Identification of a Major Locus, TNF1, That Controls BCG-Triggered Tumor Necrosis Factor Production by Leukocytes in an Area Hyperendemic for Tuberculosis

Aurelie Cobat,1 Eileen G. Hoal,2 Caroline J. Gallant,1 Leah Simkin,1 Gillian F. Black,2 Kim Stanley,2 Jean-Philippe Jais,3 Ting-Heng Yu,1 Anne Boiland-Auge,4 Ghislain Grange,5 Christophe Delacourt,6 Paul van Helden,2 Jean-Laurent Casanova,5,7 Laurent Abel,5,7 Alexandre Alcaïs,5,7,8,a and Erwin Schurr1,a

1McGill International TB Centre and Departments of Human Genetics and Medicine, McGill University, Montreal, Quebec, Canada; 2Molecular Biology and Human Genetics, MRC Centre for Molecular and Cellular Biology, DST/NRF Centre of Excellence for Biomedical TB Research, Faculty of Health Sciences, Stellenbosch University, Tygerberg, South Africa; 3University Paris Descartes, Imagine Institute, 4National Center for Genotyping, Genomic Institute, CEA, Evry, 5Laboratory of Human Genetics of Infectious Diseases, Necker Branch, INSERM U 980, University Paris Descartes, Necker Enfants-Maladies Hospital, and 6Pediatric Pneumology Unit, Necker-Enfants-Malades Hospital, Paris, France; 7St Giles Laboratory of Human Genetics of Infectious Diseases, Rockefeller Branch, The Rockefeller University, New York, New York; and 8URC, CIC, Necker and Cochin Hospitals, Paris, France

Background. Tumor necrosis factor (TNF) is a key immune regulator of tuberculosis resistance, as exemplified by the highly increased risk of tuberculosis disease among individuals receiving TNF-blocker therapy.

Methods. We determined the extent of TNF production after stimulation with BCG or BCG plus interferon gamma (IFN-γ) using a whole blood assay in 392 children belonging to 135 nuclear families from an area hyperendemic for tuberculosis in South Africa. We conducted classical univariate and bivariate genome-wide linkage analysis of TNF production using the data from both stimulation protocols by means of an extension of the maximum-likelihood-binomial method for quantitative trait loci to multivariate analysis.

Results. Stimulation of whole blood by either BCG or BCG plus IFN-γ resulted in a range of TNF release across subjects. Extent of TNF production following both stimulation protocols was highly correlated ($r = 0.81$). We failed to identify genetic linkage of TNF release when considering each stimulus separately. However, using a multivariate approach, we detected a major pleiotropic locus ($P < 10^{-5}$) on chromosome region 11p15, termed TNF locus 1 (TNF1), that controlled TNF production after stimulation by both BCG alone and BCG plus IFN-γ.

Conclusions. The TNF1 locus was mapped in the vicinity of the TST1 locus, previously identified in the same family sample, that controls tuberculin skin test (TST) negativity per se, that is, T-cell–independent resistance to Mycobacterium tuberculosis infection. This suggested that there is a connection between TST negativity per se and TNF production.

Keywords. TNF; tuberculosis; multivariate linkage analysis; pleiotropic locus.

High fatality rates in tuberculosis persist despite the widespread use of BCG vaccine, efficient multidrug therapy, and tuberculosis control programs. The precarious situation is further complicated by the increasing incidence of tuberculosis among individuals with human immunodeficiency virus (HIV), which is more difficult to diagnose, and the steep increase and dissemination of multidrug-resistant and extremely drug-resistant tuberculosis cases, which are difficult to treat. Exposure to Mycobacterium tuberculosis does not imply that progression to disease will occur, and there is strong interindividual variability in susceptibility to both infection with M. tuberculosis and the progression...
of latent tuberculosis to clinical tuberculosis disease [1, 2]. However, the identity of the molecular factors that put a person at increased risk of becoming infected or developing tuberculosis disease is largely unknown [1]. Nevertheless, a critical role of tumor necrosis factor (TNF) in the control of M. tuberculosis infection and progression of tuberculosis disease has long been recognized. A lack of functional TNF in the murine model has been shown to exacerbate both acute and chronic M. tuberculosis infection, likely by abrogating granuloma maintenance and macrophage activation [3–6]. More recently, studies in zebrafish and nonhuman primates have confirmed the important role of TNF in the physiopathology of M. tuberculosis infection and progression of infection [7, 8]. The strongest evidence for a key role of TNF in human tuberculosis susceptibility is provided by TNF inhibitors that, when given to rheumatoid arthritis patients, can cause reactivation of latent tuberculosis or exacerbate ongoing tuberculosis [9].

Disentangling the genetic mechanisms that control the production of TNF in response to mycobacterial infection is a promising strategy that allows the unbiased identification of innate factors involved in the regulation of this important antimycobacterial effector molecule. In humans, heritability of TNF production triggered by mycobacterial antigens has been estimated at 30% to 60% [10, 11], which is consistent with the 53% heritability reported for lipopolysaccharide-stimulated TNF production [12]. Moreover, evidence for a major gene controlling M. tuberculosis–triggered TNF production has been detected by complex segregation analysis of tuberculosis cases and their household contacts in Uganda [13]. However, few studies have aimed to identify the underlying genetic variants that predispose to higher or lower TNF production. Two studies in rats identified several loci influencing TNF production [14, 15], and one genome-wide linkage scan in humans identified CD53 as an important regulator of innate TNF levels in the context of osteoarthritis [16]. However, in a genome-wide linkage scan of TNF production in response to M. tuberculosis culture filtrate in Uganda, there was no region that reached genome-wide significance [17].

A possible reason for the difficulties in mapping loci that control TNF production could be the monophenotypic analyses employed, that is, the concept of measuring TNF production after stimulation with a single trigger. TNF production is clearly a multivariate phenotype due to the multitude of TNF inducers and their complex regulatory feedbacks. Importantly, there can be considerably more information for genetic analysis in the analysis of multivariate phenotypes than in univariate analyses [18–22]. Hence, we used both univariate and multivariate linkage analyses to map loci that control TNF production in the context of mycobacterial infection in a unique population sample of 135 large nuclear families from a tuberculosis hyperendemic region of South Africa [23, 24]. In these families, we determined the production of TNF in whole blood assays following 2 distinct stimuli and then applied both univariate and multivariate linkage analyses. Only through the use of a multivariate framework were we able to identify a major locus located on chromosomal region 11p15, TNF locus 1 (TNFL), that regulates TNF production in response to mycobacterial stimulation. TNFL is located in the vicinity of TST1, a locus reported to control TST negativity per se, that is, T-cell–independent resistance to M. tuberculosis in the same South African families [25].

**MATERIALS AND METHODS**

**Subjects and Families**

Nuclear families (ie, parents and offspring) with at least 2 children were enrolled from Cape Town, South Africa, as described previously [23–25]. Subjects who had had clinical tuberculosis disease in the 2 years preceding the study were excluded. Individuals who were HIV positive, pregnant, or using immunomodulatory chemotherapy were also excluded at the time of enrollment. BCG vaccination at birth is routine in the study area and was therefore not a confounding factor in our study. Similarly, all individuals belong to the South African Coloured ethnic group, therefore limiting the risk of genetic heterogeneity. Informed consent was obtained from all study participants. Protocols involving human subjects were approved by the Stellenbosch University Health Research Ethics Committee (Tygerberg, South Africa), the University of Cape Town (Cape Town, South Africa), and the Research Ethics Board at the Research Institute of the McGill University Health Centre (Montreal, Canada).

**Whole Blood Assays**

Production of TNF was measured in whole blood using enzyme-linked immunosorbent assay after 1 day of incubation with live BCG (TNF-BCG) or BCG plus interferon gamma (IFN-γ; TNF-BCG + γ), as described previously [10, 24]. Cell cultures were stimulated in quadruplicates. Cultures incubated without stimulating antigens were used as negative controls.

Analyses were performed on the mean stimulated values from which the negative control value was subtracted. Six individuals having a TNF response to BCG plus IFN-γ lower than to BCG alone were excluded from the analysis. Because none of the available covariates (ie, age, sex, and tuberculosis disease occurring >2 years before the beginning of the study) had a significant impact on TNF production (data not shown), further genetic analyses were conducted on raw phenotypes.

**Genotyping**

High-density genotyping for the linkage study was performed at the Centre National de Génotypage (Paris, France) with the Illumina Linkage IVb Panel, containing 5657 autosomal
single-nucleotide polymorphisms (SNPs). Eleven nonpolymorphic SNPs and 196 SNPs with a call rate lower than 95% were excluded. None of the remaining SNPs showed departure from Hardy-Weinberg equilibrium among the founders at the $10^{-5}$ level. In our sample, pairwise linkage disequilibrium between adjacent SNPs was very weak, with most of the SNPs pairs having $r^2$ inferior to 0.1 (mean pairwise $r^2 = 0.07$) [25]. As described previously, 4 families were excluded on the basis of a principal component analysis of the autosomal SNPs [25]. In the present study, possible ethnic diversity of families does not result in increased type I error.

**Linkage Analysis**

We performed bivariate quantitative model-free multipoint linkage analysis of TNF response to BCG and BCG plus IFN-γ by means of the new maximum-likelihood-binomial (MLB) method for quantitative traits (nMLB-QTL version 3.0) [26, 27] as implemented in an extension of the GENEHUNTER program [28]. The MLB approach considers the sibship as a whole and does not make any assumption about the distribution of the phenotype. The idea of the MLB method is to introduce an individual latent binary variable, which captures the linkage information between the observed quantitative trait and the marker. The probability that the latent variable for an individual takes the value 0 or 1 depends on the quantitative trait value of the individual and on a link function, which can be parametric (like the standard cumulative normal distribution) or empiric (derived from the observed phenotypic distribution). Here, we used the standard cumulative normal distribution as link function. The test of linkage is a maximum likelihood ratio test (LRT) that compares the likelihoods under the null hypothesis of no linkage (H₀) and the alternative hypothesis of linkage (H₁). The test statistic is asymptotically distributed as a 50:50 mixture of a $\chi^2$ distribution with 0 and 1 degree of freedom [26, 27].

To perform a multivariate linkage analysis, we applied the approach that was proposed by Mangin and colleagues [29] and successfully applied to asthma [30]. Mangin et al showed that a 2-step procedure involving a first step of canonical decomposition of the original variables followed by a combination of the likelihood ratio statistics obtained by univariate analysis of the canonical variables was asymptotically equivalent to a multivariate maximum likelihood analysis. Here, we transformed the original variables TNF-BCG and TNF-BCG + γ into a set of 2 uncorrelated new variables, PC1 and PC2, by means of principal component (PC) analysis as implemented in the PROC PRINCOMP procedure of the SAS software package, version 9.2 (SAS Institute Inc, Cary, North Carolina). Next, univariate quantitative model-free linkage analyses of PC1 and PC2 were conducted using the nMLB-QTL approach.

Finally, we defined $T$, the multivariate test statistic, as:

$$T = LRT_{PC1} + LRT_{PC2},$$

where $LRT_{PC1}$ and $LRT_{PC2}$ denote the univariate nMLB-QTL likelihood ratio tests of PC1 and PC2, respectively. Because the 2 variables PC1 and PC2 are by definition independent, the univariate LRTs follow a 50:50:50 mixture of a $\chi^2$ distribution with 0 and 1 degree of freedom. Therefore, the asymptotic distribution of $T$ is a 25:50:25 mixture of a $\chi^2$ distribution with 0, 1, and 2 degrees of freedom [31]. $P$ values $\leq 2.3 \times 10^{-3}$ and $\leq 7 \times 10^{-4}$ were considered to be significant and suggestive of linkage at the genome-wide level, respectively, as proposed in Lander and Kruglyak [32].

**RESULTS**

**Study Sample Description**

We studied 135 families (including 392 children and 195 parents) comprising 2–6 children with available TNF production value after stimulation by BCG or BCG plus IFN-γ (Table 1). The distributions of TNF values for both stimulation protocols among 392 children who passed quality control are shown in Figure 1A. Overall, addition of IFN-γ to BCG led to higher production of TNF, as shown by the shift of the distribution to higher values. The mean TNF values were 336.4 pg/mL (SD, 292.3 pg/mL) for BCG and 1560.9 pg/mL (SD, 922.3 pg/mL) for BCG plus IFN-γ. As illustrated in Figure 1B, TNF values for stimulation by BCG and BCG plus IFN-γ were highly correlated (Pearson correlation coefficient $r = 0.81$).

**Univariate Linkage Analysis**

Because our sample was composed of families with an average of 3 children, we decided to analyze the genetic control of TNF production by linkage analysis. We first performed univariate model-free linkage analyses of TNF production after BCG alone or BCG plus IFN-γ stimulation. Results are shown in Figure 2 for BCG and BCG plus IFN-γ. By analyzing the 2 phenotypes separately, none of the linkage peaks reached a genome-wide significant level ($P \leq 2.3 \times 10^{-5}$) [32]. For TNF

<table>
<thead>
<tr>
<th>No. of Parents Genotyped</th>
<th>No. of Siblings per Family</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>32</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
</tr>
<tr>
<td>Total</td>
<td>51</td>
</tr>
</tbody>
</table>
induced by BCG, we identified 2 linkage peaks with a \( P \) value \( \leq 0.01 \) on chromosome 1p36 (\( P = 3.7 \times 10^{-3} \) at position 17.3 Mb; information content [IC] = 0.95) and on chromosome 6q14 (\( P = 5.9 \times 10^{-3} \) at position 81.6 Mb; IC = 0.97) (Table 2). For TNF induced by BCG plus IFN-\( \gamma \), we identified a suggestive linkage signal on chromosome 11p15 (\( P = 9.9 \times 10^{-5} \) at position 12.6 Mb; IC = 0.90) and a weaker linkage signal on chromosome 19q13 (\( P = 4.4 \times 10^{-3} \) at position 58.9 Mb; IC = 0.97).

**Bivariate Linkage Analysis**

In the context of correlated traits, multivariate linkage analysis may detect major loci that cannot be found by univariate linkage analysis. Thus, we performed a bivariate model-free linkage analysis of TNF production, that is, a joint analysis of TNF levels after stimulation with BCG and BCG plus IFN-\( \gamma \), to search for pleiotropic QTLs that impact TNF production after stimulation by both BCG alone and BCG plus IFN-\( \gamma \). The analytical strategy used is described in Figure 3. Bivariate linkage analysis detected a highly significant pleiotropic linkage signal exceeding the stringent threshold for genome-wide significance on chromosomal region 11p15 with a multipoint \( P \) value of \( 1.94 \times 10^{-6} \) at position 12.6 Mb (Figure 4). In addition to this major pleiotropic locus, 4 weaker linkage peaks were observed on chromosomal regions 5p15 (\( P = 4.4 \times 10^{-3} \)), 4q32 (\( P = 6.0 \times 10^{-3} \)), 7q36 (\( P = 6.1 \times 10^{-3} \)), and 19q13 (\( P = 8.9 \times 10^{-3} \)).

### Table 2. Chromosomal Regions Significant at the .01 Level by Phenotypes

<table>
<thead>
<tr>
<th>Chromosomal Region</th>
<th>Range (Mb)</th>
<th>TNF-BCG</th>
<th>TNF-BCG + ( \gamma )</th>
<th>TNF-Bivariate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1p36</td>
<td>16.6–19.5</td>
<td>.004</td>
<td>.044</td>
<td>.013</td>
</tr>
<tr>
<td>4q32–33</td>
<td>162.7–171.6</td>
<td>.032</td>
<td>.022</td>
<td>.006</td>
</tr>
<tr>
<td>5p15</td>
<td>0.6–6.6</td>
<td>.006</td>
<td>.147</td>
<td>.004</td>
</tr>
<tr>
<td>6q14</td>
<td>81.1–83.1</td>
<td>.006</td>
<td>.012</td>
<td>.015</td>
</tr>
<tr>
<td>7q36</td>
<td>152.9–155.2</td>
<td>.073</td>
<td>.039</td>
<td>.006</td>
</tr>
<tr>
<td>11p15</td>
<td>5.6–21.9</td>
<td>.012</td>
<td>.93 \times 10^{-5}</td>
<td>1.94 \times 10^{-6}</td>
</tr>
<tr>
<td>19q13</td>
<td>58.1–59.1</td>
<td>.022</td>
<td>.004</td>
<td>.009</td>
</tr>
</tbody>
</table>

Abbreviations: TNF, tumor necrosis factor; TNF-BCG, tumor necrosis factor induced by BCG vaccine; TNF-BCG + \( \gamma \), tumor necrosis factor induced by BCG vaccine plus interferon-\( \gamma \).
(Table 2). A list of the known genes located under \textit{TNF1} is given in Supplementary Table 1.

Careful examination of the results for each PC at the major locus showed that both of the 2 PCs contributed to the linkage signal of \textit{TNF1} on chromosomal region 11p15 (Table 3). The first PC, which explained 90% of the overall TNF variability, contributed to the linkage signal at the major locus for 38%. This PC was composed of the equally weighted sum of TNF production after stimulation with BCG and BCG plus IFN-$\gamma$ and can be interpreted as nonspecific overall TNF production. The second PC, which explained 10% of the TNF variability, contributed 62% of the linkage signal. This second PC was composed of the difference between the 2 stimulations and can be interpreted as the specific contribution of IFN-$\gamma$ to TNF production. These results suggest that \textit{TNF1} is a gene involved in the regulation of TNF production following BCG stimulation that itself is regulated by IFN-$\gamma$.

**DISCUSSION**

We identified a major locus on chromosome region 11p15 that controls TNF production after BCG and BCG plus IFN-$\gamma$ stimulation in an area hyperendemic for tuberculosis in South Africa. Surprisingly, this major locus was mapped to the vicinity of the \textit{TST1} locus that controls TST negativity per se, that is, T-cell–independent resistance to \textit{M. tuberculosis} infection, identified in the same family sample [25] and replicated in an independent sample from Paris, France (Cobat et al, manuscript in preparation). This overlap is suggestive of a previously unknown link of genetically controlled mycobacteria-driven TNF production and resistance to infection with \textit{M. tuberculosis}. The identification of the molecular cause of the \textit{TNF1} locus will follow the strategy successfully applied in leprosy [33, 34] and tuberculosis [35–37].

Our results demonstrate that a multivariate approach is a powerful tool for the genetic analysis of a biological complex phenotype. Indeed, the increase in information achieved in multivariate analyses, resulting from the covariance between measures, can substantially boost power of linkage detection compared to univariate analyses [18]. Here, we extended the new MLB-QTL method [27] to multivariate linkage analysis using a 2-step procedure. The analysis strategy is based on a decomposition of the PCs of the phenotypes followed by a combination of the univariate statistics obtained for each PC. This
had been initially proposed by Mangin et al in the context of variance component analysis [29] and was later applied to the Haseman-Elston method [31] and the original MLB-QTL method [30]. The 2-step strategy presents the advantage of being easily applicable to any number of traits, while avoiding the increased numerical complexity and computational burden associated with the increased number of estimated parameters inherent to fully multivariate methods. Application of the employed strategy for the genetic study of multivariate traits is not limited to genome-wide linkage analysis but can be equally applied to genome-wide association scans of multivariate phenotypes.

Surprisingly, the major QTL and 2 of the 4 weaker peaks on chromosomal regions 5p15 and 19q13 were in the vicinity of linkage peaks identified for TST negativity per se, that is, T-cell–independent resistance to M. tuberculosis infection, identified in the same family sample [25]. In particular, the major pleiotropic QTL for TNF production (TNF1) was mapped at 14 Mb from TST1, the major locus controlling TST negativity per se, in this family sample. Such close proximity is consistent with a genetic location overlap of the 2 loci [38]. The overlap between TNFI and TST1 was even more pronounced when considering the location of TST1 in an independent replication sample from Paris, France (Cobat et al, manuscript in preparation). In the latter sample, TST1 was mapped at 2 Mb from TNFI. The genetic location overlap of TNFI with TST1 is suggestive of a previously unknown link of genetically controlled mycobacteria-driven TNF production and resistance to infection with M. tuberculosis.

TNF is a key effector molecule of tuberculosis resistance that is critical for the sequestration of M. tuberculosis in infectious granuloma during latent tuberculosis. TNF is not required for granuloma formation but controls granuloma integrity via its microbicidal effect on M. tuberculosis [7, 39]. Mycobacterium tuberculosis–containing granulomas appear microscopically as

**Table 3. Contribution of Each Principal Component to the Major Linkage Peak**

<table>
<thead>
<tr>
<th>Principal Component</th>
<th>Coefficient for Standardized TNF-BCG</th>
<th>Coefficient for Standardized TNF-BCG + γ</th>
<th>Proportion of Phenotypic Variance Explained</th>
<th>Contribution to TNF1</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC1</td>
<td>0.71</td>
<td>0.71</td>
<td>0.90</td>
<td>0.38</td>
</tr>
<tr>
<td>PC2</td>
<td>0.71</td>
<td>-0.71</td>
<td>0.10</td>
<td>0.62</td>
</tr>
</tbody>
</table>

Abbreviations: PC1, first principal component; PC2, second principal component; TNF-BCG, tumor necrosis factor secreted following BCG stimulation; TNF-BCG + γ, tumor necrosis factor secreted following BCG plus interferon-γ stimulation.
distinct rings of multinucleated giant cells, epithelioid macrophages, and lymphocytes surrounding central necrotic areas. *Mycobacterium tuberculosis* bacteria remain viable for many years within the granuloma, and the breakdown of granuloma integrity leads to poor mycobacterial containment and tuberculosis disease long after primary infection [3]. Although the detailed bactericidal mechanism of TNF is still unknown, the control of TNF over *M. tuberculosis* replication in newly infected macrophages makes any regulator of TNF levels following exposure to mycobacteria a strong candidate for an *M. tuberculosis* infection resistance factor. Hence, the genetic mapping of a locus controlling TNF production provides strong biological support for the interpretation of TST1 as a *M. tuberculosis* infection resistance locus [25].

**Supplementary Data**

Supplementary materials are available at Clinical Infectious Diseases online (http://cid.oxfordjournals.org/). Supplementary materials consist of data provided by the author that are published to benefit the reader. The posted materials are not copyedited. The contents of all supplementary data are the sole responsibility of the authors. Questions or messages regarding errors should be addressed to the author.

**Notes**

**Acknowledgments.** We sincerely thank all members of the community who participated in this study. We are grateful to Christine Johannisen, Magdalene Kennedy, Ronell Taylor, and Maria Gallant for their help with field work and blood collection; Gian van der Spuy, Neal Rogers, and Mari de Klerk for assisting with data capture and database management; and the Centre National de Génotypage for providing the genotypes.

**Financial support.** This work was supported by the Canadian Institutes of Health Research (grant 111172 to E. S.); the Grand Challenge exploration of the Bill and Melinda Gates Foundation (grant 5157); and the legs Poix (grant UMRS980 to A. A.). A. C. is the recipient of a Banting postdoctoral fellowship (BPF 112932) from the government of Canada. J.-L. C., L. A., and A. A. are supported by grants from the French National Agency for Research, European Research Council (ERC-2010-AdG-268777), Institut National de la Santé et de la Recherche Médicale, Assistance Publique-Hopitaux de Paris, and the National Center for Research Resources and National Center for Advancing Sciences of the National Institutes of Health (grant number 8ULTR000043).

**Potential conflicts of interest.** All authors: No reported conflicts. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.
References


