

1 *ACTR1A* has pleiotropic effects on risk of leprosy, inflammatory  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000

James J Gilchrist<sup>1,2,3\*</sup>, Kathryn Auckland<sup>3</sup>, Tom Parks<sup>3,4</sup>,  
Alexander J Mentzer<sup>3</sup>, Lily Goldblatt<sup>5</sup>, Vivek Naranbhai<sup>3</sup>, Gavin Band<sup>3</sup>,  
Kirk A Rockett<sup>3</sup>, Ousmane B Toure<sup>6</sup>, Salimata Konate<sup>6</sup>, Sibiri Sissoko<sup>6</sup>,  
Abdoulaye A Djimdé<sup>6</sup>, Mahamadou A Thera<sup>6</sup>, Ogobara K Doumbo<sup>6†</sup>,  
Samba Sow<sup>7</sup>, Sian Floyd<sup>8</sup>, Jörg M Pönnighaus<sup>9</sup>, David K Warndorff<sup>9</sup>,  
Amelia C Crampin<sup>9</sup>, Paul EM Fine<sup>8</sup>, Benjamin P Fairfax<sup>2</sup>, Adrian VS Hill<sup>3,10\*</sup>

<sup>1</sup>Department of Paediatrics, University of Oxford, Oxford, UK

<sup>2</sup>MRC–Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, UK

<sup>3</sup>Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK

<sup>4</sup>Department of Infectious Diseases, Imperial College London, London, UK.

<sup>5</sup>Balliol College, Oxford, UK.

<sup>6</sup>Malaria Research and Training Centre, University of Science, Techniques and Technologies of Bamako, Bamako, Mali.

<sup>7</sup>Center for Vaccine Development, Bamako, Mali.

<sup>8</sup>Faculty of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, London, UK.

<sup>9</sup>Malawi Epidemiology and Intervention Research Unit (formerly Karonga Prevention Study), Chilumba, Malawi.

<sup>10</sup>Jenner Institute, University of Oxford, Oxford, UK.

\*Corresponding authors

Email: james.gilchrist@paediatrics.ox.ac.uk (JJG), adrian.hill@ndm.ox.ac.uk (AVSH)

†Deceased

January 27, 2022

## 18 Abstract

19 Leprosy is a chronic infection of the skin and peripheral nerves caused by *Mycobacterium leprae*. De-  
20 spite recent improvements in disease control, leprosy remains an important cause of infectious disability  
21 globally. Large-scale genetic association studies in Chinese, Vietnamese and Indian populations have  
22 identified over 30 susceptibility loci for leprosy. There is a significant burden of leprosy in Africa, how-  
23 ever it is uncertain whether the findings of published genetic association studies are generalizable to  
24 African populations. To address this, we conducted a genome-wide association study (GWAS) of leprosy  
25 in Malawian (327 cases, 436 controls) and Malian (247 cases, 368 controls) individuals. In that analysis,  
26 we replicated five risk loci previously reported in China, Vietnam and India; MHC Class I and II, *LACC1*  
27 (2 independent loci) and *SLC29A3*. We further identified a novel leprosy susceptibility locus at 10q24  
28 (rs2015583: combined  $p = 8.81 \times 10^{-9}$ ;  $OR = 0.51$  [95% CI 0.40 – 0.64]). The leprosy risk locus is  
29 a determinant of *ACTR1A* RNA expression in CD4<sup>+</sup> T cells (posterior probability of colocalization -  
30  $PP = 0.96$ ). Furthermore, it demonstrates pleiotropy with established risk loci for inflammatory bowel  
31 disease and atopic disease. Reduced *ACTR1A* expression decreases susceptibility to leprosy and atopy  
32 but increases risk of inflammatory bowel disease. A shared genetic architecture for leprosy and inflam-  
33 matory bowel disease has been previously described. We expand on this, strengthening the evidence that  
34 selection pressure driven by leprosy has shaped the evolution of autoimmune and atopic disease in mod-  
35 ern populations. More broadly, our data highlights the importance of defining the genetic architecture  
36 of disease across genetically diverse populations, and that disease insights derived from GWAS in one  
37 population may not translate to all affected populations.

## 38 Author Summary

39 Leprosy remains a leading cause of infectious disability globally. Human genetic variation is a major  
40 determinant of susceptibility to infection, including leprosy. Large-scale genetic association studies have  
41 been pivotal in advancing our understanding of leprosy biology. These studies have been performed in  
42 Chinese, Vietnamese and Indian populations, and it remains unclear whether these insights are infor-  
43 mative of leprosy susceptibility in African populations. To address this, we performed a genome-wide  
44 association study of leprosy susceptibility in Malawi and Mali. In doing so we replicate known leprosy  
45 susceptibility loci at MHC class I and II, *LACC1* and *SLC29A3*. Furthermore, we identify a novel leprosy  
46 susceptibility locus, which modifies expression of *ACTR1A* in CD4<sup>+</sup> T cells and demonstrates pleiotropy  
47 with inflammatory bowel disease (IBD) and atopic disease. These data deepen our understanding of  
48 leprosy biology, identifying *ACTR1A* expression in CD4<sup>+</sup> T cells as a determinant of leprosy disease  
49 risk, and further define the role of this ancient pathogen in the evolution of immune-mediated diseases  
50 in modern populations.

## 51 Introduction

52 Leprosy is a chronic infectious disease affecting the skin and peripheral nerves caused by *Mycobacterium*  
53 *leprae*. It is a leading infectious cause of disability (Britton & Lockwood, 2004). The introduction of  
54 multidrug therapy (“Chemotherapy of leprosy. Report of a WHO Study Group.”, 1994), widespread  
55 use of BCG vaccination (Pönnighaus et al., 1992), and the 1991 World Health Assembly resolution to  
56 eliminate leprosy by the year 2000 have all contributed to a decline in disease burden; nevertheless,  
57 over 200,000 new cases continue to be reported annually ([https://www.who.int/gho/neglected\\_diseases/](https://www.who.int/gho/neglected_diseases/leprosy/en/)  
58 [leprosy/en/](https://www.who.int/gho/neglected_diseases/leprosy/en/)), numbers which are likely to represent a considerable underestimate of the true disease  
59 burden (Smith et al., 2015).

60 Large-scale, unbiased genetic association studies in Chinese (Liu et al., 2015; Liu et al., 2017; Wang  
61 et al., 2018; Wang et al., 2016; F.-R. Zhang et al., 2009; F. Zhang et al., 2011), Indian (Wong et al.,  
62 2010) and Vietnamese (Gzara et al., 2020) populations have identified and validated 34 genetic loci  
63 independently associated with leprosy outside the HLA region, as well as independent HLA class I and  
64 class II associations. A key feature of these studies has been the demonstration of considerable genetic  
65 heterogeneity in leprosy susceptibility between populations. For instance, while genetic variation at *TLR1*  
66 associates with leprosy risk in Indian and Turkish populations, this finding has not been replicated in  
67 Chinese and Vietnamese populations. To date, there are no published genome-wide association studies  
68 (GWAS) of leprosy in African populations. This is important as there remains a significant burden of  
69 leprosy in Africa and the observed inter-population heterogeneity reported in the Chinese, Vietnamese  
70 and Indian studies suggest that the published GWAS findings may not be generalizable to African  
71 populations. To address this, we have performed a GWAS of leprosy susceptibility in Malawian and  
72 Malian individuals.

## 73 Results

### 74 Leprosy genome-wide association study

75 Individuals with leprosy were recruited to the study following clinical and microbiological evaluation  
76 within the Karonga Prevention Study (KPS), Karonga, Malawi. Healthy controls were recruited from  
77 the same population. Following quality control and genome-wide imputation, genotypes at 10,511,695  
78 loci from 612 samples (284 cases, 328 controls) were included in the association analysis. Inspection of  
79 the QQ plot (Supplementary Fig. 1) and the genomic inflation parameter ( $\lambda = 1.0333$ ) demonstrates  
80 that inclusion of the six major principal components as covariates in the model adequately controls for  
81 confounding variation. In that analysis, we identified 142 loci, representing 38 independent genomic  
82 loci, with suggestive evidence of association ( $p < 1 \times 10^{-5}$ ) with leprosy in Malawian populations

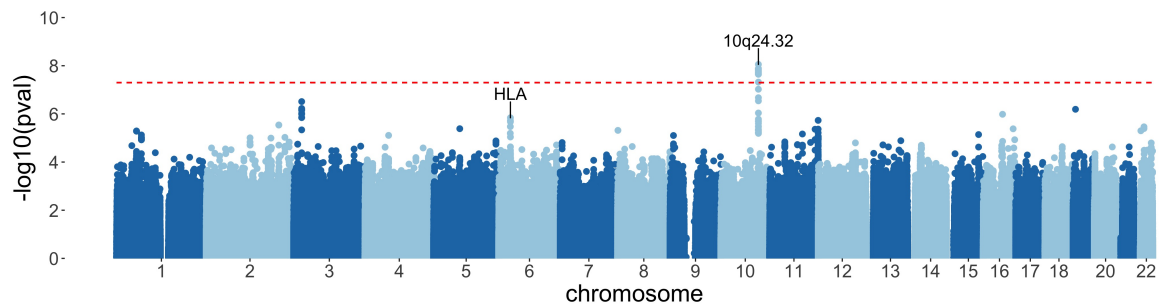


Figure 1: **Manhattan plot of leprosy in Malawi and Mali.**

Evidence for association with leprosy at genotyped and imputed autosomal SNPs and indels ( $n = 9,616,523$ ) in Malawi and Mali (492 cases, 639 controls). Association statistics represent a fixed-effects meta-analysis of additive association with disease in Malawi and Mali. The red, dashed line denotes genome-wide significance ( $p = 5 \times 10^{-8}$ ).

83 (Supplementary Fig. 2, Supplementary Table 1).

## 84 **GWAS replication and meta-analysis**

85 We sought to replicate evidence for leprosy association observed in Malawi among leprosy cases and  
86 healthy controls in Mali. Individuals with leprosy were recruited to the study at Mali's former national  
87 leprology centre, Institut Marchoux, (now Hôpital Dermatologique de Bamako), Bamako. Healthy con-  
88 trols were recruited from the same population. Among the Malian samples, 10,514,676 loci and 519  
89 individuals (208 cases, 311 controls) passed QC filters. The QQ plot (Supplementary Fig. 1) and  
90 genomic inflation parameter ( $\lambda = 1.0498$ ) of genome-wide association analysis in the Malian samples  
91 demonstrates control of confounding variation with inclusion of the major six major principal compo-  
92 nents and genotyping platform as covariates in the model. We combined evidence for leprosy association  
93 in Malawi and Mali using a fixed-effects meta-analysis (Fig. 1). Of the 142 leprosy-associated loci identi-  
94 fied in the discovery analysis, 18 SNPs, at a single genomic locus at 10q24.32 (Fig. 2A), showed evidence  
95 of replication in Mali ( $p < 0.05$ ) and overall evidence of association with leprosy exceeding genome-wide  
96 significance ( $p < 5 \times 10^{-8}$ ). The variant with the strongest evidence for leprosy association at that locus  
97 is rs2015583:  $p = 8.81 \times 10^{-9}$ ,  $OR = 0.51$  (95% CI 0.40 – 0.64). There is no evidence for heterogeneity of  
98 effect between populations at rs2015583 (heterogeneity  $p = 0.871$ , Fig. 2B), and the data best supports  
99 a model in which rs2015583 modifies risk of both paucibacillary and multibacillary leprosy (log10 Bayes  
100 factor = 6.01, Fig. 2B).

101 Fine-mapping of the leprosy association at chr10q24.32 identifies a credible set of 19 SNPs with a  
102 95% probability of containing the causal variant, spanning a 6kb region: chr10:104,226,830-104,232,809  
103 (Supplementary Table 2). Genetic variation at this locus has not been previously described as a de-  
104 terminant of leprosy risk. In keeping with the observation that the human genetics of leprosy risk is  
105 characterized by considerable heterogeneity of effect between populations, there is no evidence for leprosy  
106 association ( $p < 0.05$ ) in Chinese populations (8,156 cases, 15,610 controls) at rs2015583 ( $p = 0.587$ ,

107  $OR = 0.98$ ) or at any of the 19 variants in the credible SNP set (Dr Z Wang and Prof. F Zhang, personal  
108 communication, November 2019).

### 109 **A leprosy risk locus modifies *ACTR1A* expression in CD4<sup>+</sup> T cells**

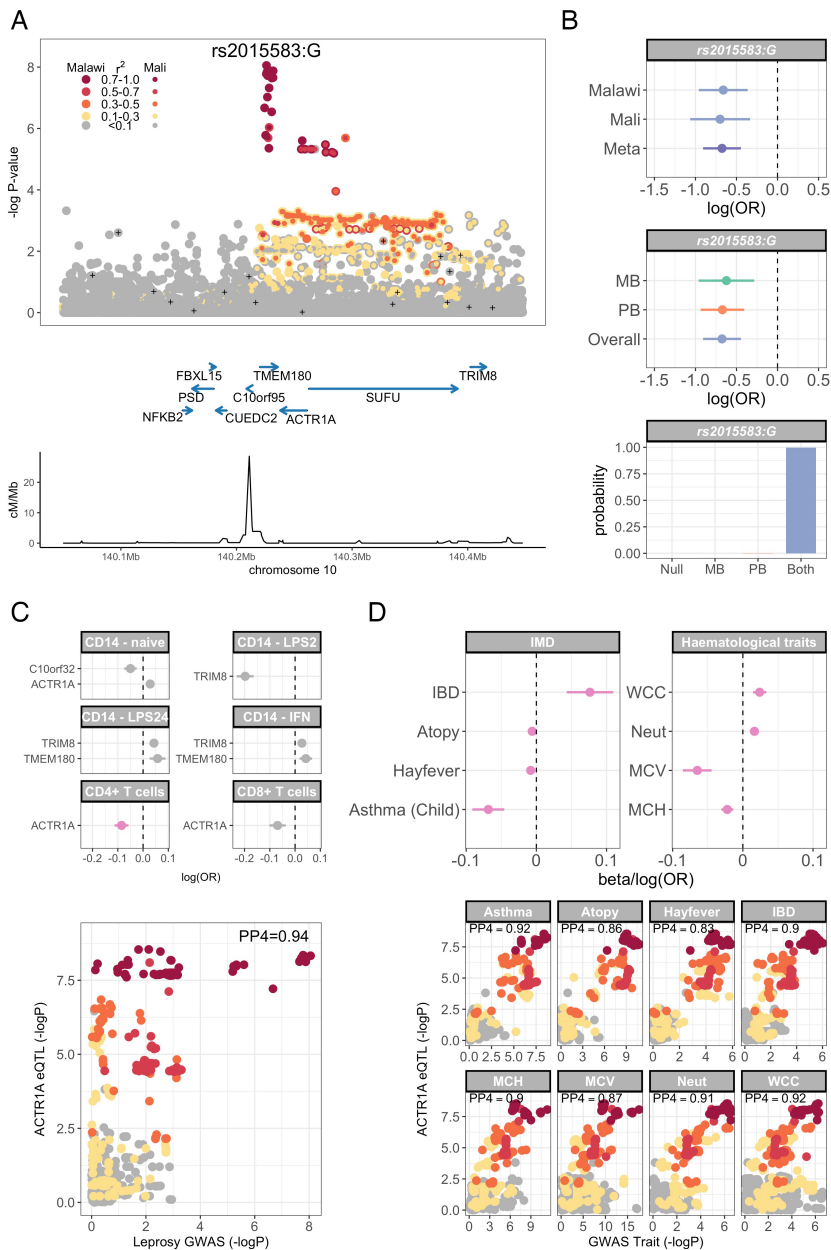
110 Trait-associated genetic variation identified by GWAS are highly enriched for regulatory variation. To  
111 explore whether leprosy-associated genetic variation at chr10q24.32 modifies leprosy risk through its  
112 effect on gene expression, we investigated whether the leprosy risk locus colocalizes with expression  
113 quantitative trait loci (eQTL) in a range of primary immune cells; monocytes (Fairfax et al., 2014), B  
114 cells (Fairfax et al., 2012), NK cells (Gilchrist et al., 2021), neutrophils (Naranbhai et al., 2015), CD4<sup>+</sup>  
115 T cells and CD8<sup>+</sup> T cells (Kasela et al., 2017). These data demonstrate colocalization of the leprosy risk  
116 locus at chr10q24.32 with an eQTL for *ACTR1A* expression in CD4<sup>+</sup> T cells (posterior probability of  
117 colocalization,  $PP4 = 0.94$ ), but not gene expression in other cell types (Fig. 2C, Supplementary Table  
118 3). The leprosy protective allele, rs2015583:G, is associated with reduced *ACTR1A* expression in CD4<sup>+</sup>  
119 T cells ( $p = 4.69 \times 10^{-9}$ ,  $\beta = -0.085$ ).

### 120 ***ACTR1A* expression has pleiotropic effects on risk of immune-mediated dis-** 121 **ease**

122 A key feature of genetic determinants of leprosy described to date has been the identification of  
123 pleiotropy at leprosy risk loci with other immune-mediated diseases, in particular inflammatory bowel  
124 disease (Jostins et al., 2012; Sun et al., 2016). To explore whether leprosy-associated genetic variation  
125 modifying *ACTR1A* expression is a determinant of other immune-mediated diseases in human popula-  
126 tions, we assessed evidence of colocalization at the CD4<sup>+</sup> T cell *ACTR1A* eQTL with immune-mediated  
127 diseases (n=29) and haematological traits (n=26). In that analysis (Fig. 2D, Supplementary Table 4),  
128 in addition to leprosy, the *ACTR1A* eQTL in CD4<sup>+</sup> T cells colocalizes with four immune-mediated dis-  
129 eases (inflammatory bowel disease, atopy, childhood-onset asthma and hayfever), and four haematological  
130 traits (white cell count, neutrophil count, mean corpuscular volume and mean corpuscular haemoglobin).  
131 Decreased *ACTR1A* expression in CD4<sup>+</sup> T cells is associated with perturbations of haematological in-  
132 dices (increased neutrophil and white cell counts, decreased mean corpuscular haemoglobin and volume),  
133 increased risk of inflammatory bowel disease and decreased risk of atopy (atopy as a composite trait,  
134 childhood-onset asthma and hayfever).

### 135 **Replication of leprosy HLA associations**

136 The observation that class I and II HLA alleles are key determinants of leprosy risk has been highly  
137 reproducible across diverse populations (Gzara et al., 2020; Wang et al., 2016; Wong et al., 2010).

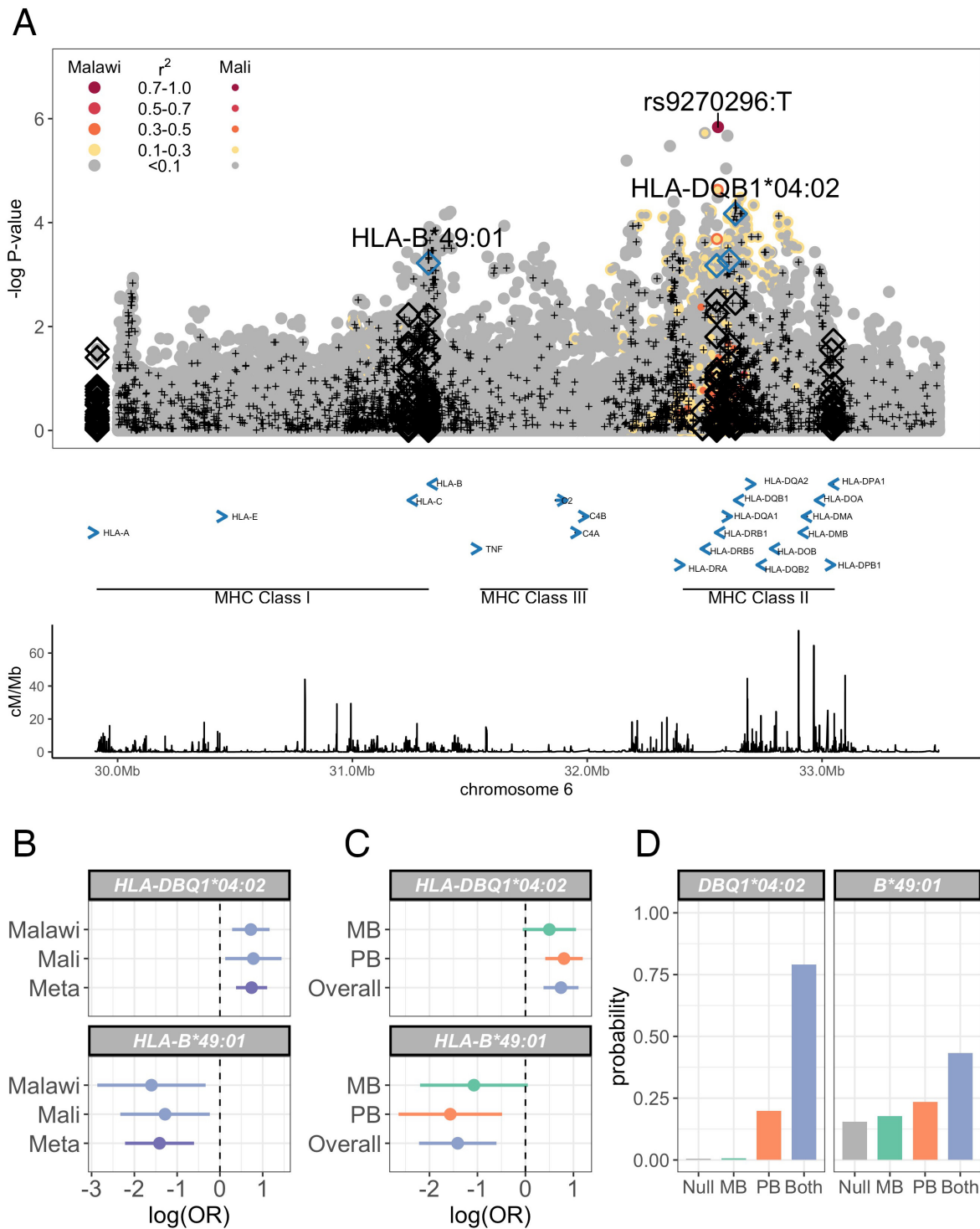


**Figure 2: Leprosy association, regulatory function and pleiotropy at chromosome 10q24.32.** (A) Regional association plot of leprosy association at chr10q24.32. Association statistics represent a fixed-effects meta-analysis of additive association with disease in Malawi and Mali. SNPs are coloured according to linkage disequilibrium to rs2015583, and genotyped SNPs marked with black pluses. (B) Log-transformed odds ratios and 95% confidence intervals of rs2015583 association with leprosy in Malawi and Mali (top) and stratified by multibacillary and paucibacillary disease (middle). Posterior probabilities of models of rs2015583 association with leprosy: “Null”, no association with leprosy; “MB”, non-zero effect in multibacillary leprosy alone; “PB”, non-zero effect in paucibacillary leprosy alone; “Both”, the same non-zero effect is shared by individuals with multibacillary and paucibacillary leprosy. (C) Log-transformed odds ratios and 95% confidence intervals of rs2015583 association with gene expression in primary immune cells (top). Associations which colocalize with the leprosy association signal (*ACTR1A* expression in CD4<sup>+</sup> T cells) are highlighted in pink. The *ACTR1A* eQTL in CD4<sup>+</sup> T cells colocalizes with the risk locus for leprosy at chr10q24.32 (bottom). SNPs are coloured according to linkage disequilibrium to rs2015583 as above. (D) Log-transformed odds ratios and 95% confidence intervals of rs2015583 association (top) with immune-mediated diseases (IBD, inflammatory bowel disease; atopy; hayfever; childhood-onset asthma) and hematological indices (WCC, white cell count; Neut, neutrophil count; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin). The *ACTR1A* eQTL in CD4<sup>+</sup> T cells colocalizes with the GWAS locus for each trait at chr10q24.32 (bottom). SNPs are coloured according to linkage disequilibrium to rs2015583 as above.

138 Motivated by this, and by evidence of association in the HLA observed in our data (Fig. 1), we explored  
139 evidence for leprosy association in the HLA in Malawi and Mali at the level of SNPs and classical HLA  
140 alleles. In a fixed-effects meta-analysis of leprosy association in Malawi and Mali (Fig. 3A, Supplementary  
141 Table 5), the peak classical allele association is a class II allele: HLA-DQB1\*04:02 ( $p = 6.74 \times 10^{-5}$ ,  
142  $FDR = 0.0063$ ,  $OR = 2.1$  95% CI 1.75 – 2.51). We also observed a leprosy association in the class I  
143 region, at HLA-B\*49:01 ( $p = 6.02 \times 10^{-4}$ ,  $FDR = 0.0156$ ), which is independent of HLA-DQB1\*04:02  
144 (Supplementary Fig. 3). No significant residual associations were observed after conditioning on both  
145 DQB1\*04:02 and HLA-B\*49:01 (Supplementary Fig. 4). There is no evidence for heterogeneity of effect  
146 between populations at HLA-DQB1\*04:02 or HLA-B\*49:01 (heterogeneity  $p = 0.8836$  and  $p = 0.7049$ ,  
147 Fig. 3B), and the data best supports a model in which both HLA-DQB1\*04:02 and HLA-B\*49:01 modify  
148 risk of both paucibacillary and multibacillary leprosy (log10 Bayes factors = 2.22 and 0.43 respectively,  
149 Fig. 3C,D).

## 150 Replication of known leprosy associations outside the HLA

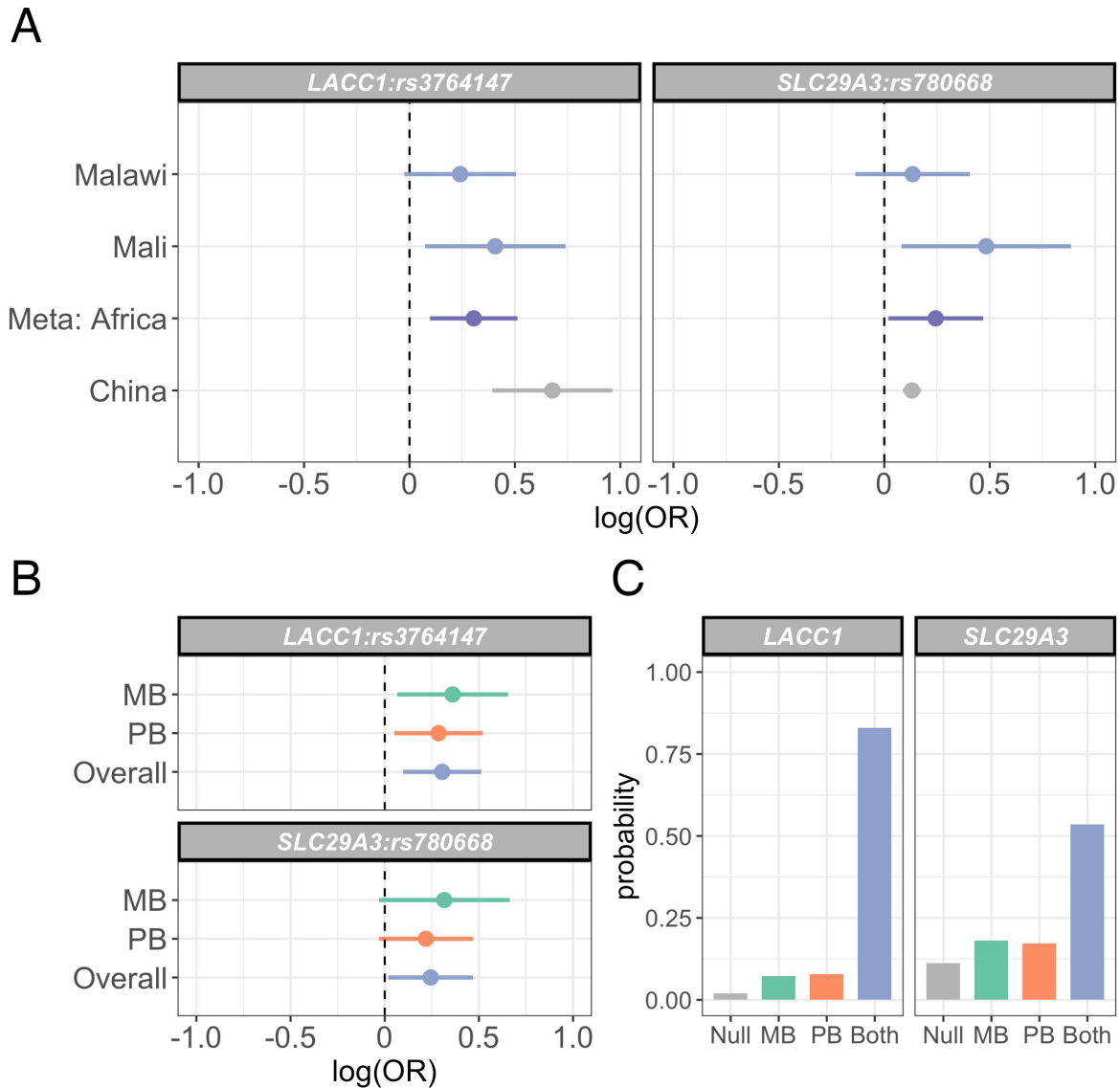
151 Our identification of a genetic variant modifying leprosy risk in African populations, but not in Chinese  
152 populations, highlights the inter-population heterogeneity that has been observed across large-scale ge-  
153 netic studies of leprosy susceptibility (Gzara et al., 2020; Wang et al., 2016; Wong et al., 2010). Our  
154 study has adequate power ( $> 80\%$ ) to replicate ( $p < 0.05$ ) findings at 7 of 34 previously-published leprosy  
155 risk loci outside the HLA (Supplementary Table 6). We were able to replicate leprosy associations at  
156 2 loci (Fig. 4A); a missense SNP in *LACCI*, rs3764147 ( $p = 0.004$ ,  $OR = 1.36$  95% CI 1.10 – 1.67),  
157 and a missense SNP in *SLC29A3*, rs780668 ( $p = 0.034$ ,  $OR = 1.28$  95% CI 1.02 – 1.60). There is  
158 no evidence for heterogeneity of effect between populations at rs3764147 or rs780668 (heterogeneity  
159  $p = 0.444$  and  $p = 0.159$ , Fig. 4A), and the data best supports a model in which both rs3764147 and  
160 rs780668 modify risk of both paucibacillary and multibacillary leprosy (log10 Bayes factors = 1.64 and  
161 0.68 respectively, Fig. 4B,C). Among the 6 loci at which we were not able to demonstrate replication of  
162 previously-published leprosy association despite adequate study power, there is no evidence that our lack  
163 of replication reflects effects restricted to multibacillary or paucibacillary disease (Supplementary Table  
164 7). We further considered whether our failure to replicate previously-reported leprosy associations could  
165 represent differential linkage disequilibrium to an undefined causal locus between study populations. To  
166 test this, we examined evidence for leprosy association within 250kb of each previously-reported leprosy  
167 risk locus outside the HLA. In that analysis we identified a promoter variant in *RAB32*, rs34271799,  
168 with suggestive evidence of association with leprosy risk in Malawi and Mali ( $p_{MALAWI} = 0.0023$ ,  
169  $p_{MALI} = 0.0034$ ,  $p_{COMBINED} = 6.00 \times 10^{-5}$ ;  $OR = 0.42$ , 95%CI = 0.28 – 0.64, Supplementary Fig.  
170 5). There was no evidence suggestive of leprosy association ( $p < 1 \times 10^{-4}$ ) within 250kb of any other  
171 previously identified leprosy risk locus.



**Figure 3: MHC leprosy association in Malawi and Mali.**

(A) Regional association plot of leprosy association across the HLA region. Association statistics represent a fixed-effects meta-analysis of additive association with disease in Malawi and Mali. SNPs are coloured according to linkage disequilibrium to rs9270296, and genotyped SNPs marked with black pluses. Imputed classical HLA alleles are plotted as diamonds, with significantly associated (FDR < 0.05) alleles highlighted in blue. (B) Log-transformed odds ratios and 95% confidence intervals of HLA-DBQ1\*04:02 and HLA-B\*49:01 associations with leprosy in Malawi and Mali. (C) Log-transformed odds ratios and 95% confidence intervals of HLA-DBQ1\*04:02 and HLA-B\*49:01 associations with leprosy stratified by multibacillary and paucibacillary disease. (D) Posterior probabilities of models of HLA-DBQ1\*04:02 and HLA-B\*49:01 associations with leprosy: “Null”, no association with leprosy; “MB”, non-zero effect in multibacillary leprosy alone; “PB”, non-zero effect in paucibacillary leprosy alone; “Both”, the same non-zero effect is shared by individuals with multibacillary and paucibacillary leprosy.





**Figure 4: Replication of leprosy associations at *LACC1* and *SLC29A3* in Malawi and Mali.** (A) Log-transformed odds ratios and 95% confidence intervals of rs3764147 and rs780668 associations with leprosy in Malawi and Mali. (B) Log-transformed odds ratios and 95% confidence intervals of rs3764147 and rs780668 associations with leprosy stratified by multibacillary and paucibacillary disease. (C) Posterior probabilities of models of rs3764147 and rs780668 associations with leprosy: “Null”, no association with leprosy; “MB”, non-zero effect in multibacillary leprosy alone; “PB”, non-zero effect in paucibacillary leprosy alone; “Both”, the same non-zero effect is shared by individuals with multibacillary and paucibacillary leprosy.

## 172 Discussion

173 In this study, we demonstrate that genetic variation at chromosome 10q24.32 is a determinant of leprosy  
174 risk in African populations. In common with many examples of trait-associated genetic variation identi-  
175 fied by GWAS (Maurano et al., 2012), variation at 10q24.32 modifies risk of leprosy through regulatory  
176 effects on gene expression, specifically *ACTR1A* expression in CD4<sup>+</sup> T cells. We expand upon this,  
177 identifying evidence of pleiotropy at 10q24.32, demonstrating a shared genetic architecture of leprosy,  
178 inflammatory bowel disease and atopy at this locus. Furthermore, we replicate previously identified  
179 leprosy susceptibility loci at *LACC1*, *SLC29A3*, and with HLA Class I and II alleles.

180 *ACTR1A* encodes actin-related protein-1 (ARP1), a component of the dynactin complex. Dynactin  
181 interacts with the cytoplasmic motor, dynein, facilitating intracellular trafficking of a wide range of intra-  
182 cellular cargos (Roberts et al., 2013; Urnavicius et al., 2015). In T cells specifically, dynactin/dynein com-  
183 plexes direct the accumulation of TCRs and secretory vesicles at the immunological synapse (Hashimoto-  
184 Tane et al., 2011; Nath et al., 2016), and are required for nuclear translocation of NF $\kappa$ B in response to T  
185 cell stimulation (Shrum et al., 2009). The identification of a CD4<sup>+</sup> T cell-specific eQTL as a determinant  
186 of leprosy reflects the established role of T cell-mediated immunity in leprosy biology. The spectrum of  
187 leprosy disease is defined by host T cell responses, with tuberculoid disease being characterised by ro-  
188 bust CD4<sup>+</sup> T cell IFN $\gamma$  responses and patients with lepromatous disease failing to mount anti-*M. leprae*  
189 cell-mediated responses (Yamamura et al., 1992). Similarly, leprosy reversal reactions, characterised by  
190 painful inflammation of leprosy lesions, are association with infiltration of IFN $\gamma$  producing CD4<sup>+</sup> T cells  
191 (Britton, 1998). In addition, the highly reproducible association between MHC class II alleles and lep-  
192 rosy strongly suggests a role for *M. leprae* antigen presentation as a determinant of leprosy susceptibility  
193 *per se*. Our data are complementary to these observations, suggesting a model in which CD4<sup>+</sup> T cell  
194 activation, determined at the level of the T cell as well as that of the antigen presenting cell, modifies  
195 susceptibility to leprosy.

196 The identification of a CD4<sup>+</sup> T cell-specific eQTL as a risk locus for inflammatory bowel disease  
197 and atopy is consistent with the known biology of autoimmune and atopic disease. GWAS-identified  
198 risk loci for both autoimmune and atopic disease are highly enriched for regulatory variation active in  
199 CD4<sup>+</sup> T cells, in particular regulatory CD4<sup>+</sup> T cells (Bossini-Castillo et al., 2019). Our data provide  
200 further support for the observation that leprosy and inflammatory bowel disease have a shared genetic  
201 architecture (Jostins et al., 2012). We expand on this observation, suggesting a model in which selection  
202 pressure imposed by *M. leprae* has shaped the evolution of atopic disease and autoimmune disease in  
203 modern populations.

204 Our identification of a genetic locus modifying leprosy susceptibility in African populations, but with  
205 no effect on leprosy risk in well-powered GWAS in Chinese populations, is consistent with the existence  
206 of heterogeneity of genetic architecture of leprosy risk between populations. In keeping with this, we

207 observe no evidence of leprosy association in Malawi or Mali at 6 of the 7 non-HLA loci at which we have  
208 adequate study power to assess this. We hypothesized that some of these inter-population differences  
209 may reflect differential effects of genetic risk loci on multibacillary and paucibacillary disease, however we  
210 find no evidence to support this in our data. Differential linkage disequilibrium between assayed variation  
211 and a shared causal locus may explain some of the observed inter-population genetic heterogeneity of  
212 leprosy risk. In keeping with this we observe modest evidence of leprosy association at *RAB32*, which is  
213 distinct from that reported in Chinese populations. Understanding whether genetic variation at *RAB32*  
214 is associated with leprosy risk in African populations, and whether this is distinct from that observed in  
215 Chinese populations, will require replication in additional study populations.

216 Here we define regulatory variation at *ACTR1A* as a novel determinant of leprosy susceptibility in  
217 African populations. Moreover, regulatory variation at *ACTR1A* has pleiotropic effects on hematological  
218 indices in European populations and risk of IBD and atopy. A shared genetic architecture for leprosy  
219 and IBD has been previously described. We expand on this, strengthening the evidence that selection  
220 pressure driven by leprosy has shaped the evolution of immune-mediated disease in modern populations.  
221 Our colocalization analyses identify *ACTR1A* as a potential therapeutic target for autoimmune and  
222 atopic disease, and deepens our understanding of leprosy biology, which will be key in informing the  
223 development of novel control strategies. More broadly, our data highlights the importance of defining  
224 the genetic architecture of disease across genetically diverse populations, and that disease insights derived  
225 from GWAS in one population may not readily translate to all affected populations.

## 226 **Materials and methods**

### 227 **Ethics and consent**

228 All cases and controls were recruited following informed consent of the participant or their parent/guardian.  
229 The study protocol detailing recruitment and sample collection within KPS, Malawi was approved by the  
230 National Health Sciences Research Committee of Malawi and by the Ethics Committee of the London  
231 School of Hygiene and Tropical Medicine. The study protocols detailing recruitment and sample collec-  
232 tion at the Institut Marchoux, Mali and at the Centre Hospitalier Universitaire Gabriel Toure, Bamako,  
233 Mali were approved by the University of Bamako Ethics Review Board. Genotyping and imputation for  
234 additional Malian controls were collected as part of the MalariaGEN project, for which the study pro-  
235 tocol was reviewed by Oxford University Tropical Research Ethics committee (OXTREC), Oxford, UK  
236 (OXTREC 020-006). The overall study design, including re-appraisal of study samples using genome-  
237 wide genotyping, was reviewed and approved by Oxford University Tropical Research Ethics committee  
238 (OXTREC), Oxford, UK (OXTREC 560-15).

## 239 Study Participants

240 Leprosy case and control samples were recruited to the study in Karonga, Malawi and Bamako, Mali.  
241 Participant recruitment in Malawi (Fitness et al., 2004; Wallace et al., 2004) and Mali (Meisner et al.,  
242 2001) have been described previously. In brief, cases of leprosy were diagnosed in both settings on the  
243 basis of clinical examination, split skin smear and histopathologic examination of biopsy material. Adults  
244 or children with “certain” or “probable” cases were considered eligible for recruitment (Ponninghaus et al.,  
245 1987). Cases were further defined as having paucibacillary (PB) or multibacillary (MB) disease on the  
246 basis of clinical examination and bacteriological index  $> 1$  on slit-skin smear or biopsy.

247 In Malawi, cases (n=327) and controls (n=436) were recruited to the study within the KPS, a long-  
248 term community-based, epidemiological study in Northern Malawi (Ponninghaus et al., 1987). Leprosy  
249 cases were identified through active population surveys in the 1980s, followed by enhanced passive case  
250 detection in the 1990s. Control samples were individuals within the KPS with no history or clinical  
251 features of leprosy, matched to case samples with respect to age, sex and geographic area of residence.

252 In Mali, between April and June 1997, patients with leprosy (n=247) presenting to Institut Marchoux,  
253 Bamako, Mali (formerly Mali’s national leprology center now Hôpital Dermatologique de Bamako) were  
254 recruited to the study as described previously (Meisner et al., 2001). Control participants (n=185),  
255 following exclusion of leprosy by clinical examination and history, were recruited in the same study  
256 setting among hospital staff and patients with diagnoses other than leprosy. In addition, we supplemented  
257 control numbers in the Malian replication study using healthy control samples (n=183), also recruited  
258 in Bamako, collected as part of the MalariaGEN project (Malaria Genomic Epidemiology, 2019; Toure  
259 et al., 2012).

260 Cases in Malawi were recruited at a median age of 47 years (range 15 to 83 years) and 187 were  
261 female (57%). Among leprosy cases in Malawi, 47 had multibacillary disease and 275 paucibacillary  
262 disease. Control samples in Malawi were recruited at a median age of 44 years (range 15 to 82 years)  
263 and 273 were female (63%). Cases in Mali were recruited at a median age of 45 years (range 13 to 85  
264 years) and 136 (55%) were female. Among leprosy cases in Mali, 165 had multibacillary disease and 82  
265 paucibacillary disease. Control samples in Mali were recruited at a median age of 30 years (range 14  
266 to 72 years) and 113 were female (62%). Additional MalariaGEN control samples were recruited at a  
267 median age of 3 years (range 0 to 15 years) and 91 (50%) were female.

## 268 Genotyping

269 Genomic DNA was extracted from study samples as described previously (Fitness et al., 2004; Meisner  
270 et al., 2001). Following quantification (Ahn et al., 1996), samples were genotyped using the Illumina  
271 African Diaspora Power Chip platform (Johnston et al., 2017) and genotypes called using GenCall  
272 in GenomeStudio (Illumina). Using consensus strand information from the array manifest file and a

273 remapping pipeline (Dr William Rayner, Wellcome Centre for Human Genetics, Oxford) we aligned  
274 genotypes such that all alleles are on the forward strand. Throughout genetic positions reflect GRCh37.

## 275 **Sample quality control**

276 We calculated per sample quality control (QC) metrics in PLINK (Purcell et al., 2007). For each sample  
277 we calculated the proportion of missing genotype calls, heterozygosity and the mean X and Y channel  
278 intensities. We plotted mean X and Y channel intensities (Supplementary Fig. 6) and missingness against  
279 heterozygosity (Supplementary Fig. 7), defining outlier samples using ABERRANT (Bellenguez et al.,  
280 2012). We used PLINK to estimate sample sex from genotype data, excluding samples with discordant  
281 genotype and metadata sex information. To identify duplicated and related samples, we calculated  
282 pairwise relatedness between samples in PLINK. We considered samples with relatedness  $> 0.75$  to  
283 be duplicates, and additionally identified samples with relatedness  $> 0.2$ . In both cases we retained  
284 the sample with the highest genotyping call rate of a duplicated/related sample pair. We calculated  
285 principal components (PC) in EIGENSTRAT (Price et al., 2006). To identify population outliers, we  
286 plotted study sample PCs against a background of African Genome Variation Project (Gurdasani et  
287 al., 2015) samples, identifying outliers by visual inspection (Supplementary Fig. 8). For both PC and  
288 relatedness computations we used an LD-pruned SNP set with regions of high linkage disequilibrium  
289 (LD) excluded. The first two major principal components differentiate self-reported ethnicity in both  
290 Malawi (Supplementary Fig. 9) and Mali (Supplementary Fig. 10).

## 291 **SNP quality control**

292 Prior to genome-wide imputation, we extracted genotypes from non-duplicated, autosomal SNPs and  
293 applied the following SNP QC filters; SNP missingness  $> 10\%$ , minor allele frequency (MAF)  $< 1\%$ ,  
294 Hardy-Weinberg equilibrium (HWE)  $p < 1 \times 10^{-20}$  and plate effect  $p < 1 \times 10^{-6}$ . HWE was calculated  
295 among control samples for each cohort. Plate effect represents an association test of nondifferential  
296 missingness with the plate on which each sample was genotyped.

## 297 **Imputation**

298 Following sample (Supplementary Table 8) and SNP (Supplementary Table 9) QC, genotypes at 351,236  
299 autosomal SNPs from 612 samples (Malawi) and genotypes at 367,433 autosomal SNPs from 350 samples  
300 (Mali) were taken forward for phasing and genome-wide imputation. We performed phasing of genotypes  
301 in SHAPEIT2 (Delaneau et al., 2012), phasing genotypes across each chromosome for all samples from  
302 each country jointly. We then imputed untyped autosomal genotypes using IMPUTE2 (v2.3.0) (B. Howie  
303 et al., 2011; B. N. Howie et al., 2009), in 5Mb chunks using the 1000 Genomes Phase III as a reference  
304 panel. We used 250kb buffer regions and effective sample size of 20,000.

## 305 HLA imputation

306 We used HLA\*IMP:03 (Motyer et al., 2016) to impute classical HLA alleles into our datasets. As input  
307 to HLA\*IMP we used genotypes passing sample and SNP QC thresholds in the HLA region (chr6:28Mb-  
308 36Mb). HLA\*IMP:03 uses a multi-population reference panel, including individuals of African ancestry.  
309 HLA imputation performed well, with estimated imputation accuracies ranging between 95% and 99.8%.  
310 We took forward HLA alleles present in both Malawi and Malawi (MAF > 1%), including 93 classical  
311 alleles (42 class I and 51 class II) in downstream analysis.

## 312 Additional cross-platform quality control

313 We noted that relatively few Mali control samples (n=142) were available for inclusion in the association  
314 analysis. To address this, we used genotypes from additional control samples (n=183) of representative  
315 ethnicity collected as part of the MalariaGEN project (Malaria Genomic Epidemiology, 2019; Toure et  
316 al., 2012). These samples have been genotyped on the Illumina Omni 2.5M platform. Sample genotypes  
317 have been phased using SHAPEIT2 (Delaneau et al., 2012) and untyped genotypes imputed genome-wide  
318 using IMPUTE2 (v2.3.2) (B. Howie et al., 2011; B. N. Howie et al., 2009) with 1000 Genomes Phase  
319 III as a reference panel. The SNP and sample QC used in processing these samples (Malaria Genomic  
320 Epidemiology, 2019) is highly analogous to the QC we applied to our study samples. MalariaGEN SNP  
321 QC excluded poorly genotyped SNPs using the following metrics; SNP missingness (thresholds 2.5-10%  
322 dependent on study population), MAF < 1%, HWE  $p < 1 \times 10^{-20}$ , plate effect  $p < 1 \times 10^{-3}$  and a  
323 recall test quantifying changes in genotype following a re-clustering process  $p < 1 \times 10^{-6}$ . MalariaGEN  
324 sample QC excluded samples prior to imputation according to the following metrics; channel intensity,  
325 missingness, heterozygosity (outlier thresholds determined by ABERRANT (Bellenguez et al., 2012)),  
326 population outliers and duplicated samples (relatedness > 0.75). Of note, related samples (relatedness  
327 > 0.2) are retained for imputation purposes.

328 We defined a shared subset of SNPs genotyped and passing SNP QC on both platforms ( $n = 26, 136$ ),  
329 from which we calculated relatedness estimates and PCs. We removed one of related pairs (relatedness  
330 > 0.2) from the MalariaGEN samples, which resulted in a final sample size of 519 (208 cases, 311  
331 controls). Inspection of the PCs demonstrate no further population outliers, and that the 10 major PCs  
332 are nondifferential with respect to genotyping array (Supplementary Fig. 10).

## 333 Association analysis

334 Following imputation, SNPs were taken forward for association analysis if they passed the following QC  
335 metrics; MAF > 4%, imputation info score > 0.4, HWE  $p < 1 \times 10^{-10}$ . For the Mali samples, these QC  
336 filters were applied overall and for samples genotyped on each genotyping platform individually. At each

337 variant passing QC we tested for association with leprosy case-control status in a logistic regression model  
338 in SNPTEST (Marchini et al., 2007) in each cohort. At loci of interest, we used multinomial logistic  
339 regression, implemented in SNPTEST, to estimate the effect of the genetic variation on leprosy risk  
340 stratified by multibacillary and paucibacillary disease. We used control status as the baseline stratum  
341 and cases of multibacillary and paucibacillary leprosy as strata. To account for confounding variation, in  
342 particular population structure, we included the six major principal components of genotyping data in  
343 all models. In addition, in Mali, we included genotyping platform as an additional categorical covariate.  
344 At variants passing QC thresholds in both cohorts, we then performed genome-wide meta-analysis under  
345 a frequentist fixed-effects model using BINGWA (Band et al., 2015). For association analysis using HLA  
346 allele imputations we coded posterior probabilities of each HLA allele to represent carriage of 0, 1 or 2  
347 copies of that allele. Association analysis and meta-analysis was performed in SNPTEST and BINGWA  
348 as above. For HLA association analysis we corrected for the number of classical alleles tested ( $n=93$ )  
349 and considered  $FDR < 0.05$  to be significant.

## 350 Bayesian comparison of models of association

351 We compared models of association at loci of interest with multibacillary and paucibacillary leprosy, as  
352 estimated by multinomial logistic regression, using a Bayesian approach. We considered four models of  
353 effect, defined by the prior distributions on the effect size:

354 “Null”: effect size = 0, i.e. no association with leprosy.

355 “MB”: effect size  $N(0, 0.2^2)$  for multibacillary disease, but no effect in paucibacillary disease.

356 “PB”: effect size  $N(0, 0.2^2)$  for paucibacillary disease, but no effect in multibacillary disease.

357 “Both”: effect size  $N(0, 0.2^2)$  and fixed between multibacillary and paucibacillary disease ( $\rho = 1$ ).

358 For each model we calculated approximate Bayes factors (Wakefield, 2009) and posterior probabilities,  
359 assuming each model to be equally likely a priori. Statistical analysis was performed in R.

## 360 Definition of credible SNP sets

361 We used a Bayesian approach to identify a set of SNPs with 95% probability of containing the causal  
362 locus at the leprosy susceptibility locus at chr10q24.32. Approximate Bayes’ factors (Wakefield, 2009)  
363 were calculated for each SNP in the region (a 200kb surrounding rs2015583) with a prior distribution of  
364  $N(0, 0.2^2)$ . All SNPs were considered equally likely to be the causal variant a priori. A set of SNPs with  
365 95% probability of containing the causal SNP was defined as the smallest number of SNPs for which the  
366 summed posterior probabilities exceed 0.95.

## 367 eQTL mapping and colocalization analysis

368 We used the R package coloc (Giambartolomei et al., 2014) to identify evidence of causal variants  
369 shared by eQTL in primary immune cells and GWAS-identified trait associated loci (including leprosy).  
370 Coloc adopts a Bayesian approach to compare evidence for independent or shared association signals  
371 for two traits at a given genetic locus. We tested for colocalization between leprosy susceptibility at  
372 the chr10q24.32 locus and previously-published eQTL mapping studies in naïve and stimulated primary  
373 immune cells from individuals of European ancestry (Fairfax et al., 2014; Fairfax et al., 2012; Gilchrist  
374 et al., 2021; Kasela et al., 2017; Naranbhai et al., 2015); NK cells (n = 245), B cells (n = 283), monocytes  
375 (n = 414), CD4<sup>+</sup> T cells (n = 293), CD8<sup>+</sup> T cells (n = 283), neutrophils (n = 101), LPS-stimulated  
376 monocytes (2 hours, n = 261; 24 hours, n = 322) and IFN $\gamma$ -stimulated monocytes (n = 267). We  
377 considered evidence for colocalization for each gene within a 250kb window of the peak leprosy associ-  
378 ation (rs2015583). We considered a posterior probability > 0.8 supporting a shared causal locus to be  
379 significant.

380 To assess evidence for pleiotropy with other disease traits we again used coloc to test for the presence of  
381 a shared causal locus between the *ACTR1A* eQTL in CD4<sup>+</sup> T cells and 55 GWAS traits (Supplementary  
382 Table 4); hematological indices (n = 26) and immune-mediated diseases (n = 13) from the UK Biobank  
383 (<http://www.nealelab.is/uk-biobank/>, accessed 26th March 2021), immune-mediated diseases from the  
384 NHGRI-EBI GWAS Catalog (n = 13, [ftp://ftp.ebi.ac.uk/pub/databases/gwas/summary\\_statistics/](ftp://ftp.ebi.ac.uk/pub/databases/gwas/summary_statistics/), ac-  
385 cessed 26th March 2021), and inflammatory bowel disease traits from the International Inflammatory  
386 Bowel Disease Genetics Consortium (n = 3, <https://www.ibdgenetics.org/downloads.html>, accessed 26th  
387 March 2021).

## 388 Data Availability

389 On publication, genotype and phenotype data for Malawian and Malian cases and controls will be made  
390 available via the European Genotype-Phenome Archive. Genotype and phenotype data describing the  
391 MalariaGEN Malian samples have been deposited in the European Genome-Phenome Archive (EGA;  
392 study accession EGAS00001001311). Access to MalariaGEN datasets on EGA is by application to an  
393 independent data access committee. On publication, a full set of association summary statistics will  
394 be made available for download through the the NHGRI-EBI GWAS Catalog ([https://www.ebi.ac.uk/  
395 gwas/downloads/summary-statistics](https://www.ebi.ac.uk/gwas/downloads/summary-statistics)).



## 396 Acknowledgements

397 This publication uses genotyping data from the MalariaGEN consortial project, as described in Malaria  
398 Genomic Epidemiology Network, et al. Nature Communications, 2019 (<https://doi.org/10.1038/s41467-019-13480-z>). We thank Stuart Mucklow and Giles Warner for their assistance in collecting samples in  
400 Mali. JJG and AJM are funded by National Institute for Health Research (NIHR) Clinical Lectureships.  
401 During this work AVSH was supported by a Wellcome Trust Senior Investigator Award (HCUZZ0) and  
402 by a European Research Council advanced grant (294557). The research was supported by the Wellcome  
403 Trust Core Award Grant Number 203141/Z/16/Z with additional support from the NIHR Oxford BRC.  
404 The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the  
405 Department of Health and Social Care.

## 406 References

- 407 Ahn, S. J., Costa, J., & Emanuel, J. R. (1996). Picogreen quantitation of dna: Effective evaluation of  
408 samples pre- or post-pcr. *Nucleic Acids Res*, *24*(13), 2623–2625. [https://doi.org/10.1093/nar/](https://doi.org/10.1093/nar/24.13.2623)  
409 [24.13.2623](https://doi.org/10.1093/nar/24.13.2623)
- 410 Band, G., Rockett, K. A., Spencer, C. C. A., & Kwiatkowski, D. P. (2015). A novel locus of resistance  
411 to severe malaria in a region of ancient balancing selection. *Nature*, *526*(7572), 253–257. <https://doi.org/10.1038/nature15390>
- 412
- 413 Bellenguez, C., Strange, A., Freeman, C., Donnelly, P., & Spencer, C. C. A. (2012). A robust clustering al-  
414 gorithm for identifying problematic samples in genome-wide association studies. *Bioinformatics*,  
415 *28*(1), 134–135. <https://doi.org/10.1093/bioinformatics/btr599>
- 416 Bossini-Castillo, L., Glinos, D. A., Kunowska, N., Golda, G., Lamikanra, A., Spitzer, M., Soskic, B., Cano-  
417 Gamez, E., Smyth, D. J., Cattermole, C., Alasoo, K., Mann, A., Kundu, K., Soranzo, N., Dun-  
418 ham, I., Roberts, D., & Trynka, G. (2019). Immune disease variants modulate gene expression in  
419 regulatory cd4+ t cells and inform drug targets. *bioRxiv*, <https://www.biorxiv.org/content/early/2019/05/31/654632>  
420 <https://doi.org/10.1101/654632>
- 421 Britton, W. J. (1998). The management of leprosy reversal reactions. *Lepr Rev*, *69*(3), 225–234. <https://doi.org/10.5935/0305-7518.19980024>
- 422
- 423 Britton, W. J., & Lockwood, D. N. J. (2004). Leprosy. *Lancet*, *363*(9416), 1209–1219. [https://doi.org/](https://doi.org/10.1016/S0140-6736(04)15952-7)  
424 [10.1016/S0140-6736\(04\)15952-7](https://doi.org/10.1016/S0140-6736(04)15952-7)
- 425 Chemotherapy of leprosy. report of a who study group. (1994). *World Health Organ Tech Rep Ser*, *847*,  
426 1–24.
- 427 Delaneau, O., Marchini, J., & Zagury, J.-F. (2012). A linear complexity phasing method for thousands  
428 of genomes. *Nature Methods*, *9*(2), 179–181. <https://doi.org/10.1038/nmeth.1785>

- 429 Fairfax, B. P., Humburg, P., Makino, S., Naranbhai, V., Wong, D., Lau, E., Jostins, L., Plant, K.,  
430 Andrews, R., McGee, C., & Knight, J. C. (2014). Innate immune activity conditions the effect  
431 of regulatory variants upon monocyte gene expression. *Science*, *343*(6175), 1246949. [https://](https://doi.org/10.1126/science.1246949)  
432 [doi.org/10.1126/science.1246949](https://doi.org/10.1126/science.1246949)
- 433 Fairfax, B. P., Makino, S., Radhakrishnan, J., Plant, K., Leslie, S., Dilthey, A., Ellis, P., Langford, C.,  
434 Vannberg, F. O., & Knight, J. C. (2012). Genetics of gene expression in primary immune cells  
435 identifies cell type-specific master regulators and roles of hla alleles. *Nat Genet*, *44*(5), 502–510.  
436 <https://doi.org/10.1038/ng.2205>
- 437 Fitness, J., Floyd, S., Warndorff, D. K., Sichali, L., Mwaungulu, L., Crampin, A. C., Fine, P. E. M., &  
438 Hill, A. V. S. (2004). Large-scale candidate gene study of leprosy susceptibility in the karonga  
439 district of northern malawi. *Am J Trop Med Hyg*, *71*(3), 330–340.
- 440 for South-East Asia, W. H. O. R. O. (2009). *Enhanced global strategy for further reducing the disease*  
441 *burden due to leprosy (2011-2015): Plan period: 2011-2015*. WHO Regional Office for South-  
442 East Asia.
- 443 Giambartolomei, C., Vukcevic, D., Schadt, E. E., Franke, L., Hingorani, A. D., Wallace, C., & Plagnol,  
444 V. (2014). Bayesian test for colocalisation between pairs of genetic association studies using  
445 summary statistics. *PLoS Genet*, *10*(5), e1004383. <https://doi.org/10.1371/journal.pgen.1004383>
- 446 Gilchrist, J. J., Makino, S., Naranbhai, V., Lau, E., Danielli, S., Hameiri-Bowen, D., Lee, W., Ng,  
447 E., Whalley, J., Knight, J. C., & Fairfax, B. P. (2021). Natural killer cells demonstrate dis-  
448 tinct eqtl and transcriptome-wide disease associations, highlighting their role in autoimmunity.  
449 *bioRxiv*, <https://www.biorxiv.org/content/early/2021/05/11/2021.05.10.443088.full.pdf>. [https://](https://doi.org/10.1101/2021.05.10.443088)  
450 [doi.org/10.1101/2021.05.10.443088](https://doi.org/10.1101/2021.05.10.443088)
- 451 Gurdasani, D., Carstensen, T., Tekola-Ayele, F., Pagani, L., Tachmazidou, I., Hatzikotoulas, K., Karthikeyan,  
452 S., Iles, L., Pollard, M. O., Choudhury, A., Ritchie, G. R. S., Xue, Y., Asimit, J., Nsubuga, R. N.,  
453 Young, E. H., Pomilla, C., Kivinen, K., Rockett, K., Kamali, A., . . . Sandhu, M. S. (2015). The  
454 african genome variation project shapes medical genetics in africa. *Nature*, *517*(7534), 327–332.  
455 <https://doi.org/10.1038/nature13997>
- 456 Gzara, C., Dallmann-Sauer, M., Orlova, M., Van Thuc, N., Thai, V. H., Fava, V. M., Bihoreau, M.-T.,  
457 Boland, A., Abel, L., Alcais, A., Schurr, E., & Cobat, A. (2020). Family-based genome-wide  
458 association study of leprosy in vietnam. *PLoS Pathog*, *16*(5), e1008565. [https://doi.org/10.](https://doi.org/10.1371/journal.ppat.1008565)  
459 [1371/journal.ppat.1008565](https://doi.org/10.1371/journal.ppat.1008565)
- 460 Hashimoto-Tane, A., Yokosuka, T., Sakata-Sogawa, K., Sakuma, M., Ishihara, C., Tokunaga, M., & Saito,  
461 T. (2011). Dynein-driven transport of t cell receptor microclusters regulates immune synapse  
462 formation and t cell activation. *Immunity*, *34*(6), 919–931. [https://doi.org/10.1016/j.immuni.](https://doi.org/10.1016/j.immuni.2011.05.012)  
463 [2011.05.012](https://doi.org/10.1016/j.immuni.2011.05.012)

- 464 Howie, B., Marchini, J., & Stephens, M. (2011). Genotype imputation with thousands of genomes. *G3*  
465 (*Bethesda*), 1(6), 457–470. <https://doi.org/10.1534/g3.111.001198>
- 466 Howie, B. N., Donnelly, P., & Marchini, J. (2009). A flexible and accurate genotype imputation method  
467 for the next generation of genome-wide association studies. *PLoS Genet*, 5(6), e1000529. <https://doi.org/10.1371/journal.pgen.1000529>
- 469 Johnston, H. R., Hu, Y.-J., Gao, J., O'Connor, T. D., Abecasis, G. R., Wojcik, G. L., Gignoux, C. R.,  
470 Gourraud, P.-A., Lizee, A., Hansen, M., Genuario, R., Bullis, D., Lawley, C., Kenny, E. E.,  
471 Bustamante, C., Beaty, T. H., Mathias, R. A., Barnes, K. C., & Qin, Z. S. (2017). Identifying  
472 tagging snps for african specific genetic variation from the african diaspora genome. *Sci Rep*, 7,  
473 46398. <https://doi.org/10.1038/srep46398>
- 474 Jostins, L., Ripke, S., Weersma, R. K., Duerr, R. H., McGovern, D. P., Hui, K. Y., Lee, J. C., Schumm,  
475 L. P., Sharma, Y., Anderson, C. A., Essers, J., Mitrovic, M., Ning, K., Cleynen, I., Theatre, E.,  
476 Spain, S. L., Raychaudhuri, S., Goyette, P., Wei, Z., ... Cho, J. H. (2012). Host-microbe inter-  
477 actions have shaped the genetic architecture of inflammatory bowel disease. *Nature*, 491(7422),  
478 119–124. <https://doi.org/10.1038/nature11582>
- 479 Kasela, S., Kisand, K., Tserel, L., Kaleviste, E., Remm, A., Fischer, K., Esko, T., Westra, H.-J., Fairfax,  
480 B. P., Makino, S., Knight, J. C., Franke, L., Metspalu, A., Peterson, P., & Milani, L. (2017).  
481 Pathogenic implications for autoimmune mechanisms derived by comparative eqtl analysis of  
482 cd4+ versus cd8+ t cells. *PLoS Genet*, 13(3), e1006643. <https://doi.org/10.1371/journal.pgen.1006643>
- 484 Liu, H., Irwanto, A., Fu, X., Yu, G., Yu, Y., Sun, Y., Wang, C., Wang, Z., Okada, Y., Low, H., Li, Y.,  
485 Liang, H., Chen, M., Bao, F., Li, J., You, J., Zhang, Q., Liu, J., Chu, T., ... Zhang, F. (2015).  
486 Discovery of six new susceptibility loci and analysis of pleiotropic effects in leprosy. *Nat Genet*,  
487 47(3), 267–271. <https://doi.org/10.1038/ng.3212>
- 488 Liu, H., Wang, Z., Li, Y., Yu, G., Fu, X., Wang, C., Liu, W., Yu, Y., Bao, F., Irwanto, A., Liu, J., Chu,  
489 T., Andiappan, A. K., Maurer-Stroh, S., Limviphuvadh, V., Wang, H., Mi, Z., Sun, Y., Sun,  
490 L., ... Zhang, F. (2017). Genome-wide analysis of protein-coding variants in leprosy. *J Invest*  
491 *Dermatol*, 137(12), 2544–2551. <https://doi.org/10.1016/j.jid.2017.08.004>
- 492 Malaria Genomic Epidemiology, N. (2019). Insights into malaria susceptibility using genome-wide data  
493 on 17,000 individuals from africa, asia and oceania. *Nat Commun*, 10(1), 5732. <https://doi.org/10.1038/s41467-019-13480-z>
- 495 Marchini, J., Howie, B., Myers, S., McVean, G., & Donnelly, P. (2007). A new multipoint method for  
496 genome-wide association studies by imputation of genotypes. *Nat Genet*, 39(7), 906–913. <https://doi.org/10.1038/ng2088>

- 498 Maurano, M. T., Humbert, R., Rynes, E., Thurman, R. E., Haugen, E., Wang, H., Reynolds, A. P., Sand-  
499 strom, R., Qu, H., Brody, J., Shafer, A., Neri, F., Lee, K., Kutayavin, T., Stehling-Sun, S., John-  
500 son, A. K., Canfield, T. K., Giste, E., Diegel, M., . . . Stamatoyannopoulos, J. A. (2012). System-  
501 atic localization of common disease-associated variation in regulatory dna. *Science*, *337*(6099),  
502 1190–1195. <https://doi.org/10.1126/science.1222794>
- 503 Meisner, S. J., Mucklow, S., Warner, G., Sow, S. O., Lienhardt, C., & Hill, A. V. (2001). Association of  
504 nramp1 polymorphism with leprosy type but not susceptibility to leprosy per se in west africans.  
505 *Am J Trop Med Hyg*, *65*(6), 733–735. <https://doi.org/10.4269/ajtmh.2001.65.733>
- 506 Monot, M., Honoré, N., Garnier, T., Zidane, N., Sherafi, D., Paniz-Mondolfi, A., Matsuoka, M., Taylor,  
507 G. M., Donoghue, H. D., Bouwman, A., Mays, S., Watson, C., Lockwood, D., Khamesipour, A.,  
508 Dowlati, Y., Jianping, S., Rea, T. H., Vera-Cabrera, L., Stefani, M. M., . . . Cole, S. T. (2009).  
509 Comparative genomic and phylogeographic analysis of mycobacterium leprae. *Nat Genet*, *41*(12),  
510 1282–1289. <https://doi.org/10.1038/ng.477>
- 511 Motyer, A., Vukcevic, D., Dilthey, A., Donnelly, P., McVean, G., & Leslie, S. (2016). Practical use of meth-  
512 ods for imputation of hla alleles from snp genotype data. *bioRxiv*, <https://www.biorxiv.org/content/early/2016/12/12>  
513 <https://doi.org/10.1101/091009>
- 514 Naranbhai, V., Fairfax, B. P., Makino, S., Humburg, P., Wong, D., Ng, E., Hill, A. V. S., & Knight, J. C.  
515 (2015). Genomic modulators of gene expression in human neutrophils. *Nat Commun*, *6*, 7545.  
516 <https://doi.org/10.1038/ncomms8545>
- 517 Nath, S., Christian, L., Tan, S. Y., Ki, S., Ehrlich, L. I. R., & Poenie, M. (2016). Dynein separately  
518 partners with ndel and dynactin to orchestrate t cell focused secretion. *J Immunol*, *197*(6),  
519 2090–2101. <https://doi.org/10.4049/jimmunol.1600180>
- 520 Ponnighaus, J. M., Fine, P. E., & Bliss, L. (1987). Certainty levels in the diagnosis of leprosy. *Int J Lepr*  
521 *Other Mycobact Dis*, *55*(3), 454–462.
- 522 Pönnighaus, J. M., Fine, P. E., Sterne, J. A., Wilson, R. J., Msosa, E., Gruer, P. J., Jenkins, P. A., Lucas,  
523 S. B., Liomba, N. G., & Bliss, L. (1992). Efficacy of bcg vaccine against leprosy and tuberculosis  
524 in northern malawi. *Lancet*, *339*(8794), 636–639. [https://doi.org/10.1016/0140-6736\(92\)90794-4](https://doi.org/10.1016/0140-6736(92)90794-4)
- 525 Ponnighaus, J. M., Fine, P. E., Bliss, L., Sliney, I. J., Bradley, D. J., & Rees, R. J. (1987). The lepra  
526 evaluation project (lep), an epidemiological study of leprosy in northern malaŵi. i. methods.  
527 *Lepr Rev*, *58*(4), 359–375. <https://doi.org/10.5935/0305-7518.19870038>
- 528 Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006).  
529 Principal components analysis corrects for stratification in genome-wide association studies. *Nat*  
530 *Genet*, *38*(8), 904–909. <https://doi.org/10.1038/ng1847>
- 531 Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar,  
532 P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). Plink: A tool set for whole-genome

- 533 association and population-based linkage analyses. *Am J Hum Genet*, 81(3), 559–575. <https://doi.org/10.1086/519795>
- 534
- 535 Roberts, A. J., Kon, T., Knight, P. J., Sutoh, K., & Burgess, S. A. (2013). Functions and mechanics  
536 of dynein motor proteins. *Nat Rev Mol Cell Biol*, 14(11), 713–726. <https://doi.org/10.1038/nrm3667>
- 537
- 538 Shrum, C. K., Defrancisco, D., & Meffert, M. K. (2009). Stimulated nuclear translocation of nf-kappab  
539 and shuttling differentially depend on dynein and the dynactin complex. *Proc Natl Acad Sci U*  
540 *S A*, 106(8), 2647–2652. <https://doi.org/10.1073/pnas.0806677106>
- 541 Smith, W. C., van Brakel, W., Gillis, T., Saunderson, P., & Richardus, J. H. (2015). The missing millions:  
542 A threat to the elimination of leprosy. *PLoS Negl Trop Dis*, 9(4), e0003658. <https://doi.org/10.1371/journal.pntd.0003658>
- 543
- 544 Sun, Y., Irwanto, A., Toyo-Oka, L., Hong, M., Liu, H., Andiappan, A. K., Choi, H., Hitomi, Y., Yu, G.,  
545 Yu, Y., Bao, F., Wang, C., Fu, X., Yue, Z., Wang, H., Zhang, H., Kawashima, M., Kojima, K.,  
546 Nagasaki, M., . . . Liu, J. (2016). Fine-mapping analysis revealed complex pleiotropic effect and  
547 tissue-specific regulatory mechanism of tnfsf15 in primary biliary cholangitis, crohn’s disease and  
548 leprosy. *Sci Rep*, 6, 31429. <https://doi.org/10.1038/srep31429>
- 549 Toure, O., Konate, S., Sissoko, S., Niangaly, A., Barry, A., Sall, A. H., Diarra, E., Poudiougou, B.,  
550 Sepulveda, N., Campino, S., Rockett, K. A., Clark, T. G., Thera, M. A., & Doumbo, O. (2012).  
551 Candidate polymorphisms and severe malaria in a malian population. *PLoS One*, 7(9), e43987.  
552 <https://doi.org/10.1371/journal.pone.0043987>
- 553 Urnavicius, L., Zhang, K., Diamant, A. G., Motz, C., Schlager, M. A., Yu, M., Patel, N. A., Robinson,  
554 C. V., & Carter, A. P. (2015). The structure of the dynactin complex and its interaction with  
555 dynein. *Science*, 347(6229), 1441–1446. <https://doi.org/10.1126/science.aaa4080>
- 556 Wakefield, J. (2009). Bayes factors for genome-wide association studies: Comparison with p-values. *Genet*  
557 *Epidemiol*, 33(1), 79–86. <https://doi.org/10.1002/gepi.20359>
- 558 Wallace, C., Fitness, J., Hennig, B., Sichali, L., Mwaungulu, L., Pönnighaus, J. M., Warndorff, D. K.,  
559 Clayton, D., Fine, P. E. M., & Hill, A. V. S. (2004). Linkage analysis of susceptibility to leprosy  
560 type using an ibd regression method. *Genes Immun*, 5(3), 221–225. <https://doi.org/10.1038/sj.gene.6364062>
- 561
- 562 Wang, Z., Mi, Z., Wang, H., Sun, L., Yu, G., Fu, X., Wang, C., Bao, F., Yue, Z., Zhao, Q., Wang, N.,  
563 Cheng, X., Liu, H., & Zhang, F. (2018). Discovery of 4 exonic and 1 intergenic novel susceptibility  
564 loci for leprosy. *Clin Genet*, 94(2), 259–263. <https://doi.org/10.1111/cge.13376>
- 565 Wang, Z., Sun, Y., Fu, X., Yu, G., Wang, C., Bao, F., Yue, Z., Li, J., Sun, L., Irwanto, A., Yu, Y.,  
566 Chen, M., Mi, Z., Wang, H., Huai, P., Li, Y., Du, T., Yu, W., Xia, Y., . . . Zhang, F. (2016).

- 567 A large-scale genome-wide association and meta-analysis identified four novel susceptibility loci  
568 for leprosy. *Nat Commun*, 7, 13760. <https://doi.org/10.1038/ncomms13760>
- 569 WHO Working Group on Leprosy Control. Meeting (1st: 1991 : Geneva, S., & Unit, W. H. O. L. (1991).  
570 Report of the first meeting of the who working group on leprosy control, geneva, 1-3 july 1991.  
571 World Health Organization.
- 572 Wong, S. H., Gochhait, S., Malhotra, D., Pettersson, F. H., Teo, Y. Y., Khor, C. C., Rautanen, A.,  
573 Chapman, S. J., Mills, T. C., Srivastava, A., Rudko, A., Freidin, M. B., Puzyrev, V. P., Ali, S.,  
574 Aggarwal, S., Chopra, R., Reddy, B. S. N., Garg, V. K., Roy, S., . . . Vannberg, F. O. (2010).  
575 Leprosy and the adaptation of human toll-like receptor 1. *PLoS Pathog*, 6(7), e1000979. <https://doi.org/10.1371/journal.ppat.1000979>
- 577 Yamamura, M., Wang, X. H., Ohmen, J. D., Uyemura, K., Rea, T. H., Bloom, B. R., & Modlin, R. L.  
578 (1992). Cytokine patterns of immunologically mediated tissue damage. *J Immunol*, 149(4), 1470–  
579 1475.
- 580 Zhang, F.-R., Huang, W., Chen, S.-M., Sun, L.-D., Liu, H., Li, Y., Cui, Y., Yan, X.-X., Yang, H.-T.,  
581 Yang, R.-D., Chu, T.-S., Zhang, C., Zhang, L., Han, J.-W., Yu, G.-Q., Quan, C., Yu, Y.-X.,  
582 Zhang, Z., Shi, B.-Q., . . . Liu, J.-J. (2009). Genomewide association study of leprosy. *N Engl J*  
583 *Med*, 361(27), 2609–2618. <https://doi.org/10.1056/NEJMoa0903753>
- 584 Zhang, F., Liu, H., Chen, S., Low, H., Sun, L., Cui, Y., Chu, T., Li, Y., Fu, X., Yu, Y., Yu, G., Shi, B.,  
585 Tian, H., Liu, D., Yu, X., Li, J., Lu, N., Bao, F., Yuan, C., . . . Zhang, X. (2011). Identification  
586 of two new loci at il23r and rab32 that influence susceptibility to leprosy. *Nat Genet*, 43(12),  
587 1247–1251. <https://doi.org/10.1038/ng.973>

## 588 Supplementary Figures

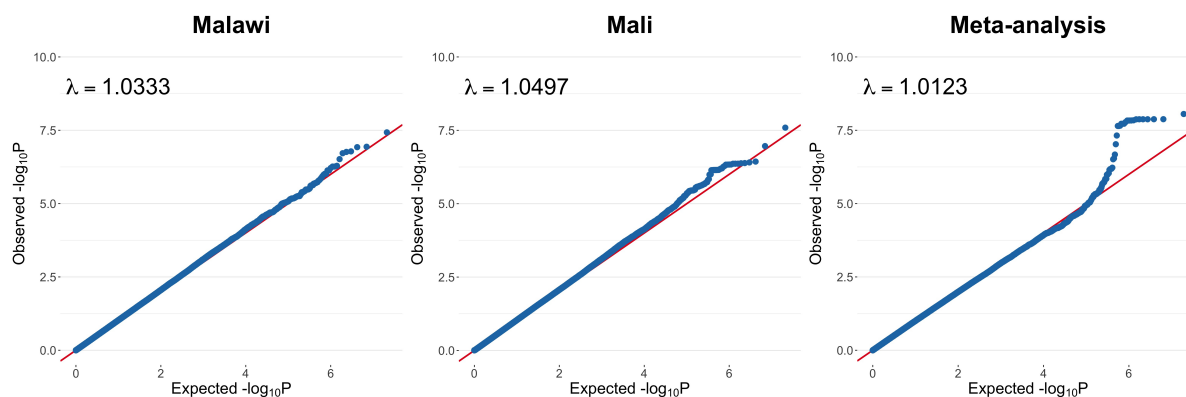


Figure S1: **Quantile-quantile plots of leprosy association.**

QQ-plots of leprosy association in Malawi (284 cases, 328 controls, SNPs = 10,511,695), Mali (208 cases, 311 controls, SNPs = 10,514,676) and fixed-effects meta-analysis of both populations (cases = 492, controls = 639, SNPs = 9,616,523).

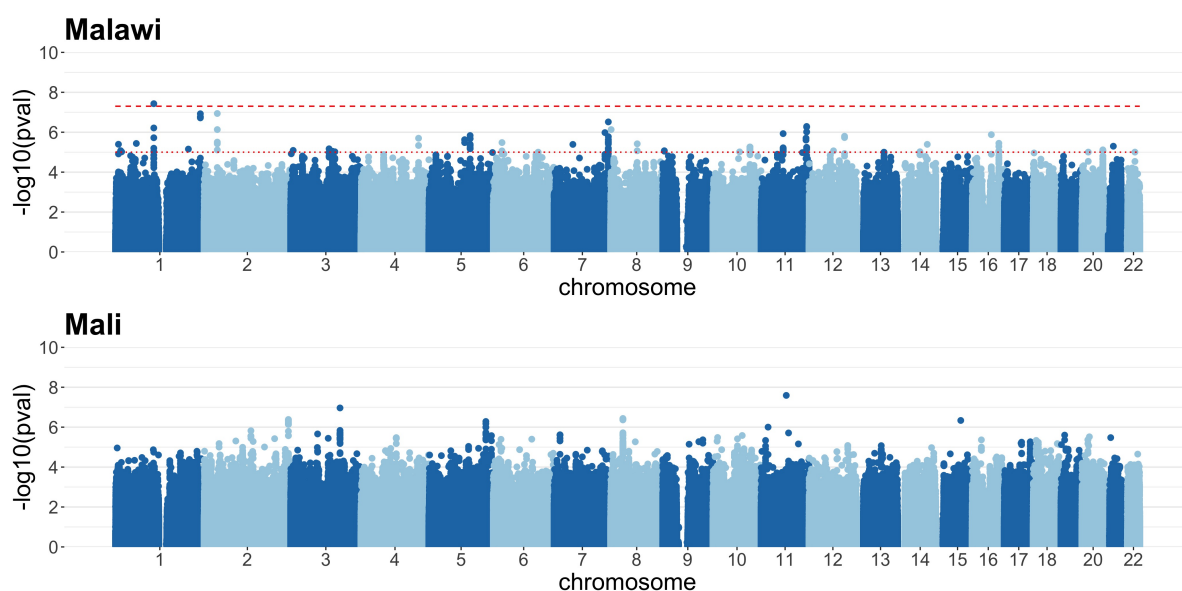


Figure S2: **Manhattan plots of leprosy association.**

Manhattan plots of leprosy association in discovery (Malawi, 284 cases, 328 controls, SNPs = 10,511,695) and replication (Mali, 208 cases, 311 controls, SNPs = 10,514,676) samples. P-value thresholds are annotated on the Malawi Manhattan plot: dashed line,  $p = 1 \times 10^{-8}$  (genome-wide significance); dotted line,  $p = 1 \times 10^{-5}$  (threshold for suggestive association).

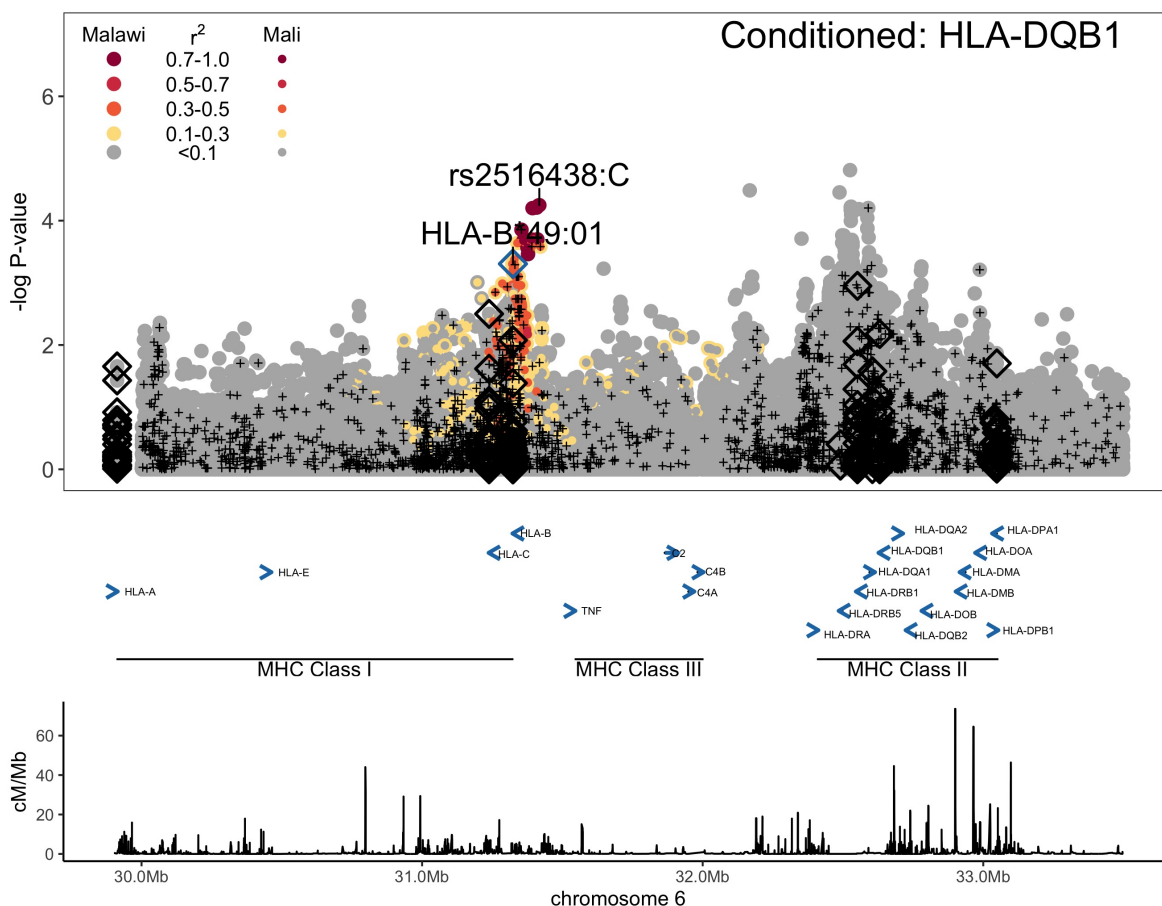


Figure S3: **HLA Leprosy association conditioned on HLA-DQB1\*04:02.**

Association statistics represent a fixed-effects meta-analysis of additive association with disease in Malawi and Mali conditioned on HLA-DQB1\*04:02. SNPs are coloured according to linkage disequilibrium to rs2516438, and genotyped SNPs marked with black pluses. Imputed classical HLA alleles are plotted as diamonds, with significantly associated (FDR < 0.05) alleles highlighted in blue.



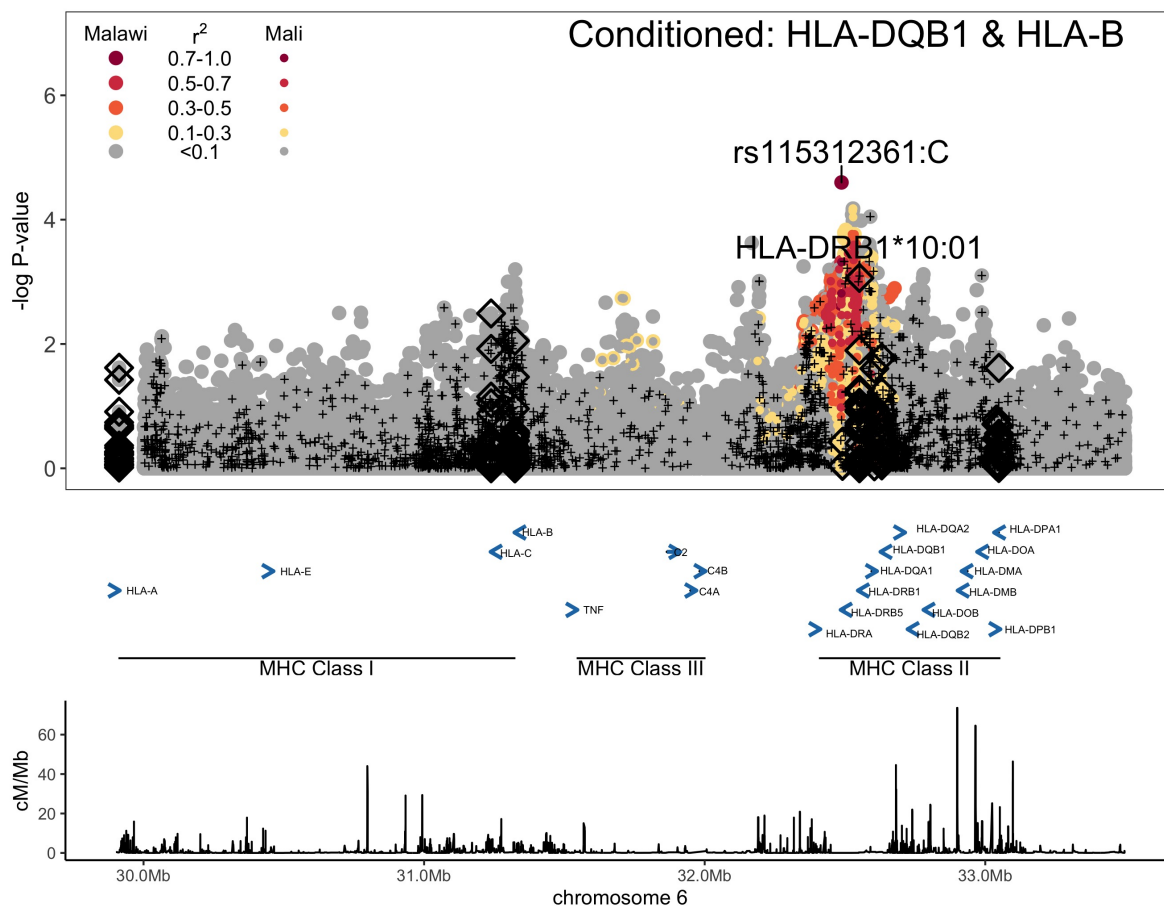


Figure S4: **HLA Leprosy association conditioned on HLA-DQB1\*04:02 and HLA-B\*49:01.** Association statistics represent a fixed-effects meta-analysis of additive association with disease in Malawi and Mali conditioned on HLA-DQB1\*04:02 and HLA-B\*49:01. SNPs are coloured according to linkage disequilibrium to rs115312361, and genotyped SNPs marked with black plusses. Imputed classical HLA alleles are plotted as diamonds. No significantly associated (FDR < 0.05) alleles remain after conditioning on HLA-DQB1\*04:02 and HLA-B\*49:01.

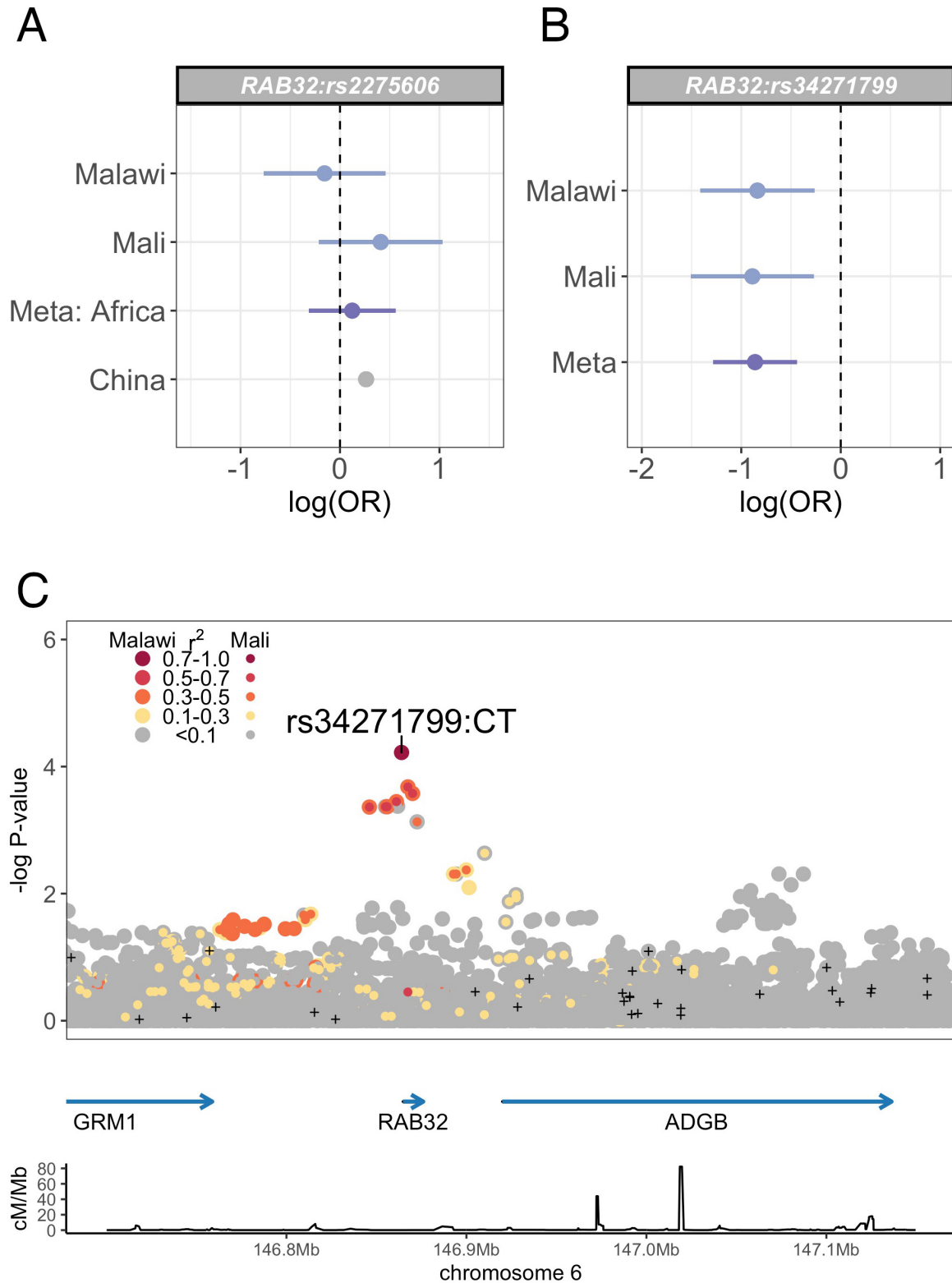


Figure S5: **Evidence for leprosy association at the RAB32 locus.**

(A) Log-transformed odds ratios and 95% confidence intervals of rs2275606 association (peak association in Chinese GWAS data) with leprosy in Malawi, Mali and China. (B) Log-transformed odds ratios and 95% confidence intervals of rs34271799 association with leprosy in Malawi and Mali. (C) Regional association plot of leprosy association at the RAB32 locus. Association statistics represent a fixed-effects meta-analysis of additive association with disease in Malawi and Mali. SNPs are coloured according to linkage disequilibrium to rs34271799, and genotyped SNPs marked with black pluses.

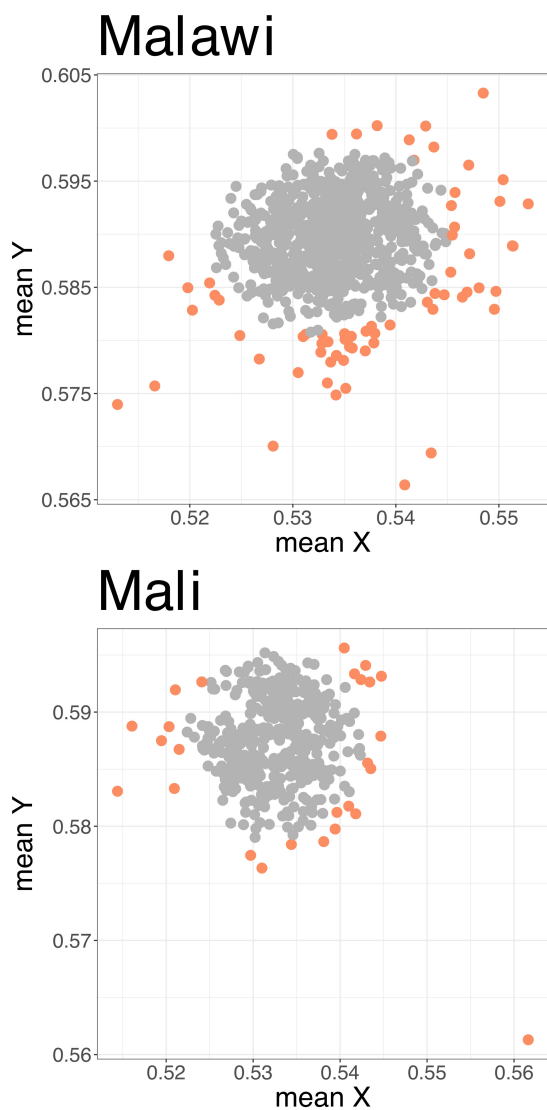


Figure S6: **Sample X and Y channel intensities.**

(A) Mean X and Y channel intensities for Malawi (top) and Mali (bottom) samples. Outlying samples were identified using ABERRANT and are highlighted (orange).

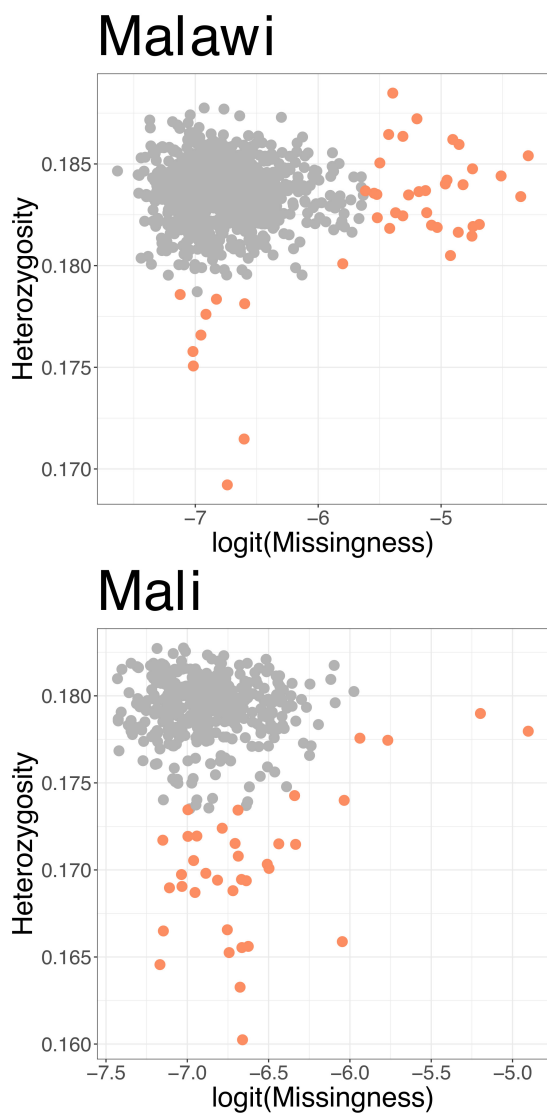


Figure S7: **Sample missingness and heterozygosity.**

(A) Mean sample genotype missingness plotted against heterozygosity for Malawi (top) and Mali (bottom) samples. Outlying samples were identified using ABERRANT and are highlighted (orange).

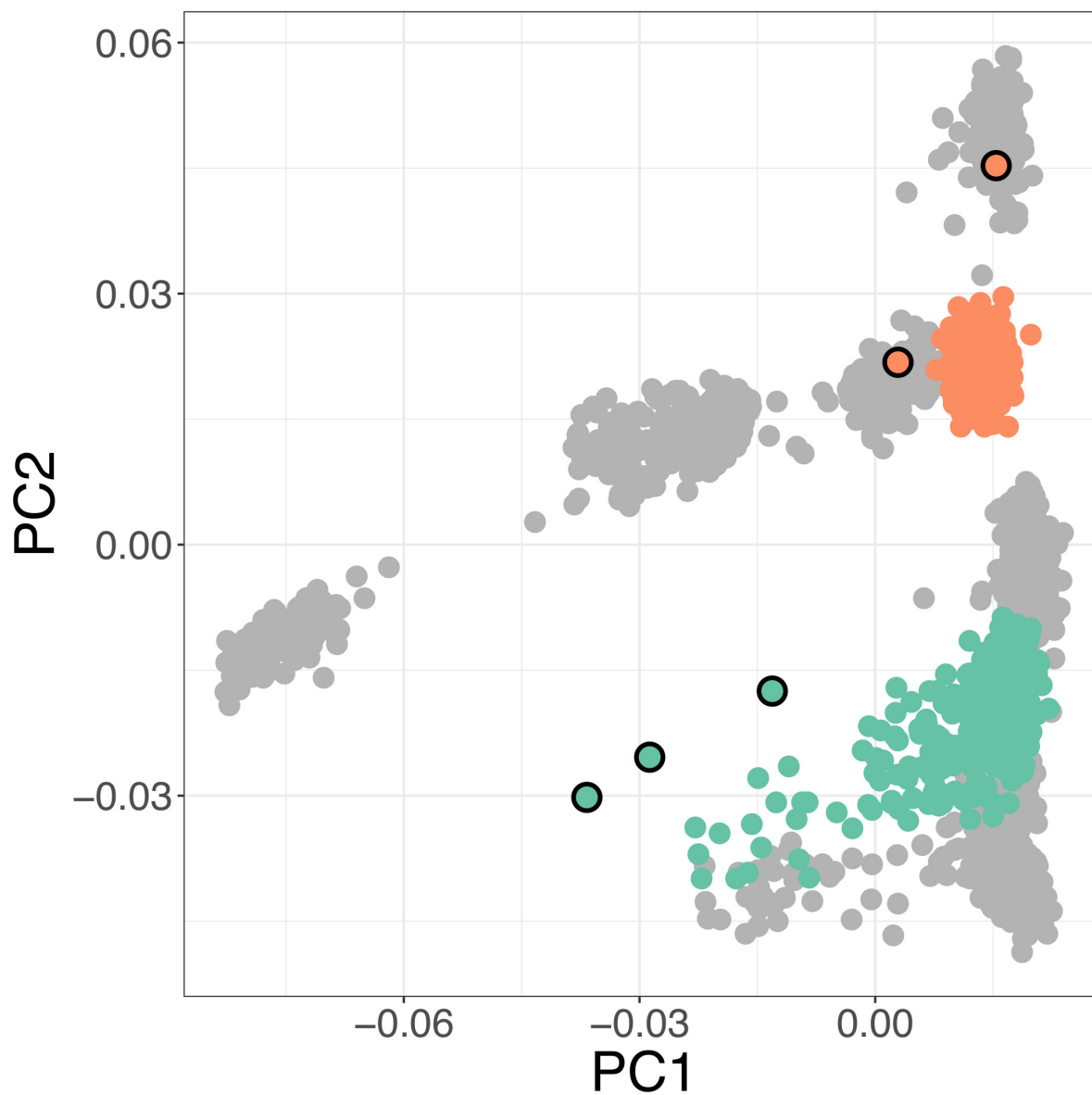
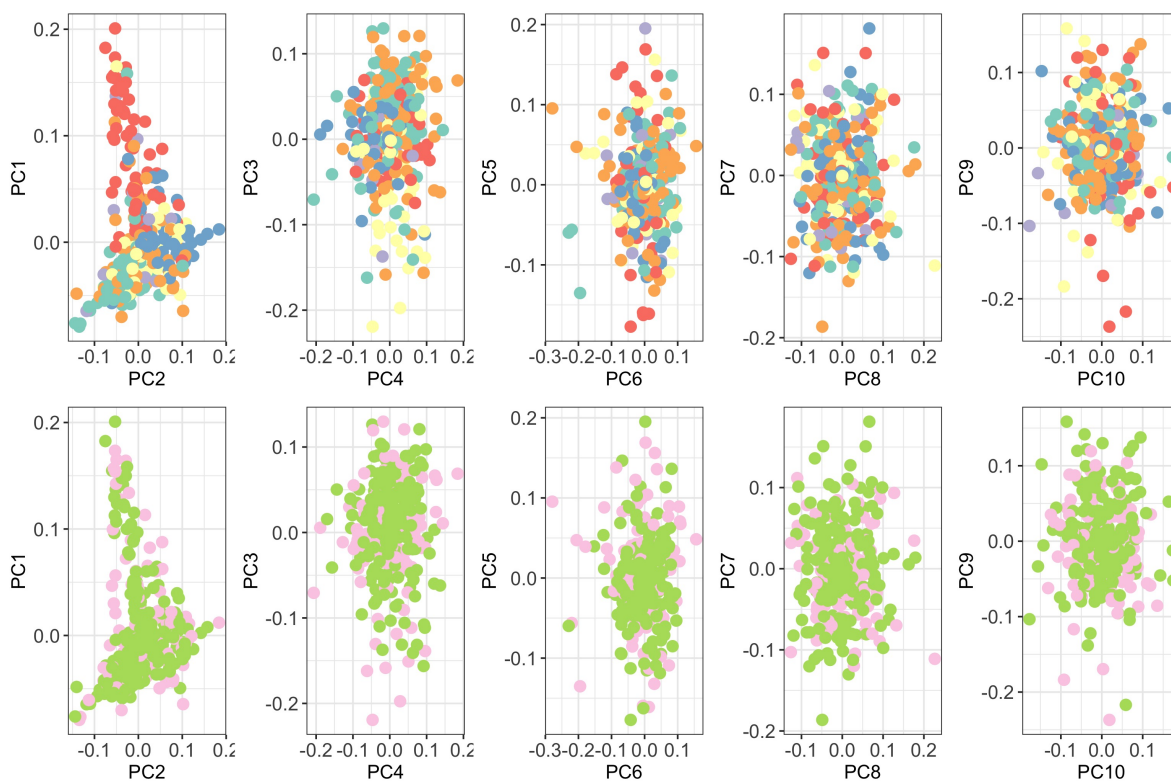


Figure S8: **Population outliers.**

Plot of the major two principal components of genome wide genotyping data. Malawi study samples are plotted in orange and Mali study samples in green, against a background of African Genome Variation Project samples (gray). Outliers are highlighted (black rings).



**Figure S9: Principal components of Malawian genome-wide genotyping data.**

Individuals are color-coded according to self-reported ethnicity (top) and case-control status; cases in pink, controls in green (bottom).

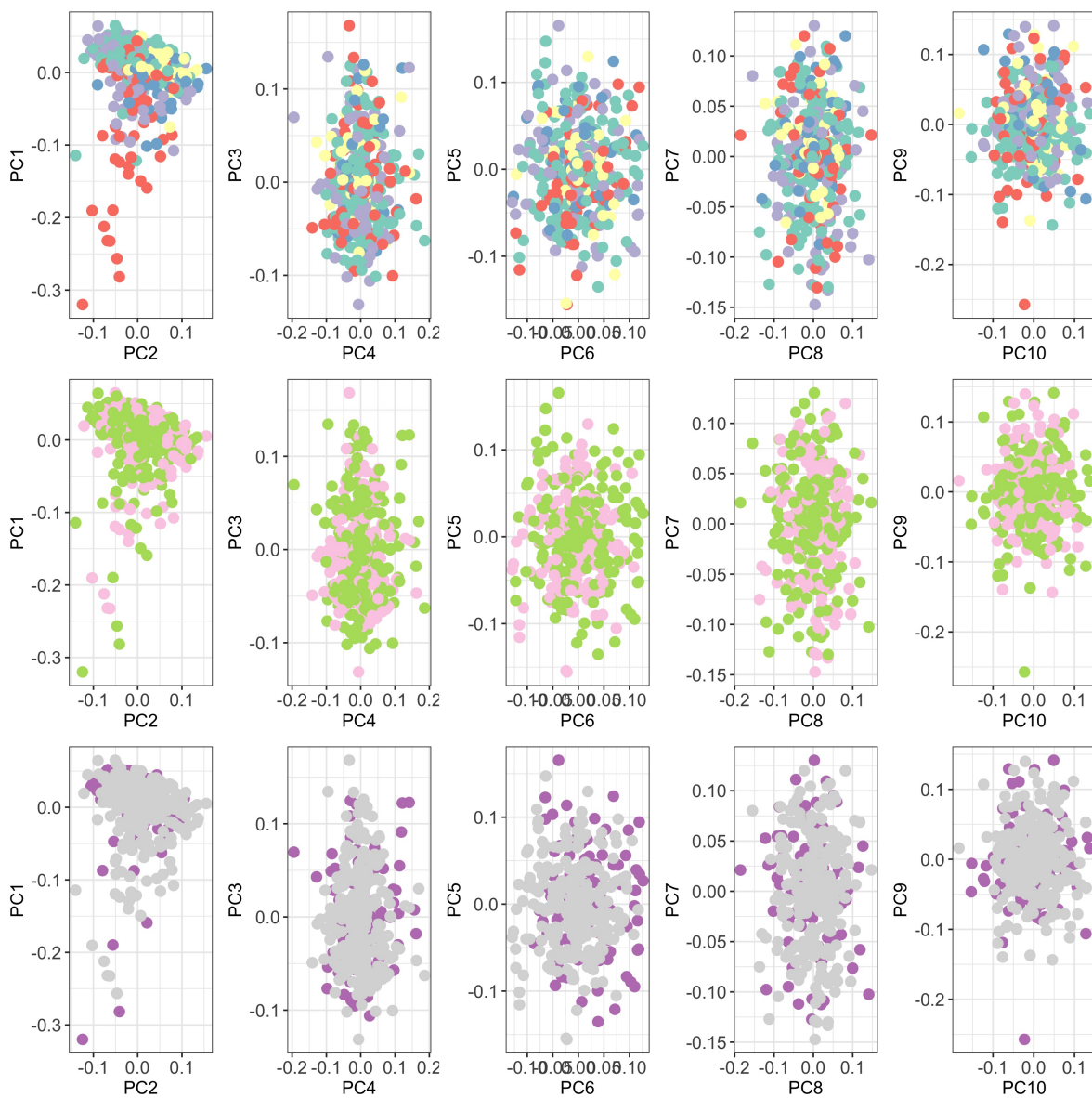


Figure S10: **Principal components of Malian genome-wide genotyping data.**

Individuals are color-coded according to self-reported ethnicity (top), case-control status (middle; cases in pink, controls in green), and genotyping platform (bottom; Omni 2.5M in purple, Africa Diaspora Power Chip in gray).