homologous recombination and a driving force in mycobacterial genome evolution, is disrupted by single nucleotide insertion in BCG Russia. This mutation results in a frameshift and premature translational stop and most probably contributed to the genome stability of this substrain of BCG.

Methods

Bacterial strains. *Escherichia coli* DH5α was obtained from Pharmacia (GE Healthcare, Uppsala, Sweden) and was used for plasmid propagation. *M. bovis* BCG Pasteur 1173P2 was obtained from the strain collection of the Institute Pasteur, Paris, France (strain #CIP105050). *M. bovis* BCG Denmark (1331) was obtained in the form of a vaccine production lot through the Statens Serum Institute, Copenhagen, Denmark. *M. bovis* BCG Frappier (Montreal, ATCC #35735) primary lot, dated 1973, lot number 1376141, was obtained from the American Type Culture Collection (ATCC), Rockville, MD, USA. The Frappier strain held by ATCC was initially transferred from the Institut Pasteur to Frappier's Institute in Montreal in 1937, was integrated in the Trudeau Mycobacterial Culture Collection and finally went to the ATCC. A vaccine production lot 547-1104 K. 1491 of M. *bovis* BCG Russia

(corresponding to ATCC #35740) was obtained through Medgamal Inc. (TD Allergen, Moscow, Russia). This was the first documented daughter strain distributed by Institut Pasteur in 1924[21] going directly to Russia. It is referred as an 'early strain' with regards to its genetic characteristics (regions of difference, insertional elements, antigen expression pattern).

The identity of the BCG substrains was confirmed using morphological features (roughling, generation time), PCR amplification and sequencing of described regions of difference as described by[22].

Media, transformation and DNA damage induction. The media for growing *E. coli* DH5α, and substrains of *Mycobacterium bovis* BCG have been described previously[23]. *M. bovis* BCG substrains were revitalised by addition of 1 mL Middlebrook 7H9 medium to a lyophilized batch. For DNA damage induction experiments *M. bovis* BCG was grown in Middlebrook 7H9 medium with addition of 10% (vol/vol) OADC in motionless tissue culture flasks in a 37 °C incubator. Published protocols were followed for the preparation of electrocompetent cells of mycobacteria and for electroporation[23]. The *recA* mutants of BCG Pasteur, Denmark and Frappier were generated by allelic replacement using a suicide knock-out vector (**Supplementary Methods** online).

To induce DNA damage, mitomycin C (0.2 μg ml^{-1}) was added to 20 mL of growing cultures (at an OD_{600} of 0.6) and incubated for 24 h. Thereafter, bacteria were harvested by centrifugation and prepared for cell lysis.

DNA isolation, sequencing and alignment. DNA was extracted from solid cultures grown on 7H10. The *recA* gene was amplified from genomic DNA (10 ng) by *Pfu* DNA polymerase using primers #1 and #2. The product was cleaned via gel purification and sequenced using primers #1 to #9 (**Suppl. Table 2** online) using an ABI PRISM 310 DNA sequencer (Applied Biosystems, Forster City, CA, USA). The sequences were aligned using Lasergene (DNASTAR, Madison, WI, USA) SeqMan

(version 6). The aligned sequence was compared via BLAST software to the *recA* sequence of *M. bovis* BCG substrain Pasteur 1173P2[1].

Preparation of cell-free extracts. Untreated and mitomycin-C-treated bacteria were harvested by centrifugation, washed with PBS buffer (pH 7.2) and resuspended in 300 µL PBS containing 0.5% protease inhibitor cocktail (Sigma). Bacteria were lysed in screw-cap tubes in a water immersion (ice-chilled) sonicator (Elma, Singen, Germany), using maximum duty-cycle during 1 h 20 min with a cooling interval every 20 min. The crude lysate was clarified by spinning at 1 000 x g during 10 min.

Western blotting. Cell-free extracts corresponding to 50 µg total protein (quantification via Bradford assay) were used for Western blot analyses. Following SDS-PAGE in a 13.3% polyacrylamide gel, proteins were blotted to a PVDF membrane using a semi-dry blotter (Biorad, Hercules, CA, USA) at constant current (0.3 A) for 30 min. Equal loading of proteins was confirmed by Coomassie staining, and by blotting with antibodies against the constitutively expressed mycobacterial protein (KatG).

1. Brosch, R. et al. Genome plasticity of BCG and impact on vaccine efficacy. *Proc. Natl. Acad. Sci. USA* **104**, 5596-601 (2007).

2. WHO. Fact sheet N°104: Tuberculosis. Vol. 2007 (ed. Center, W.M.) 1 (WHO, Geneva, 2007).

3. Brodin, P. et al. Dissection of ESAT-6 system 1 of Mycobacterium tuberculosis and impact on immunogenicity and virulence. *Infect. Immun.* **74**, 88-98 (2006).

4. Hsu, T. et al. The primary mechanism of attenuation of bacillus Calmette-Guerin is a loss of secreted lytic function required for invasion of lung interstitial tissue. *Proc. Natl. Acad. Sci. USA* **100**, 12420-5 (2003).

5. Lewis, K.N. et al. Deletion of RD1 from Mycobacterium tuberculosis mimics bacille Calmette-Guerin attenuation. *J. Infect. Dis.* **187**, 117-23 (2003).

6. Mahairas, G.G., Sabo, P.J., Hickey, M.J., Singh, D.C. & Stover, C.K. Molecular analysis of genetic differences between Mycobacterium bovis BCG and virulent M. bovis. *J. Bacteriol.* **178**, 1274-82 (1996).

7. Behr, M.A., Schroeder, B.G., Brinkman, J.N., Slayden, R.A. & Barry, C.E., 3rd. A point mutation in the mma3 gene is responsible for impaired methoxymycolic acid production in Mycobacterium bovis BCG strains obtained after 1927. *J. Bacteriol.* **182**, 3394-9 (2000).

8. WHO, Expert Committee on Biological Standardization. Requirement for dried BCG vaccine—amendment 1987. *World Health Organisation Technical Report Series* **771**, 35-36, 212 (1987).

9. Ho, M., Corbel, M., Knezevic, I. & Roumiantzeff, M. Report on a WHO consultation on the characterisation of BCG vaccines, WHO, Geneva, Switzerland 8–9 Dec 2004. *Vaccine* **23**, 5700–5704 (2005).

10. Sander, P. et al. A recA deletion mutant of Mycobacterium bovis BCG confers protection equivalent to that of wild-type BCG but shows increased genetic stability. *Vaccine* **21**, 4124-7 (2003).

11. Tamas, I. et al. 50 million years of genomic stasis in endosymbiotic bacteria. *Science* **296**, 2376-9 (2002).

12. Datta, S. et al. Crystal structures of Mycobacterium tuberculosis RecA and its complex with ADP-AlF(4): implications for decreased ATPase activity and molecular aggregation. *Nucleic. Acids Res.* **28**, 4964-73 (2000).

13. Cox, M.M. Regulation of bacterial RecA protein function. *Crit. Rev. Biochem. Mol. Biol.* **42**, 41-63 (2007).

14. Papavinasasundaram, K.G. et al. Slow induction of RecA by DNA damage in Mycobacterium tuberculosis. *Microbiology* **147**, 3271-9 (2001).

15. Behr, M.A. BCG - different strains, different vaccines? *Lancet Infect. Dis.* **2**, 86-92 (2002).

16. Colditz, G.A. et al. Efficacy of BCG vaccine in the prevention of tuberculosis. Meta-analysis of the published literature. *Jama* **271**, 698-702 (1994).

17. Cole, S.T. et al. Massive gene decay in the leprosy bacillus. *Nature* **409**, 1007-11 (2001).

18. Ho, T.B., Robertson, B.D., Taylor, G.M., Shaw, R.J. & Young, D.B. Comparison of Mycobacterium tuberculosis genomes reveals frequent deletions in a 20 kb variable region in clinical isolates. *Yeast* **17**, 272-82 (2000).

19. Brosch, R. et al. Comparative genomics uncovers large tandem chromosomal duplications in Mycobacterium bovis BCG Pasteur. *Yeast* **17**, 111-23 (2000).

20. Gordon, S.V. et al. Genomics of Mycobacterium bovis. *Tuberculosis (Edinb.)* **81**, 157-63 (2001).

21. Dubos, R.J., Pierce, C.H. & Schaefer, W.B. Differential characteristics in vitro and in vivo of several substrains of BCG. III. Multiplication and survival in vivo. *Am. Rev. Tuberc.* **74**, 683-98 (1956).

22. Mostowy, S., Tsolaki, A.G., Small, P.M. & Behr, M.A. The in vitro evolution of BCG vaccines. *Vaccine* **21**, 4270-4 (2003).

23. Frischkorn, K. et al. Investigation of mycobacterial recA function: protein introns in the RecA of pathogenic mycobacteria do not affect competency for homologous recombination. *Mol. Microbiol.* **29**, 1203-14 (1998).

Note: Supplementary information is available on the Nature Medicine website.

Acknowledgements We are grateful to Elaine O. Davis and to Steward T. Cole for providing RecA and KatG antibodies; to Anatoly A. Kotlobay for the free vial of BCG Russia, and to Brigitte Berger-Bächi and all members of the laboratory for helpful discussions. This work was supported in part by the University of Zurich and the Niedersächsischer Verein zur Bekämpfung der Tuberkulose Lungen- und Bronchialerkrankungen e. V. The *M. bovis* BCG Russia was a gift from Medgamal (TD Allergen, Moscow, Russia). The company had no influence in study design; collection, analysis, and interpretation of data; writing of the paper; and decision to submit it for publication.

Author contributions P.K. performed the experiments and wrote parts of the paper. P.S. and E.B. designed the study, analysed the results and were involved in writing of the manuscript. All authors commented on the manuscript.

Author information DNA sequence of *M. bovis* BCG substrain Russia *recA* has been deposited with GenBank data base under accession EU442641. Reprints and permissions information is available at npg.nature.com/reprintsandpermissions

Correspondence and requests for materials should be addressed to P.S. (e-mail: psander@immv.uzh.ch).

Figure 1 Comparative alignment of DNA sequencing electropherograms for *M. bovis* BCG Pasteur 1173P2 and BCG Russia displaying the insertional SNP (C_{414}) in the *recA* gene of BCG Russia. As a consequence Leu_{139} is changed to Ala, which is immediately followed by a premature stop codon (D140*). Nucleotide and amino acid numbering starts at the *recA* initiation codon.

Figure 2 Western blot analysis of cell-free protein extracts from different BCG preparations (with and without SOS response induction by mitomycin C; 50 µg protein per lane). Detection was carried out using a polyclonal mouse α-RecA antibody (1:1000), mature RecA protein has an approximate molecular weight of 38 kDa. A polyclonal mouse α-KatG antibody was used as loading control (1:2500); KatG protein has a molecular weight of 81 kDa. Bands migrating around 50 kDa are unspecific; secondary antibody was HRP-conjugated goat-anti-mouse antibody (1:2500).

Figure 3 Genealogy of BCG vaccine substrains, modified from[1] with permission, copyright (2007) National Academy of Sciences, U.S.A., displaying the original virulent ancestor strain *Mycobacterium bovis* (isolated by Nocard in 1908) and the subsequent series of genomic alteration including deletions of regions of difference (RD), single nucleotide polymorphisms, and duplications of genomic regions. The *recA* alteration (recA_D140*) is added to the scheme.

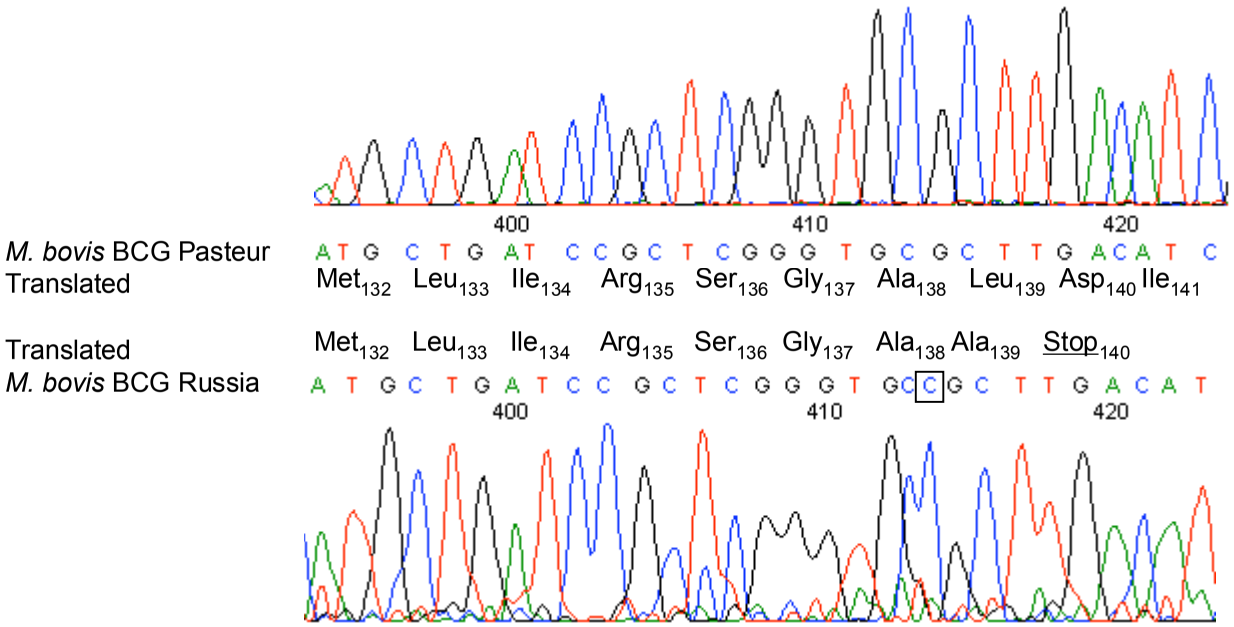
400
410
420
M. bovis BCG Pasteur
A T G C T G A T C C G C T C G G G T G C G C T T G A C A T C
Translated
Met132 Leu133 Ile134 Arg135 Ser136 Gly137 Ala138 Leu139 Asp140 Ile141
Translated
Met132 Leu133 Ile134 Arg135 Ser136 Gly137 Ala138 Ala139 Stop140
M. bovis BCG Russia
A T G C T G A T C C G C T C G G G T G C C G C T T G A C A T
400
410
420

Fig. 1

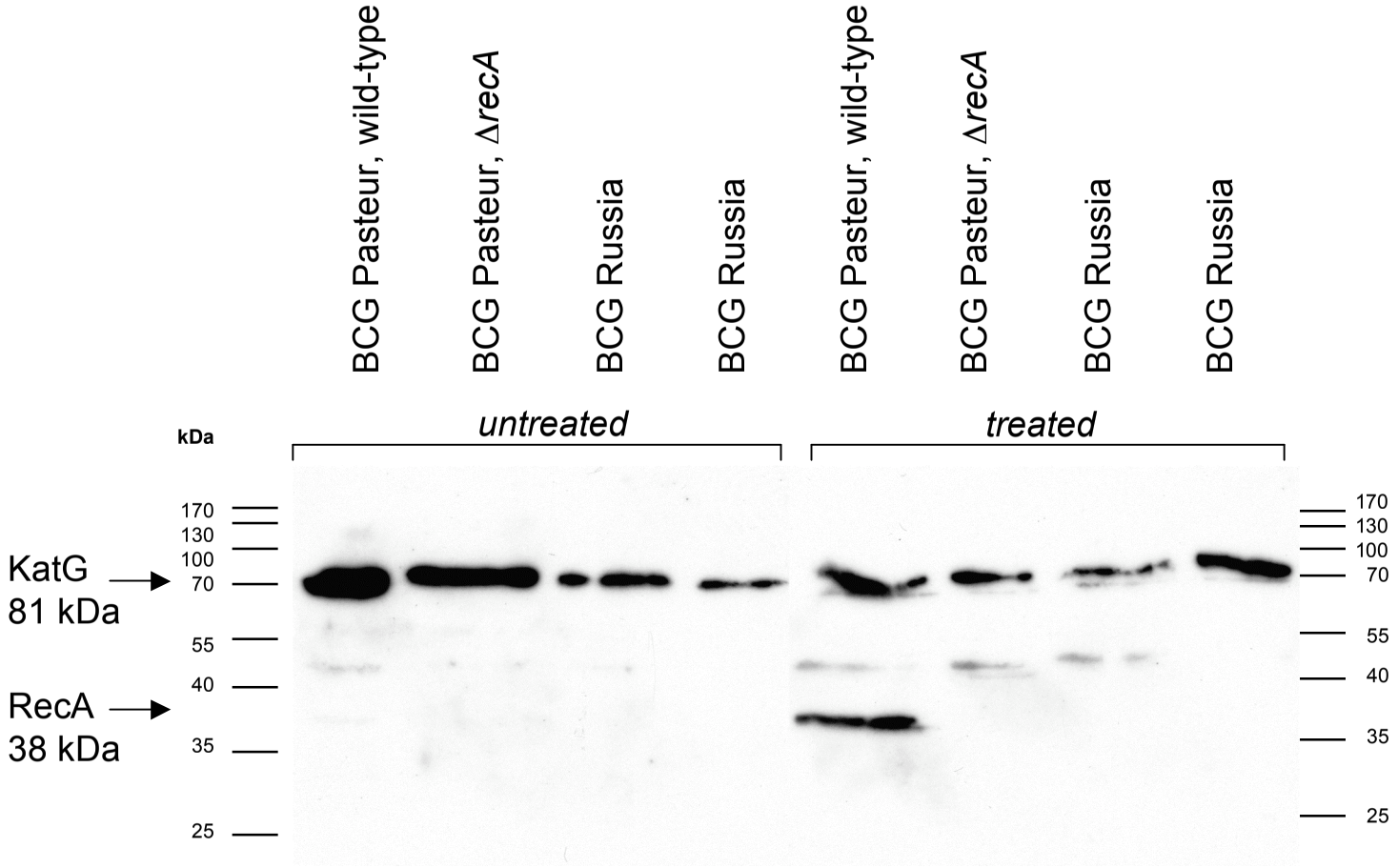
BCG Pasteur, wild-type
BCG Pasteur, ΔrecA
BCG Russia
BCG Russia
BCG Pasteur, wild-type
BCG Pasteur, ΔrecA
BCG Russia
BCG Russia
untreated
treated
kDa
170
130
100
70
55
40
35
25
KatG
81 kDa
RecA
38 kDa

Fig. 2

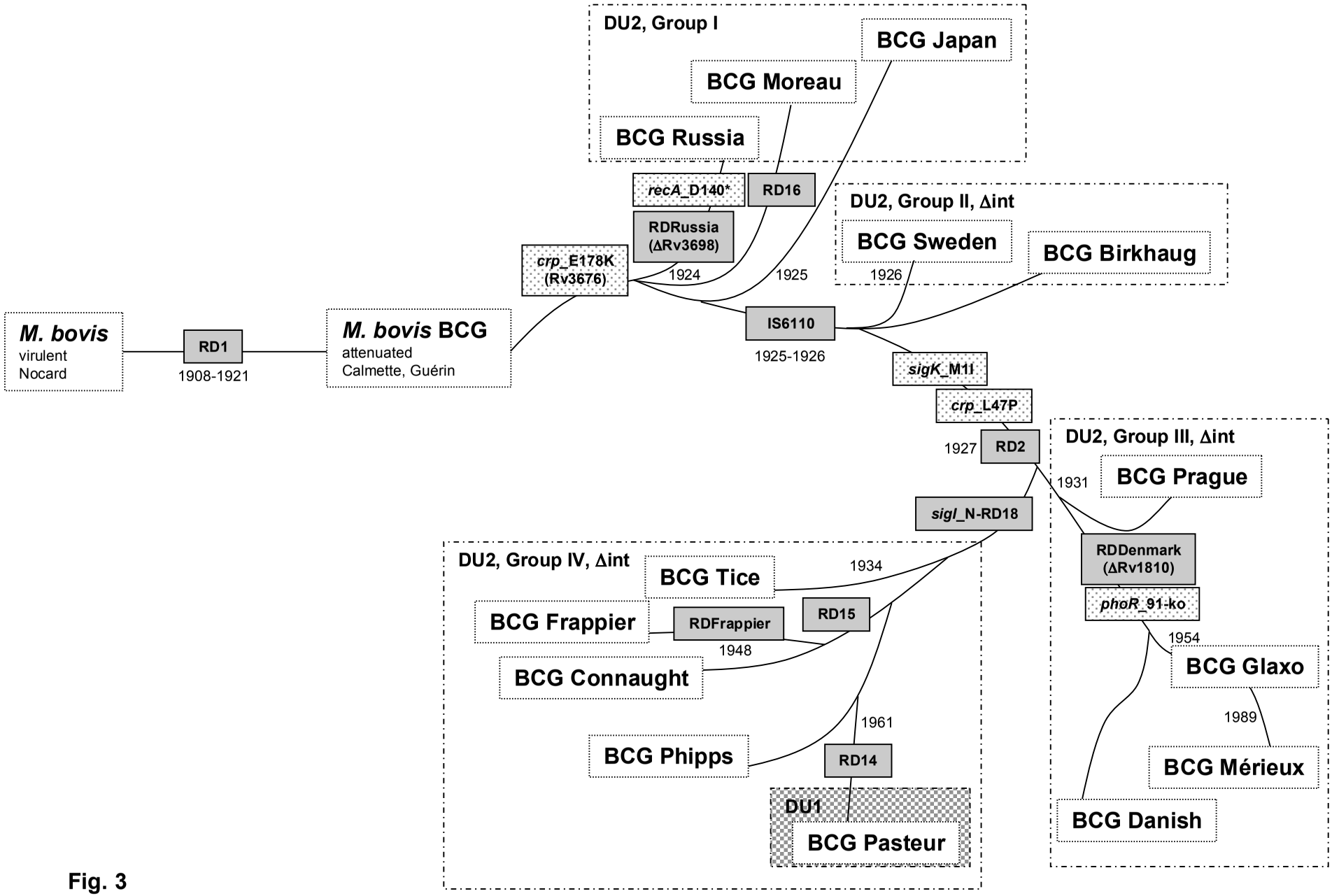
DU2, Group I
BCG Japan
BCG Moreau
BCG Russia
recA_D140*
RD16
RDRussia (ΔRv3698)
crp_E178K (Rv3676)
1924
1925
DU2, Group II, Δint
BCG Sweden
1926
BCG Birkhaug
M. bovis
virulent
Nocard
RD1
1908-1921
M. bovis BCG
attenuated
Calmette, Guérin
IS6110
1925-1926
sigK_M1I
crp_L47P
1927
RD2
DU2, Group III, Δint
1931
BCG Prague
sigI_N-RD18
RDDenmark (ΔRv1810)
phoR_91-ko
1954
BCG Glaxo
1989
BCG Mérieux
BCG Danish
DU2, Group IV, Δint
1934
BCG Tice
BCG Frappier
RDFrappier
RD15
1948
BCG Connaught
1961
BCG Phipps
RD14
DU1
BCG Pasteur

Fig. 3