Supplementary Methods

Sample Collections

Most of the sample collections used here have been described in considerable detail [1-8].

For the IPD group, DNA samples were collected from patients with IPD (defined by the isolation of *Streptococcus pneumoniae* from a normally sterile site) in Oxfordshire, UK [5]. The first control group (1) comprises UK cord blood samples (567 individuals), healthy adult blood donors (78 individuals), and healthy adults from a Human Random Control (HRC) panel (96 individuals). Genotype frequencies were near-identical between adult and cord blood controls (data available upon request); the association between S180L and IPD remained significant even after the omission of the cord blood controls, i.e. analysing solely the adult controls against the IPD cases (3x2 $\chi^2 = 6.8$, P=0.03).

Individuals with thoracic empyema were recruited upon entry to the UK MIST1 trial of intrapleural streptokinase for the treatment of pleural infection [6]. Entry criteria were as follows: the presence of pleural fluid that was macroscopically purulent or positive on bacterial culture or Gram's stain, or pleural fluid pH < 7.2 in the setting of clinical evidence of infection [6]. The bacteriology of the cohort has been described in detail: analysis of pleural fluid was performed using both standard bacterial culture and amplification and sequencing of the bacterial 16s ribosomal RNA gene. Using these techniques, a bacteriological diagnosis was made in 59% of the available samples. The breakdown of bacterial species within this group was as follows: *Streptococcus pneumoniae* 27%, *Streptococcus intermedius-anginosus-constelatus (milleri)* group 24%, other streptococcal species 7%, *Staphylococcus aureus* 8%, anaerobes 16%, Gram-negative bacteria 6%, others 12%. The second control group

(control (2)) was collected independently of the first and consisted of 370 healthy adult blood donors from the UK. UK individuals who were not of European ancestry were excluded from the studies.

For the Gambian malaria study, children aged 1 to 10 years were enrolled between August 1988 and November 1990 at the Royal Victoria Hospital of Banjul, and the Medical Research Council hospital of Fajara, The Gambia [1]. The criteria for diagnosis of malaria were as follows:

1. Parasitaemia greater than 2500/μl of blood.
2. Appropriate clinical history
3. Laboratory investigations exclude other diagnoses

Cerebral malaria was diagnosed if the following criteria were fulfilled:

1. Blantyre coma score of ≤ 2, persisting for > 30 minutes after effective treatment of hypoglycaemia or convulsions)
2. Repeated prolonged seizure (> 30 minutes) in a child with *P. falciparum* parasitaemia and no other apparent cause of fits or coma.

Severe anaemia secondary to malaria was defined as a haemoglobin level of < 50 g/L on admission with concomitant parasitaemia. Children were scored for severe malaria if they suffered from either cerebral malaria or severe malarial anaemia. Children with mild malaria had an uncomplicated febrile illness with *P. falciparum* parasitaemia and no other apparent causes of fever.

The controls were recruited at both hospitals and health centres in the study area, and were frequency matched to the cases for age and area of residence. These children had mild, mostly infectious diseases (e.g. mild cough and cold) and illness that did not require

hospital admission and did not have parasitaemia on blood films. These children were not matched specifically for ethnic groups. Instead, correction for the potential confounding effect of ethnicity was performed using binary logistic regression.

The Kenyan children (<13 years old) with bacteraemia were recruited from the Kilifi District hospital between 1998 and 2002 [4]. The 593 bacteraemic cases comprised patients with isolated Gram-positive and Gram-negative infections, diagnosed using standard blood culture techniques. The most frequent organisms isolated were *Streptococcus pneumoniae* (25.5%), *Salmonella typhi* (16.6%), *Haemophilus influenzae* (14.6%), and *Escherichia coli* (8.7%), as well as other less common bacteria. The community controls were frequency matched to a subset of the cases (two per case) on the basis of time (recruited within 14 days), location of homestead, age, and sex. Only children with complete data for HIV, malnutrition, and malaria status were included in the analysis. Ethical approval was given by the Kenya Medical Research Institute (KEMRI) National Scientific Steering and Research Committees. Full details of the study can be obtained from Berkley *et al.* 2005 [4].

The Vietnamese subjects were adults suffering from severe malaria at the Centre for Tropical Diseases, Ho Chi Minh City, Vietnam recruited between 1991 and 1996. Controls were ethnically matched adult non-malaria cases recruited from other departments of the same hospital. Subjects with mild malaria were not included in this study. Ethical approval was granted by the Ethical and Scientific committee of the Centre for Tropical Diseases in Ho Chi Minh City, Vietnam.

For the Kenyan malaria study, children admitted to the high dependency unit of Kilifi District Hospital with severe malaria[8] were recruited as cases. The controls were children who had been recruited from the community surrounding the hospital for a study of bacteraemia, matched by age, sex and household location to children admitted to the hospital with bacteraemia.

For the West African EUTB study, samples were drawn from three countries - The Gambia, Guinea Bissau and the Republic of Guinea - using standardised study design and diagnostic criteria [2, 3]. Suspected cases were identified by field workers on the basis of clinical criteria including cough (or more than 3 weeks duration), fever, weight loss, night sweats, chest pain, haemoptysis and PPD positive skin response and referred for clinical examination. Cases confirmed by sputum positivity on smear or culture for *Mycobacterium tuberculosis* were included in the study. Individuals were tested for HIV status. HIV rates differed between countries. In The Gambia, the rates were 7.5% cases vs. 5.2% controls, in Guinea Bissau 23.8% cases vs. 16.3% controls and in the Republic of Guinea 9.3% cases vs. 5.0% controls. 67.3% of cases were male reflecting the sex bias amongst sufferers of tuberculosis. There was no significant difference in frequency of ethnic groups between cases and controls in any of the three countries studied ($P = 0.70$, 0.62 and 0.22 respectively).

For the Algerian TB study, a retrospective and prospective study of tuberculosis (TB) was initiated at the Institut Pasteur d'Algérie from March 1999 – June 2003. The study collected blood samples from TB patients and their immediate families; i.e. siblings and parents. The patients and their families were from the Algerian capital of Algiers, and were grouped according to their residency in one of the administrative “sectors of

health". These sectors are: Bab El Oued (BEO), Mustapha (MUS) and Les Sources (SCE), located in the city centre; Sidi Moussa (SDM), in the South; Château Neuf (CHN), in the South-East; and Staoueli (STL) in the East. Each family was given a code consisting of the 3-letter code for the sector in which they lived and a family number. The age of the donors ranged from 7 – 83. The initial case of pulmonary TB diagnosed and treated in a family was designated the index case. Diagnosis was based on positive histology and bacteriology. Follow-ups of the index case and its family were carried out to detect secondary TB in the former, and pulmonary or extra-pulmonary TB in the latter. Recruitment of the family involved explaining the basics of the study to the index case or the head of the family. Once consent was obtained, the whole family was asked to attend the clinic for a blood sample. 10mL of blood was taken from each member of the family, and placed in 200µL of EDTA. The blood was stored at -20^0C until DNA extraction. DNA extraction was with "Nucleon BACC2" kits, as per the protocol therein.

Ethical approval for all studies has been granted by the relevant authorities.

PCR conditions

For assays running on the Sequenom MassArray platform, a standard touch-down PCR protocol was used for both uniplex and multiplex assays. The cycling conditions were as follows:

1. 95ºC for 15 minutes
2. 94ºC for 20 seconds
3. 65ºC for 30 seconds
4. 72ºC for 30 seconds

Steps 2 to 4 to be repeated for 5 cycles

5. 94ºC for 20 seconds
6. 58ºC for 30 seconds
7. 72ºC for 30 seconds

Steps 5 to 7 to be repeated for 5 cycles

8. 94ºC for 20 seconds
9. 53ºC for 30 seconds
10. 72ºC for 30 seconds

Steps 8 to 10 to be repeated for 38 cycles

A final extension at 72ºC for 3 minutes.

General PCR conditions for amplifying products prior to sequencing:

1. 95ºC for 15 minutes
2. 95ºC for 30 seconds
3. 55-65ºC for 30 seconds
4. 72ºC for 1 minute

Steps 2-4 are repeated for 38 cycles

5. A final extension at 72ºC for 5 minutes.

Details for genotyping of marker Ser180Leu by RFLP.

Forward primer: CTCCAGGGGCCGAGGGCTGCACCATCCCC[C→A]TGCTG

Reverse primer: TACTGTAGCTGAATCCCGTTCC

Restriction enzyme used: *BstX*I (NEB)

Assay conditions: 1unit of enzyme, incubated at 55ºC overnight. Visualized by transillumination on 4% agarose gel.

Supplementary Methods References

1. Hill, A.V. *et al.* Common West African HLA antigens are associated with protection from severe malaria. Nature. 1991 Aug 15; **352(6336)**: 595-600.

2. Lienhardt, C. *et al.* Investigation of environmental and host-related risk factors for tuberculosis in Africa. I. Methodological aspects of a combined design. Am J Epidemiol. 2002 Jun 1; **155(11)**: 1066-73.

3. Bennett, S. *et al.* Investigation of environmental and host-related risk factors for tuberculosis in Africa. II. Investigation of host genetic factors. Am J Epidemiol. 2002 Jun 1;**155(11):**1074-9.

4. Berkley, J.A. *et al.* Bacteremia among children admitted to a rural hospital in Kenya. N Engl J Med. 2005 Jan 6;**352(1)**:39-47.

5. Roy, S. *et al.* MBL genotype and risk of invasive pneumococcal disease: a case-control study. Lancet 2002; **359**: 1569-73.

6. Maskell, N.A., *et al*. UK Controlled trial of intrapleural streptokinase for pleural infection. N Engl J Med 2005; **352**: 865-74.

7. Chapman, S.J., *et al*. PTPN22 and Invasive Bacterial Disease. Nat Genet 2006; **38**: 499-500.

8. Marsh, K., Forster, D., Waruiru, C., Mwangi, I., Winstanley, M., Marsh, V., Newton, C., Winstanley, P., Warn, P., Peshu, N., *et al.* (1995) Indicators of life-threatening malaria in African children. N Engl J Med. **332**(21):1399-404