Abstract: This invention refers to the mapping of insertions, deletions and point mutations of BCG Moreau from determining the genomic sequence of BCG Moreau and his comparison with the genome of other mycobacteriae. Moreover, this invention identifies nucleotide and peptide sequences, which provide methodologies for genetic verification of BCG Moreau, and for the distinction between BCG Moreau and other mycobacteriae, including in biological and clinical samples. The present invention supplies, furthermore, methods and kits for identification of such sequences, of the products of expression of these sequences or antibodies generated from this expression. The invention furthermore provides methodologies for monitoring and preparation of vaccines and kits for verifying the gene transcription/expression of certain antigens of the bacillus. From this invention are also provided secondary products with immunomodulatory activity, for example, in the treatment of bladder cancer and asthma.
"METHOD, KIT AND INITIATORS FOR GENETIC IDENTIFICATION OF MYCOBACTERIUM BOVIS BCG VARIETY MOREAU-RJ, METHOD FOR MONITORING FEASIBILITY OF BCG MOREAU-RJ AND EXPRESSION OF GENES, AND PRODUCTS WITH IMMUNOLOGICAL ACTIVITY"

FIELD OF INVENTION

This invention refers to the mapping of insertions, deletions and point mutations of BCG Moreau from the determination of the genome sequence of BCG Moreau and its comparison to the genome of other mycobacteria. Moreover, this invention identifies nucleotidic and peptidic sequences, which provide methodologies for genetic verification of BCG Moreau, and for distinction of BCG Moreau of other mycobacteria, including in biological and clinical samples. This invention further supplies methods and kits for the identification of such sequences, of products from the expression of these sequences or of antibodies generated from this expression. The invention still does not provide methodologies for monitoring and preparation of vaccines and kits for verification of the gene transcription/expression of certain antigens of the bacillus. From this invention are provided secondary products with immunodulatory activity, for example, in the treatment of bladder cancer and asthma.

BACKGROUND OF THE INVENTION

Tuberculosis (TB) is responsible annually for the death of approximately 2 million persons in the whole world and more than 8 million new cases of the disease are recorded each year. The large majority of deaths (99%) and of new cases recorded (95%) occur in countries of middle or
low income, making clear the narrow relationship between the larger incidence of TB and precarious socio-economic condition. Mycobacterium tuberculosis is the main agent that causes tuberculosis in human beings. However, Mycobacterium bovis, an etiological agent of bovine TB, is also capable of infecting humans, provoking a clinical condition of TB that is clinically indistinguishable from TB caused by M. tuberculosis. The vaccine used worldwide against tuberculosis is BCG (Calmette-Guerin bacillus), an attenuated variety of Mycobacterium bovis obtained in the early XX century after 230 passages (subcultures) in the glycerol potato-bile medium by A. Calmette and C. Guerin, in France. Between the years 1924 and 1926, at least 34 countries received BCG cultures from the Pasteur Institute. Since then, and until the establishment of the lyophilized seed batches, BCG has been sub-cultivated and independently propagated in the different countries. In 1947, an international meeting was held to discuss the future of the BCG vaccine, resulting in the comparison of the more than 50 existing sub-varieties of BCG through extensive studies in animals and a survey of all clinical data available until that time. Among the different currently used BCG varieties are Connaught, Danish, Glaxo (Merieux), Moreau (Rio de Janeiro), Pasteur and Tokyo. Although BCG is the vaccine most widely used in the world to date, several doubts were not elucidates, such as, for example, the reasons why:

the effectiveness of protection varies (from 0 to 80%, depending on the clinical test).
the different varieties of BCG present variable degrees of virulence and protection when assessed in animal models and also in humans; the action mechanism(s) involved in protection are little known; the conversion capacity of the tuberculin test varies among the different varieties of BCG and pathways of immunization; although extraordinarily safe for a live vaccine, the increase in the prevalence of the HIV infection in countries where TB is endemic raises important questions in connection with the safety of the vaccine; the use of BCG as an oral vaccine is, generally speaking, restricted and its effectiveness varies. However, BCG Moreau RDJ has proved quite efficient in this respect; BCG (or fractions thereof) can be used as an adjuvant in certain immunization situations, but this application needs to be studied in greater depth; and BCG (or fractions thereof) can be used as an immunodulator in illnesses such as bladder cancer and asthma, although the action mechanism is not totally clear yet.

Recent studies using molecular and genomic techniques, associated to an extensive historic survey, were capable of establishing a "genealogy" of the different BCG varieties. The set of these data demonstrates that, when compared to the complete M. tuberculosis genome, circa 129 genes (ORFs) are absent from the genome of various BCG varieties (deleted regions). Some deletions are common to all the
BCGs analyzed, such as a region referred to as RD1. Others seem to be specific to a certain variety, such as the RD16, which is absent from the BCG Moreau variety. Although valuable, these studies are not capable of identifying subtler differences among the different varieties, such as point mutations or small deletions/insertions, accessible only through full sequencing of the genome (an example of important point mutations is evidenced by the loss of the synthesis capacity of methoxymethyl acids in BCG varieties derived since 1927, which loss results from the mutation of a single nucleotide; however, this mutation may be associated to a decrease in virulence). One could think that the variation in the protective capacity of BCG, observed in the different clinical tests performed, must be solely due to the BCG used; however, this is not the only explanation since contradictory results were obtained when the same variety was used in different countries. Other factors such as the form of preparation of the vaccine, virulence, form of distribution (e.g., lyophilized), the feasibility of bacillae, genetic characteristics of the target-population and standard of previous infection by common mycobacteriae of the environment, among others, certainly perform an important role in this scenario.

As shown in Table 1, BCG Moreau can be considered to be an "old" variety, closer to the original BCG obtained by Calmette and Guerin, presenting a single deleted region (RD16), in addition to the common RD1 region deleted from all the BCG varieties compared to the sequence of M. tuberculosis. On the other hand, the BCG Moreau variety
does not present several deletions common to the most "recent" BCG's.

<table>
<thead>
<tr>
<th>BCG Variety</th>
<th>Deleted Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pasteur</td>
<td>RD1, RD2, RD14</td>
</tr>
<tr>
<td>Phipps (Philadelphia)</td>
<td>RD1, RD2</td>
</tr>
<tr>
<td>Frappier (Montreal)</td>
<td>RD1, RD2, RD8, RDFrappier</td>
</tr>
<tr>
<td>Connaught (Toronto)</td>
<td>RD1, RD2, RD8</td>
</tr>
<tr>
<td>Tice (Chicago)</td>
<td>RD1, RD2</td>
</tr>
<tr>
<td>Denmark (Danish 1331)</td>
<td>RD1, RD2, RDDenmark</td>
</tr>
<tr>
<td>Birkhaug</td>
<td>RD1</td>
</tr>
<tr>
<td>Swedwen (Gothenburg)</td>
<td>RD1</td>
</tr>
<tr>
<td>Glaxo</td>
<td>RD1, RD2, RDDenmark/Glaxo</td>
</tr>
<tr>
<td>Prague</td>
<td>RD1, RD2</td>
</tr>
<tr>
<td>Japan</td>
<td>RD1</td>
</tr>
<tr>
<td>Russia</td>
<td>RD1, RDRussia</td>
</tr>
<tr>
<td>Moreau (Brazil)</td>
<td>RD1, RD16</td>
</tr>
</tbody>
</table>

RD = Deleted Region

Thus, we can observe that the known varieties of BCG still need to grow in effectiveness, to produce less side effects, and an expansion of quality control in production of the vaccine.

SUMMARY OF INVENTION

This invention, in its most general aspect, comprises mapping of insertions, deletions and point mutations of BCG Moreau based on the determination of the genomic sequence of BCG Moreau RDJ and its comparison to the genome of other bacteria.

Moreover, this invention provides methodologies for genetic verification of BCG Moreau, as well as for
traceability, verification of feasibility of the BCG bacillus in vaccine preparations and kits for verification of the gene transcription(expression of certain antigens of the bacillus during the growth process and in vaccine preparations.

Based on the methods of this invention, it is possible to estimate the capacity of discriminating between one bacteria containing the deletions described and other mycobacteria, which do not have these deletions. Thus, one estimates the capacity of discriminating, through the analysis of clinical samples or of tests in patients, among individuals vaccinated with BCG Moreau and individuals infected with M. tuberculosis, M. bovis, or another mycobacteria of the M. tuberculosis complex. It is also possible, from knowing precisely the genomic sequence of BCG Moreau RDJ, the construction of bacillae genetically modified through the introduction of exogenous DNA fragments in the genome, or through the introduction of sequences in extrachromosomal elements, complementation of absent genes in BCG Moreau RDJ etc.

An additional objective of the invention refers to kits for monitoring and traceability of BCG Moreau.

**BRIEF DESCRIPTION OF FIGURES**

Figure 1 is a scheme demonstrating the arrangement of DNA of the regions INDEL No. 22 and 23.

Figure 2 shows an analysis in agarose gel, whose purpose is to demonstrate the genetic identification capacity.
Figure 3 shows another analysis in agarose gel whose purpose is to demonstrate genetic identification capacity.

Figure 4 is a scheme demonstrating the deletion of the INDEL No. 7 region and the position of the oligonucleotides for amplification.

**DETAILED DESCRIPTION OF THE INVENTION**

Having in view the precarious knowledge of the genetic and genomic characteristics of BCG Moreau RDJ, (e.g. mapping of mutations such as INDELs: insertions and deletions in the genome, and point mutations, responsible for the absence of capacity or modification of the capacity of codifying or expressing correctly proteins in these regions), the full sequencing and characterization of the genome of the BCG Moreau RDJ vaccine variety becomes essential.

The studies in this invention were started from the sample of the seed inventory for production of the vaccine BCG Moreau RDJ of Fundação Ataulfo de Paiva (FAP). Simultaneously, detailed knowledge of the genome permits the development of fast, modern and extremely accurate methodologies for the identification and control of the vaccine's quality, as well as for monitoring and improvement of the production evolution stages. Such methodologies involve the development of molecular and proteomic evaluation kits (PCR, micro- and macroarray).

From this invention, it is possible to build the recombinant multivalent BCG vaccine with an increased degree of protection.
In view of this, this invention permits improvement of the effectiveness, decrease of side effects and greater quality control in the production of the vaccine used in fighting the disease.

The effectiveness of the vaccine may be increased by reintroducing important antigens produced by genes that disappeared from the BCG Moreau-RDJ genome after years of cultivation and passage. With the mapping of the BCG Moreau genome, it is possible to improve the vaccine through electrical engineering, "linking" or "deleting" genes according to their function.

Moreover, this invention offers genome evaluation mechanisms of the BCG produced, which will permit to ensure the genetic stability and quality of different production and storage conditions, so that one is sure of what is being produced. Moreover, from this invention it is possible to increase the quality of the vaccine by consciously modifying manufacturing conditions.

From the invention it is possible to develop vaccines recombinant with BCG (Bacillus Calmette-Guerin). This type of vaccine is obtained by inserting in the BCG the genes that codified antigens of illnesses in general. Thus, by genetically adding to the BCG segments of DNA responsible for the synthesis of antigens of diseases such as tetanus and chickenpox, one increases the protection power of the vaccine, which may grant immunity against several illnesses.

This invention will be described based on the Examples below, which must not be understood as limiting the same.
EXAMPLE 1

In this Example, it will be described how the genome library was built. In this phase, the DNA of the bacillus will be divided into smaller pieces, which will be inserted in plasmids (DNA molecules) of bacteria and cloned. After being especially prepared, the fragments will be sequenced and analyzed in computers, which will assess the quality and will fit the sequences in the correct order, until they form the complete genome.

The BCG Moreau-RJ will be compared with other varieties, according to Table 2, to identify variations in the code and expression of the genes, which will result in detailed knowledge of the evolution and characteristics of virulence and survival.

The stages for identification of the varieties in the code and expression of the genes are as presented below.

Construction of genome libraries of *M. bovis* BCG Moreau-RJ

Two types of genome libraries of BCG Moreau-RJ were built, both from the genome DNA extracted from a culture of BCG Moreau-RDJ supplied by Fundagao Ataulpho de Paiva. Both libraries were built in the plasmid vector pBluescript from partially digested DNA genome, with the restriction enzyme BspRI (produced by LGFB, DBBM-FIOCRUZ). For this, the premise of a practically random digestion was verified by analyzing the GC content estimated for BCG Moreau-RDJ, and an *in-silico* estimate of the distribution of expected sites for Bsp RI. The progress of a partial digestion was subsequently verified by electrophoresis and visualization in agarose gel. The first library, of routine use in
genomic sequencing, contains DNA inserts in the range of 2kb. The second, for more specific purposes, contains DNA inserts in the range of 5 to 10kb.

**Stages:**

1) Extraction of genomic DNA from *M. bovis* BCG Moreau-RDJ  
2) Partial digestion with BspRI  
3) Fractioning of genomic DNA in 0.8% agarose gel and isolation of two populations (1, 5-3 kb and 5-10 kb)  
4) Preparation of cloning vector (extraction of DNA, digestion with EcoRV, dephosphorylation, evaluation of quality of vector)  
5) Linking and transformation of *Escherichia coli* DH5αF’lacIq  
6) Evaluation of the quality of libraries

**Resulting products**

Two highly representative genome were obtained (more than 50,000 independent clones per library).

**EXAMPLE 2**

**Transformation of genome libraries and preparation of plasmid DNA**

The genome libraries (BCG: smaller inserts of approximately 2kb and BCG-L with inserts larger than 5 to 10kb) were transformed weekly into *Escherichia coli*, plated in selective medium and isolated colonies peaked in 1 ml of half a CG (*deep well plates of 96 wells*). After growth of the cultures for 20 hours, the plasmid DNA was extracted through a alkaline use protocol with clarification of the Used material in MAGV-Millipore plates of 96 wells. The plasmid DNA was then precipitated with isopropanol and
stored in plates of 96 wells for sequencing. All the DNA plates are stored at -20°C. Each DNA (clone) has individual numbering, which identifies it in relation to the type of library, to the plate and to the well.

For example:

**BCG-B120-C3** (library with minor inserts, plate number 120, well C3)

Circa 215 plates of culture from 96 wells were prepared, resulting in 215 plates of plasmid DNA from 96 wells, adding to a total of 20,640 independent clones.

**EXAMPLE 3**

**Sequencing of BCG Moreau-RJ genome**

Afterwards, the sequencing of the BCG Moreau-RJ genome was sequenced.

The plasmid DNAs were obtained in Example 2 above, sequenced in two directions with initiators that hybrid in the region of the vector (F and R: initiators M13 forward and reverse). In short, two replicas of each plate of the inventory of plasmid DNA were made in sequencing plates from 96 wells. To these plates were added the reagents of the Big Dye 3.1 (ABI) sequencing kits and the initiators, respectively F or R. The sequencing reactions were submitted to cycling for the extension of the chain in the thermocycler of 96 wells (ABI), the reactions were precipitated with isopropanol and re-suspended in a racing plug. The sequencing reactions thus prepared were read in a 48 capillaries (ABI3730) automatic sequencer from the DNA Sequencing Platform RPT-01A PDTIS-FIOCRUZ. Each sequencing plate takes 4 hours to be processed. The data obtained were
captured automatically by the server of the Biocomputing PlatformRPT-04A PDTIS-FIOCRUZ using a script especially developed for this purpose.

Circa 50,000 other "reads" (20,640 templates - 41,280 individual sequences + circa 20% re-sequences) were obtained.

The DNA sequences thus obtained above were analyzed in connection with their quality, removing the regions corresponding to the vector sequence and maintaining the sequenced insert regions with quality higher than 20 (parameters used by the program) through programs Phred and Phrap. These data were then analyzed by the program Consed, which does the genome sequence assembly from fragments obtained from the sequence, joining them in contigs when there regions for superposition.

Thus, 37,555,553 bp were included in the assembly, of which 23,429,107 of high quality (Q>20), resulting in 5.4 times the genome's coverage in high quality stretches. The general coverage was 8.5 times, and after assembly 76 resulting contigs were obtained.

EXAMPLE 4

Assembly, verification and junction of contigs

In the initial phase of random genome sequencing are obtained almost only single sequences. As the number of sequences increases, DNA sequences that joint isolated fragments joining the contigs start to appear. At this point, the number of contigs begins to rise until it reaches a plateau. From this stage on, the union of separate contigs starts to occur and their number to
decrease. Ideally, we would reach a single contig, corresponding to the single, circular, chromosome of the bacteria. However, this stage is not reached by random sequencing (shotgun) for the following reasons:

1) Occurrence of genome regions with low representativeness in genome libraries, which do not go as far as being sequenced with this approach (inherent to the random approach, where the distribution of the frequency of specific fragments in the library corresponds to a Gauss curve).

2) Existence of genome regions with high G+C content (characteristic of the mycobacteriae in general), which form secondary structures in the form of a staple, leading to disconnection of the DNA polymerase of the mold DNA tape, preventing the sequencing reaction.

Then begins the finalization of the sequencing aimed at the junction of all contigs (gap closure) and verification of regions with doubts, where a directed approach is used to resolve these specific problems. Three distinct approaches were used for finalization of the BCG Moreau-RJ genome:

a) Primer walking

In this approach initiators are used (oligonucleotides; primers) for sequencing of the internal region of a specific clone, the initiators are drawn from sequence data already obtained. Each initiator is single in the genome region. Circa 230 specific initiators were drawn and used at this stage.

b) PCR for gap closure
In this approach, two oligonucleotides are drawn to hybrid in the flanking regions of a gap. These oligonucleotides are then used to amplify this region from the genomic DNA of the BCG Moreau-RDJ. The DNA fragment obtained in this PCR reaction is purified and submitted to sequencing using the same oligonucleotides as initiators of the sequencing reactions. Eighty-four oligonucleotides were drawn and used at this stage.

c) Modifications in the sequencing protocol for resolution of secondary structure regions

PCR for resolution of the secondary structure regions: This approach is essentially identical to the one described in item b, above. However, in this case, the oligonucleotides drawn are longer, permitting one to perform a PCR reaction (and sequencing) at a higher temperature, which is to the detriment of the formation of secondary structures in the DNA. At this stage, 48 nucleotides were drawn and used.

EXAMPLE 5

Comparative analysis of the genomic sequence, and structural and functional annotation

The detailed annotation of BCG Moreau-RJ genome involves analysis of the ORFs of the genome and its regulatory sequences, analysis of the metabolic pathways, comparison with genomes from other mycobacteriae (M. bovis, M. tb, BCG Pasteur etc), and prediction of the impact of deletions/insertions and specific mutations for BCG Moreau-RJ on gene expression, virulence and vaccine protection of same.
At the end of the experiments reported, the inventors identified 23 locations in the genome where there are differences of insertion (1 case) and deletion (22 cases) when compared to the genomic sequence of the BCG Moreau RDJ with the genomic sequence of M. bovis. For each of these regions, the potential impact of the INDEL on the encoding capacity of this region, and on the region's regulatory capacity. ORFS (Regions in Open Reading Phase) were identified deleted in BCG Moreau RDJ or whose regulation of the expression was affected by the mutations - INDELS.

The scope of protection of this invention, then, are sequences of nucleotides deleted in the BCG Moreau RDJ genome, and present in the genome of M. tuberculosis or M. bovis, or inserted in BCG Moreau RDJ and absent from the genome of M. tuberculosis or M. bovis.

Nucleotidic sequence or sequences of nucleotide in this invention means double strand DNA, single strand DNA, or products of transcription of these DNAs.

The proteins codified by the regions deleted may permit the identification of individuals vaccinated with BCG Moreau, of patients infected with M. tuberculosis, M. bovis, or other pathogenic mycobacteriae. This identification may occur through analysis of the clinical sample, such as blood, saliva, secretions, serum, biopsy, aspirate, lymph, aspirate or bronchial wash, pleural liquid or other material, or even through the skin test or other directly applied to patients.

Typically, such test would consist of the following steps:
1) Isolation of DNA or RNA from biological material;
2) Detection of specific DNA sequences of mycobacterium present in the sample, or detection of mycobacterium RNA, after cDNA production;
3) Analysis of these sequences through sequencing, or through hybridization with specific oligonucleotides or with DNA or RNA fragments prepared for this purpose.

The analysis of these sequences can be made through electrophoresis in agarose or polycrylamide gel, revealing visually, after coloration, DNA bands of the size contemplated from analysis of the organism's genome (BCG Moreau RDJ or M. tuberculosis or other mycobacteria).

The analysis can be made also through hybridization with a probe marked non-radioactively or radioactively. The probe, of the complementary sequence to the DNA or RNA to be detected can consist of one oligonucleotide, or of a fragment of DNA or RNA. To increase the test's sensitivity, it may include the use of amplification techniques, such as PCR reaction or real time PCR.

With respect now to Figure 1, the same refers to a scheme demonstrating the DNA arrangement of the regions INDEL No. 22 and 23. The arrows indicate the positions of the amplified oligonucleotides. The numeric coordinates refer to the sequence of M. bovis. More specifically, Figure 1 demonstrates approaches to discriminate for example, M bovis from M. tuberculosis and from Moreau. If we presume the numbering of the oligonucleotides in Figure 1, the numeric coordinates 1, 2, 3, 4 and 5 (from left to right), we have:
(i) 1 and 2 amplify the product of M. bovis and M. tuberculosis, but do not amplify Moreau;

(ii) 1 and 3 amplify BCG, M. bovis and M. tuberculosis, but will result in quite a smaller fragment in the case of Moreau.

(iii) 4 and 5 amplify fragment in Moreau and in M. tuberculosis, but not in M. bovis.

It must be stressed that the position of complementariness of the oligonucleotides is indicated in a limited way in Figure 1, but there is a margin to move them along the sequence.

As examples were listed sets of primers flanking the junction regions mentioned in Table 2. These primers will amplify, in PCR reaction, a fragment of the size indicated in the table when BCG Moreau RDJ is present in the sample, and a larger fragment when M. tuberculosis or M. bovis is present. Alternatively, primers may be used, which would amplify a specific fragment when M. tb or M. bovis is present, but no product when BCG Moreau RDJ is present, since, in this case, primers are chosen that hybrid in the region that was deleted in BCG Moreau RDJ.

Table 3 shows the sequence of oligonucleotides of this invention.

With reference to Figure 2, 1% agarose gel colored with etide bromide containing the amplification products of the genomic DNA of M. tuberculosis H37Rv (channel 2), BCG Moreau (channel 3) and BCG Pasteur (channel 4) with the initiators RD237 and RD238. Channel 1 contains a molecular weight marker (molecular sizes indicated to the left of
Figure 2). This analysis in agarose is a typical PCR experiment to demonstrate the effectiveness of the kit of this invention for genetic identification. In the experiment shown in Figure 2, initiators 237 and 238 were used (according to Table 3), corresponding to SEQ ID NO. 47 and SEQ ID NO. 48, which amplify junction region 16.

Figure 3 shows another 1% agarose gel colored with etide bromide containing the amplification products of the genomic DNA of M. tuberculosis H37Rv (channels 2, 5, 8 and 11), BCG Moreau (channels 3, 6, 9 and 12) and BCG Pasteur (channels 4, 7, 10 and 13) with the initiators RD 251/253, RD 251/252, RD 254/253 and RD 254/252, as indicated in Figure 3. Channel 1 contains a molecular weight marker (molecular sizes are indicated to the left of Figure 3).

Figure 4 is a scheme demonstrating deletion of the region INDEL No. 7 and the position of the oligonucleotides for amplification, used in the experiment of Figure 3. The shaded bar represents the deleted region in BCG Moreau RDJ in relation to M. bovis.

Thus, from this invention, it is possible to use the kits developed for traceability and verification of the genetic identity of BCG Moreau RDJ, for example, as an essential step in quality control of the BCG Moreau RDJ vaccine, verifying the identity of the inoculum for growth in culture, verification of seed batches, verification of samples of diverse origins and of the infectious or vaccine strain in isolated clinical samples of patients. This can be of great importance when, for example, a patient suspects that he is suffering the side effects of
vaccination with BCG Moreau RDJ, but where there is equally the possibility of infection with a pathogenic mycobacterium. In this case, the analysis of a clinical sample, using methodologies as those described above, can indicate the presence or not of BCG Moreau RDJ in the sample.

The secondary products of this invention, which have an immunodulatory activity, may be used in diseases such as bladder cancer, asthma and in cases of immunodeficiency.

The Listing of Seguences attached to this report lists the 65 sequences of interest in this invention, where SEQ ID NO:47 to SEQ ID NO:65 represent the sequences of oligonucleotides of particular interest.

Finally, it is within the scope of this invention, the complete genome of BCG Moreau RJ obtained from the mapping of insertions, deletions and point mutations of BCG Moreau based on the determination of the genomic sequence of BCG Moreau and its comparison with the genome of other mycobacteriae. The complete sequence of the BCJ Moreau RJ genome, which has 4,600,000 base pairs, obtained from the invention, was deposited in the EMBL database, whose access number is still not available.

The invention described here and the aspects touched upon must be considered as possible materializations. However, it should be clear that the invention is not limited to these materializations; and experts in the art will notice that any particular characteristic introduced into it must be understood only as something that was described to facilitate understanding and cannot be made
without one removing oneself from the inventive concept described. The limiting characteristics of the objects of this invention are related to the claims that are part of this description.
<table>
<thead>
<tr>
<th>No.</th>
<th>Begin</th>
<th>End</th>
<th>Type</th>
<th>Affected genes</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>79594</td>
<td>79594</td>
<td>Insertion of 36pb</td>
<td>Mb0072</td>
<td>In phase of insertion of 12 aminoacids (possible maturase)</td>
</tr>
<tr>
<td>2</td>
<td>425072</td>
<td>425173</td>
<td>Deletion 102 pb</td>
<td>intergenic</td>
<td>Deletion of grampo between hpsR (3') + PPE 8 (3')</td>
</tr>
<tr>
<td>3</td>
<td>581752</td>
<td>581958</td>
<td>Deletion 207 pb</td>
<td>intergenic</td>
<td>Deletion of 207 pb in 5' UTR of regX3/3'UTR of senX3 &gt; change SD of regX3</td>
</tr>
<tr>
<td>4</td>
<td>842369</td>
<td>842417</td>
<td>Deletion of 48pb</td>
<td>PE-PGRS 10</td>
<td>In phase, deletion of 16 aa</td>
</tr>
<tr>
<td>5</td>
<td>1192568</td>
<td>1192696</td>
<td>Deletion of 99pb</td>
<td>PE-PGRS 20</td>
<td>In phase, deletion of 33 aa</td>
</tr>
<tr>
<td>6</td>
<td>1284836</td>
<td>1284907</td>
<td>Deletion of 72pb</td>
<td>Mb1188c</td>
<td>In phase, deletion of 24aa (protein conserved rich in Ala-Pro)</td>
</tr>
<tr>
<td>7</td>
<td>1764646</td>
<td>1773892</td>
<td>Deletion of 9246pb</td>
<td>Mb1598A - 1613c</td>
<td>Deletion &quot;RD03&quot;: phego region phiRv1</td>
</tr>
<tr>
<td>8</td>
<td>1940956</td>
<td>1941070</td>
<td>Deletion 114pb</td>
<td>3' Mb1759 &gt; 5' MB1758</td>
<td>Intergenic - 3'UTR of possible linked to penicillin Mb1759c</td>
</tr>
<tr>
<td>9</td>
<td>1980768</td>
<td>1980803</td>
<td>Deletion of 36 pb</td>
<td>wag22b (PE-PGRS family)</td>
<td>Deletion phase of 12 aminoacids</td>
</tr>
<tr>
<td>10</td>
<td>1980909</td>
<td>1980923</td>
<td>Deletion of 15 pb</td>
<td>wag22b (PE-PGRS family)</td>
<td>Deletion phase of 5 aminoacids</td>
</tr>
<tr>
<td>11</td>
<td>2321316</td>
<td>2321393</td>
<td>Deletion of 78pb</td>
<td>Mb2108</td>
<td>In phase, deletion of 26aa</td>
</tr>
<tr>
<td>12</td>
<td>2382327</td>
<td>2382442</td>
<td>Deletion of 116pb</td>
<td>5' leu U</td>
<td>Deletion in region intergenic (loss of 2 repetições) above of leu U</td>
</tr>
<tr>
<td>13</td>
<td>3011662</td>
<td>3011835</td>
<td>Deletion of 174pb</td>
<td>PE-PGRS 47</td>
<td>In phase deletion of 58aa</td>
</tr>
<tr>
<td>14</td>
<td>3120222</td>
<td>3120075</td>
<td>Deletion of 54pb</td>
<td>PE-PGRS 48</td>
<td>In phase deletion of 18aa</td>
</tr>
<tr>
<td>15</td>
<td>3201078</td>
<td>3202053</td>
<td>Deletion of 976 pb</td>
<td>fadD26 + ppsA</td>
<td>Loss the second half of fadD16 + start ppsA (co-traduzido) &gt; may affect the translation of the remaining operon</td>
</tr>
<tr>
<td>16</td>
<td>3733171</td>
<td>3733275</td>
<td>Deletion of 105 pb</td>
<td>PE-PGRS 51</td>
<td>In phase deletion of 35aa</td>
</tr>
<tr>
<td>17</td>
<td>3769705</td>
<td>3777603</td>
<td>RD16 = Deletion of 7899 pb</td>
<td>Mb3433 - Mb3439c</td>
<td>ORFs affected: Mb3433, Mb3434, Mb3436c, Mb3437c, Mb3438c, Mb3439c,</td>
</tr>
<tr>
<td>18</td>
<td>3888915</td>
<td>3889052</td>
<td>Deletion of 138pb</td>
<td>PE-PGRS 55</td>
<td>In phase deletion of 48aa</td>
</tr>
<tr>
<td>19</td>
<td>3890900</td>
<td>3891010</td>
<td>Deletion of 111pb</td>
<td>PE-PGRS 57</td>
<td>In phase deletion of 37aa</td>
</tr>
<tr>
<td>20</td>
<td>4094370</td>
<td>4094428</td>
<td>Deletion of 59pb</td>
<td>intergenic</td>
<td>Between DNA Q and Mb3739 -</td>
</tr>
<tr>
<td>21</td>
<td>4286586</td>
<td>4296088</td>
<td>RD-01 = Deletion of 9503pb</td>
<td>3901-3909c</td>
<td>ORFs affected: Mb3901, PE35, PPE68, esxB, esxA, Mb3906, Mb3907, Mb3908, Mb3909c</td>
</tr>
<tr>
<td>22</td>
<td>4307376</td>
<td>4307376</td>
<td>Insertion of 1950 pb</td>
<td></td>
<td>3917c</td>
</tr>
<tr>
<td>23</td>
<td>4306955</td>
<td>4307375</td>
<td>Deletion of 422 pb</td>
<td></td>
<td>3917c</td>
</tr>
</tbody>
</table>
Table 3 shows the sequence of oligonucleotides of this invention

<table>
<thead>
<tr>
<th>Seq. No.</th>
<th>INDEL</th>
<th>Name</th>
<th>Bovis Begin</th>
<th>Bovis End</th>
<th>TB (H37Rv) Begin</th>
<th>TB (H37Rv) End</th>
<th>Sequence 5'→3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>47</td>
<td>16</td>
<td>RD237</td>
<td>3200739</td>
<td>3200756</td>
<td></td>
<td></td>
<td>tgccgggtttctctgtc</td>
</tr>
<tr>
<td>48</td>
<td>16</td>
<td>RD238</td>
<td>3202391</td>
<td>3202410</td>
<td></td>
<td></td>
<td>aagggcgtcagagaggt</td>
</tr>
<tr>
<td>49</td>
<td>18</td>
<td>RD244</td>
<td>377047</td>
<td>377048</td>
<td></td>
<td></td>
<td>gccgggaaccttctgttta</td>
</tr>
<tr>
<td>50</td>
<td>18</td>
<td>RD245</td>
<td>3777932</td>
<td>3777952</td>
<td></td>
<td></td>
<td>gcgtctcataatgaggttg</td>
</tr>
<tr>
<td>51</td>
<td>22</td>
<td>RD247</td>
<td>4285838</td>
<td>7285855</td>
<td></td>
<td></td>
<td>cacgctgacagagctga</td>
</tr>
<tr>
<td>52</td>
<td>22</td>
<td>RD249</td>
<td>4296369</td>
<td>4296389</td>
<td></td>
<td></td>
<td>cgggtacgctgattatagt</td>
</tr>
<tr>
<td>53</td>
<td>22</td>
<td>RD248</td>
<td>4286921</td>
<td>4286941</td>
<td></td>
<td></td>
<td>caccggcaatgaataatcttt</td>
</tr>
<tr>
<td>54</td>
<td>22</td>
<td>RD250</td>
<td>4295718</td>
<td>4295737</td>
<td></td>
<td></td>
<td>ccgtgaccaatgattgatg</td>
</tr>
<tr>
<td>55</td>
<td>18</td>
<td>RD243</td>
<td>3769255</td>
<td>3769274</td>
<td></td>
<td></td>
<td>gccgaggaagatgagaagaa</td>
</tr>
<tr>
<td>56</td>
<td>18</td>
<td>RD246</td>
<td>3777006</td>
<td>3777024</td>
<td></td>
<td></td>
<td>gccgggaacccctcttgtta</td>
</tr>
<tr>
<td>57</td>
<td>23</td>
<td>RD.235</td>
<td>4306445</td>
<td>4306463</td>
<td>4370077</td>
<td>4370095</td>
<td>ccccgagaacattagcgtg</td>
</tr>
<tr>
<td>58</td>
<td>23</td>
<td>RD.236</td>
<td>4307879</td>
<td>4307899</td>
<td>4373918</td>
<td>4373938</td>
<td>cagctccacatgatttacgaa</td>
</tr>
<tr>
<td>59</td>
<td>23</td>
<td>RD.316</td>
<td>4371762</td>
<td>4371784</td>
<td>ttgatagcgtgatgaaat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>23</td>
<td>RD.317</td>
<td>4372706</td>
<td>4372723</td>
<td>ccggctagctaatagat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>61</td>
<td>23</td>
<td>RD.318</td>
<td>4307236</td>
<td>4307253</td>
<td>4370868</td>
<td>4370855</td>
<td>cccgctcagatgttta</td>
</tr>
<tr>
<td>62</td>
<td>8</td>
<td>RD.251</td>
<td>1764434</td>
<td>1764451</td>
<td>atggccagaaggggtgtgat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>63</td>
<td>8</td>
<td>RD.252</td>
<td>1765302</td>
<td>1765322</td>
<td>acgcctctgatatcgttggta</td>
<td></td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>8</td>
<td>RD.253</td>
<td>1775275</td>
<td>1775294</td>
<td>ggcgttcagactcactgat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>8</td>
<td>RD.254</td>
<td>1773564</td>
<td>1773582</td>
<td>ttctgcctggctaatgac</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


CLAIMS

1. Nucleic acid characterized for hybridizing with a junction marker to detect the presence or absence of a deletion, or the presence of an original fragment.

2. Nucleic acid according to claim 1 characterized by having one of the sequences SEQ ID NO. 1 to SEQ ID NO: 46.

3. Nucleic acid according to claim 1 characterized for containing the junction of a deletion or insertion marker according to Table 2.

4. Nucleic acid according to claim 1, characterized by the nucleic acid hybridizing with a genome or molecule of RNA or cDNA, where the deletion is present, but not with a genome or RNA molecule, where the deletion is not present.

5. Nucleic acid according to claim 1, characterized by the genome, where the deletion is present being Mycobacterium bovis BCG Moreau.

5. Nucleic acid according to claim 1, characterized by the genome where the deletion is present being a variant of Mycobacterium tuberculosis, Mycobacterium bovis or Mycobacterium bovis BCG.

6. Initiator characterized by having one of the sequences SEQ ID NO. 47 to SEQ ID NO. 65, which are used to amplify a fragment containing the junction region.

7. Initiator characterized for hybridizing with a region in the genome where a deletion is present.

8. Kit characterized for comprising sequences SEQ ID NO.1 to SEQ ID NO. 65 for genetic identification of BCG Moreau in any type of sample, permitting to identify if an individual has been vaccinated with BCG Moreau or with
another BCG.

9. Kit characterized for comprising sequences SEQ ID NO.1 to SEQ ID NO. 65 to separate a vaccinated individual with BCG Moreau and individual who is a carriers of tuberculosis.

10. Kit according to claim 8 characterized by the sample being selected from the group, consisting of blood, saliva, secretions, serum, biopsy, aspirate, lymph, aspirate of bronchial wash, pleural liquid or other material, or even through a skin test or other test directly applied to patients.

11. Kit characterized by comprising fragments of DNA, RNA or cDNA, which permits through hybridization or amplification to express genes in BCG Moreau, or feasibility of the bacillus during production, storage and transport of a vaccine, as well as persistence of the bacteria in vaccinated individuals.

12. Use of BCG Moreau RDJ as obtained from this invention as a vaccine for mammals.

13. Use of BCG Moreau RDJ as obtained from this invention as an antitumor in bladder cancer, asthma and as an immunodulator.

14. Recombinant proteins or peptides corresponding to the insertion region in BCG Moreau, characterized by these proteins or peptides detecting an immune answer, where antibodies are induced by BCG Moreau.

15. Recombinant proteins or antibodies against the same, based on insertions, or modified proteins, specific to BCG Moreau, which permit to monitor the gene or
proteotome expression in BCG Moreau, for control, for example of quality or feasibility.

16. BCG Moreau RJ characterized for being obtained from mapping of insertions, deletions and point mutations from the genome sequence of BCG Moreau and its comparison to the genome of other mycobacteriae, whose sequence completes the genome of BCJ Moreau RJ has 4,600,000 base pairs, deposited in the EMBL database.

17. Genetically modified mycobacteriae containing an exogenous nucleic acid containing one or more deletion markers specified in Table 2.

18. Mycobacterium genetically modified according to claim 17 characterized by the mycobacteria being M. bovis BCG, and where the deletion marker is deleted as specified in Table 2.

19. Mycobacterium according to claim 17, characterized for comprising a vehicle physiologically acceptable for injection.

20. Genetically modified characterized for containing a deletion resulting from a homologous recombination in a deletion marker as specified in Table 2.

21. Mycobacteria genetically modified according to claim 20, characterized by the mycobacteria being Mycobacterium bovis.

22. Mycobacteria genetically modified according to claim 20, characterized by the mycobacteria being Mycobacterium bovis BCG.

23. Mycobacteria according to claim 21 characterized for comprising a vehicle physiologically acceptable for
24. Mycobacteria according to claim 22, characterized for comprising a vehicle physiologically acceptable for injection.

25. Mycobacteria genetically modified according to claim 20, characterized by the mycobacteria being Mycobacterium tuberculosis.

26. Mycobacteria according to claim 25, characterized for comprising a vehicle physiologically acceptable for the injection.

27. Method for distinguishing a stump from the complex of M. tuberculosis, characterized for being performed by determination of the presence of a deletion or insertion marker according to Table 2, where the deletion is absent in at least one lineage being analyzed, and where the presence of said deletion marker indicates that the lineage is not the candidate lineage.

28. Method according to claim 27, characterized by the stage of determination of the presence or not of a marker involving hybridization of the nucleic acid with the deletion marker.

29. Method according to claim 27, characterized by the stage of determination of the presence or not of a marker involving a link of antibodies with a polypeptide codified by said deletion marker.

30. Method according to claim 27, characterized by the stage of determination of the presence or not of a marker involving PCR amplification of the region containing said deletion marker.
31. Method according to claim 27, characterized by the stage of determination of the presence or not of a marker involving hybridization with a junction sequence associated to said deletion marker.