Helminth infections and allergies in Ghana

Amoah, A.S.

chapter 6

Urban-rural differences in the gene expression profiles of Ghanaian children

Abena S. Amoah[1,2], Benedicta B. Obeng[1,2], Linda May[2] , Yvonne C. Kruize[2], Irene A. Larbi[1], Michael Kabesch[3], Michael D. Wilson[1], Franca C. Hartgers[2] , Daniel A. Boakye[1] and Maria Yazdanbakhsh[2]

Affiliations:

[1] Department of Parasitology, Noguchi Memorial Institute for Medical Research, Accra, Ghana
[2] Department of Parasitology, Leiden University Medical Center, Leiden, the Netherlands;
[3] Department of Pediatric Pneumology and Allergy, KUNO University
Children's Hospital Regensburg, Regensburg, Germany

- Genes and Immunity 2014; 15(5):313-9-

Abstract

Recent studies indicate that urbanization is having a pronounced effect on disease patterns in developing countries. To understand the immunological basis of this, we examined mRNA expression in whole blood of genes involved in immune activation and regulation in 151 children aged 5-13 years attending rural, urban low socioeconomic status (SES) and urban high SES schools in Ghana. Samples were also collected to detect helminth and malaria infections.

Marked differences in gene expression were observed between the rural and urban areas as well as within the urban area. The expression of both interleukin (IL)-10 and programmed cell death protein 1 (PD-1) increased significantly across the schools from urban high SES to urban low SES to rural (p-trend<0.001). Although IL-10 gene expression was significantly elevated in the rural compared to the urban schools (p <0.001), this was not associated with parasitic infection. Significant differences in the expression of Toll-like receptors (TLRs) and their signalling genes were seen between the two urban schools. Genetic differences could not fully account for the gene expression profiles in the different groups as shown by analysis of *IL-10*, *TLR-2* and *TLR-4* gene polymorphisms.

Immune gene expression patterns are strongly influenced by environmental determinants and may underlie the effects of urbanization seen on health outcomes.

Key words

Gene expression, polymorphisms, urbanization, Interleukin-10, Toll-like Receptors, helminths,

Introduction

Urbanization worldwide, particularly in developing nations, is changing mortality and morbidity patterns from largely infectious to non-communicable diseases [1]. Although governed by genetics, the increase in non-communicable (often inflammatory) diseases is thought to be driven by changes in environmental factors [2]. Moreover, in rapidly expanding urban centres, socioeconomic differences are resulting in heterogeneous environmental exposures that are determining disease patterns. For instance, some urban areas are characterized by high population density, overcrowding, limited access to potable water and poor sanitation [3], all of which increase exposure to pathogens [4]. Conversely, other urban environments are more affluent with wealthier inhabitants having greater access to clean water, food and adequate sanitation but at the same time being more susceptible to chronic conditions such as hypertension [5], obesity [6] and cardiovascular diseases [7]. Further along this spectrum, rural environments in developing countries remain largely agrarian with lifestyles characterized by traditional diet [8], limited access to health-care [9] and in many areas, continual exposure to pathogens.

Understanding the factors associated with changing environments and the link to the alteration of the immune system would be important for both communicable and non-communicable disease prevention strategies. Although risk factors associated with the rural to urban transition, particularly in relation to inflammatory diseases, have been studied extensively [10,11], little is known about actual changes that take place in the immune system as a function of the rural-urban gradient.

Examining gene expression patterns is one approach towards dissecting differences in immune responses between urban and rural populations since variability in gene expression is a result of not only genetic but also environmental factors [12]. Moreover, differential gene expression can be a key mechanism in disease manifestation [13]. Although many studies have addressed the genetic determinants of gene expression, few have examined the impact of geographical location in generating transcriptional variation. A study conducted among the genetically homogenous Amazighs of Morocco living in three geographically distinct areas demonstrated that locality can have a dominant impact on gene expression profiles with up to one third of the leukocyte transcriptome being associated with geographical area differences [14, 15]. However, this study did not explore the possible factors within the distinct geographical locations that may account for observed profile differences. The high burden of parasitic infection in many rural areas, in addition to differences in socioeconomic status (SES) within urban areas, can result in very different exposures to microorganisms and thus lead to differences in transcriptomal profiles.

We investigated whether contrasting geographical locations in one region of Ghana with a large urban centre and rural areas that are endemic for parasitic infections [16] have an impact on messenger RNA (mRNA) expression of selected immune genes among children. Our target study locations were a rural area, an urban low SES area and an urban high SES area.

Of particular interest were genes involved in immune activation and regulation particularly in response to parasites including helminths. Given the role of Toll-like receptors (TLRs) in the recognition of pathogen-associated molecular patterns (PAMPs) linked to microbial infection [17], various genes involved in the TLR signalling pathway were part of our selection. In environments typified by chronic helminth infections, key factors involved in the immune regulation of helminth infections such as the regulatory cytokines IL-10 and TGF-β [18] were part of our selection. In addition, immune markers involved in T-cell activation and polarization such as FOXP3 were included since CD4+CD25+FOXP3+ regulatory cells form a key population of regulatory cells involved in infections in general [19]. The gene for immunoglobulin E (IgE) antibody, which is strongly associated with immune responses to helminths, was included to compare with IgE antibody levels in circulation and thus act as a control. We hypothesized that the expression of genes involved in immune regulatory processes associated with helminth infections would be high in the rural area followed by the urban low SES and then urban high SES areas.

Methods

Study area

Study participants were recruited from 3 schools located in distinct geographical locations of the Greater Accra region of Ghana. This region is the second most populous in Ghana and is situated in the south-east of the country. The 3 schools were selected to reflect the dynamic environmental changes associated with urbanization in Ghana.

The rural school was located in Ayikai Doblo, a community in the Ga West district which is approximately 20-30 km north of Accra City Centre. The vast majority of people in Ayikai Doblo are of the Ga-Adangme ethnic group. The main income earning activities in this community are farming, trading and commercial sand collection. The area remains endemic for the waterborne helminth infection *S. haematobium* [16]. The urban low SES school was located in Jamestown a coastal community in the city of Accra and is inhabited predominantly by indigenous Ga-Adangme people. Jamestown can best be described as a "large high-density low-income formal settlement" [20] and is characterized by overcrowding as well as poor sanitation. The main economic activities in this area revolve around fishing and petty trading. The urban high SES school was situated on the University of Ghana campus at Legon and can be classified as middle-to-high income with the majority of those attending this school being the children of faculty and employees of the university. The school is ethnically diverse with not only Ga-Adangmes but also other Ghanaian ethnic groups.

Study Population

The study population consisted of schoolchildren aged between 5 - 13 years randomly selected from a larger investigation into immune responses, parasitic infections and atopic sensitization in Ghanaian children [21].

RNA isolation from whole blood

For each study subject, blood was drawn into a heparinized tube and immediately following venipuncture, 0.8 ml of whole blood was added to 3.6 ml of Nuclisens lysis buffer (Biomérieux, Boxtel, The Netherlands) to stabilize RNA. Samples were stored for a maximum of 2 weeks at 4°C after which they were transferred to -80 °C for long-term storage. Detailed RNA isolation methodology has been described previously [21]. Briefly, a Nuclisens isolation kit (Biomérieux, Boxtel, The Netherlands) was used for the isolation of total nucleic acid according to manufacturer's instructions. Prior to the isolation, the samples were treated with RNAse-free DNAse (Invitrogen, Breda, The Netherlands) to remove genomic DNA.

cDNA synthesis and Real-time PCR

The cDNA synthesis and real-time PCR methodology followed has been described in detail by Hartgers *et al.* [21] Briefly, reverse transcription of RNA was carried out using moloney murine leukaemia virus reverse transcriptase (Invitrogen, Breda, The Netherlands). Gene expression was determined by real-time quantitative PCR using the ABI PRISM 7500 system (Applied Biosystems, Foster City, California, USA). PCR reactions were performed in duplicate according to Taqman™ assay instructions using Taqman probes and qPCR Core kit reagents (Eurogentec, Seraing, Belgium).

Normalization of gene expression was done using the housekeeping gene 18S rRNA. Following the normalization procedure, the donor with the lowest expression was set to 1. Expression levels for other donors for each gene were determined relative to this donor. A description of genes examined is shown in Table S1 of the supplementary material.

Genotyping

Polymorphisms of *IL-10, TLR-2* and *TLR-4* genes were selected on the basis of a larger investigation to establish whether genetic variants associated with allergic phenotypes in developed countries were of relevance in Ghana. Therefore, variants associated with allergy phenotypes in European populations were targeted.

For IL-10, tagging SNPs from the IL-10 gene region were selected based on genotype data available through the HapMap project (www.hapmap.org). The HapMap reference population for the tagging selection was the CEPH (Utah Residents with Northern and Western European Ancestry). In addition, SNPs in close linkage disequilibrium with the tagging SNPs as well as an additional IL-10 promoter SNP of functional importance in IL-10 cytokine production (rs10494879) were included. The linkage disequilibrium plot for the IL-10 SNPs genotyped is shown in Figure S5 (Supplementary material).

TLR-2 and *TLR-4* polymorphisms were selected from the 10 human *TLR* genes using the Innate Immunity Program for Genomic Applications mutation screen [22]. Polymorphisms with a minor allele frequency ≥ 0.03 associated with amino acid

changes as well as SNPs with a minor allele frequency > 0.1 associated with altered transcription factor binding in *TLR* regulatory regions were selected [22].

All SNPS were genotyped by matrix-assisted laser desorption / ionization time-of-flight mass spectrometry using the MassARRAY system (Sequenom Inc, San Diego, California, USA) as has been described in detail elsewhere [22, 23]. A total of 330 samples from our study population were genotyped successfully for the IL-10 gene and 318 for the *TLR-2* and *TLR-4* genes. Genotype frequencies in the study population were examined for deviation from Hardy-Weinberg Equilibrium (HWE) as part of quality control for the genotyping process. MAFs for the SNPs genotyped were also compared to Yoruba in Ibadan, Nigeria genotype data from HapMap (www.hapmap.org) as well as to studies on genetic variation and inflammatory responses conducted in Northern Ghana [24, 25] (see Table S3 and S6 - Supplementary material).

Total IgE and CRP

Serum levels of total IgE as well as CRP were assessed for each participant. Total IgE was measured by enzyme linked immunosorbent assay as described in detail elsewhere [26]. The concentration of CRP in serum samples was determined by immunoturbidimetric assay using the automated P-800 system (Hitachi, Tokyo, Japan). Detailed methodology has been described elsewhere [27].

Detection of parasitic infections

Helminths

Stool samples were collected for the detection of intestinal helminth eggs by the Kato-Katz technique using 25 mg of stool. A urine sample was also collected to determine *S. haematobium* infection using the standard filtration method in which 10ml of urine is filtered through a nylon nucleopore filter (pore size, 10 µm) in a Swin-lok filtration device (Whatman,'s-Hertogenbosch, Netherlands). Helminth eggs were detected by microscopy.

Malaria detection

For each subject, a small quantity of blood was used to prepare a Giemsa-stained thick smear slide to detect malaria parasites by microscopy.

Statistical analyses

Area differences in the distribution of subject characteristics were examined by Pearson's χ^2 tests for categorical variables and Mann-Whitney U test for continuous variables. A p-value less than 0.05 was taken as the level for statistical significance. Messenger RNA expression levels were not normally distributed and so were log-transformed (base 10). Z-scores [(individual level – mean level) / standard deviation] were generated on log-transformed mRNA data. Analysis of Covariance (ANCOVA) models were used to examine the association between area and mRNA expression levels adjusting for age

and gender as *a priori* confounders. The Bonferroni correction was applied for between area pair-wise comparisons. We also used ANCOVA models to examine variations in mRNA expression according to helminth infections and malaria infection.

Haploview software package [28] was used to estimate minor allele frequencies of *IL-10, TLR-2* and *TLR-4* gene polymorphisms. Deviations from HWE were tested by χ^2 tests. Area differences in the SNP MAFs were also examined. Linear regression models were generated to examine the relationship between individual *IL-10* SNPs and IL-10 mRNA expression correcting for age, gender and area assuming an additive model. The same was done for the effects of *TLR-2* and *TLR-4* SNPs on TLR-2 and TLR-4 mRNA expression levels respectively.

Statistical analysis was performed using IBM SPSS version 20.0 (SPSS Inc., Chicago, IL, USA). GraphPad Prism version 5.00 for Windows (GraphPad Software, San Diego California USA) was used to generate figures.

Informed consent and ethical approval

The parents / guardians of all study participants were given informed consent forms to sign or thumbprint if they wished to enrol their wards in the study. They were also provided with study information sheets which were explained verbally at Parent Teacher Association (PTA) meetings held at the urban low SES and rural schools. Ethical approval for this project was granted by the Noguchi Memorial Institute for Medical Research Institutional Review Board (approval number CPN015/02-03).

Results

Characteristics of study participants

Whole blood samples were collected from 151 children for gene expression profiling. Table 1 shows the characteristics of the study participants. The rural area was endemic for helminths with 54.2% of children being positive for *Schistosoma haematobium* and 38.3% having at least one intestinal helminth infection. Malaria infection was detected among 52.2% of rural participants. Both urban schools were free of *S. haematobium* infection with no detectable malaria infection among the urban low SES children and one case among urban high SES children. Intestinal helminth infections were present in both urban schools and affected 13.5% of urban low SES participants compared to 2.0% of urban high SES subjects. Serological analysis showed that the geometric mean c-reactive protein (CRP) level was 5 times higher among rural children compared to their urban counterparts and slightly lower in the urban low SES compared to urban high SES school. We also observed that the geometric mean total IgE level was significantly elevated among rural compared to both urban low SES and urban high SES children ($p < 0.001$).

Living in the rural area was associated with elevated total IgE (but not CRP) after taking parasitic infections into account

We analyzed whether area differences in CRP and total IgE levels still remained after taking parasitic infections into account. Given that these parameters were not normally distributed, CRP was categorized into a binary variable using the geometric mean (1.1 mg/ml) as a cut-off while total IgE was log-transformed. After adjusting for current helminth and malaria infections, living in the rural area still was strongly associated with elevated total IgE ($p < 0.001$) but not CRP ($p=0.232$). In addition, malaria infection was independently associated with CRP ($p < 0.05$) after adjusting for area.

Gene expression profiles can vary as a function of urban-rural area & SES

Gene expression levels (expressed as z-scores) in peripheral blood samples of children in the three areas were compared and are shown in Figure 1.The results of detailed Bonferroni pairwise comparisons of between-area differences in gene expression levels are shown in Table S2 (Supplementary material). IgE mRNA expression was strongly elevated among children in the rural area, where helminths were highly prevalent, compared to both urban schools. A correlation was observed between total IgE and IgE mRNA (Spearman's rho correlation coefficient = 0.61, $p < 0.001$). The expression levels for genes with immunosuppressive activities such as IL-10 and PD-1 were highest among rural participants followed by the urban low SES school and lastly the urban high SES school (p-trend < 0.001). Messenger RNA levels for IL-10 and for PD-1 were also correlated with each other (Spearman's rho correlation coefficient = 0.55, $p < 0.001$) as shown in Figure S1 (Supplementary material). Other genes involved in immune regulation such as TGF-β and FOXP3 were lowest in the rural area and while TGF-β levels showed a gradient increase across the schools from the rural to urban low SES to urban high SES, FOXP3 expression was similar in both urban schools.

The expression of some genes involved in TLR signalling, specifically TLR-2, TLR-4, CD14, NOD-2, SOCS-3 and LIR-7 were all high in the blood of urban high SES children but lower in both urban low SES and rural children. Post-hoc pair-wise comparison tests indicated that observed differences between the urban schools were significant ($p < 0.05$).

Gene expression profiles can be affected by parasitic infections

The influence of parasitic infections on gene expression profiles, independent of area differences, was assessed by examining expression in the rural area only. Children with current *S. haematobium* infection had significantly lower CD14, LIR-7 and CD28 mRNA expression levels ($p < 0.05$) as shown in Figure 2. The mean expression levels of other genes involved in the TLR signalling pathway were lower among *S. haematobium* positives compared to negatives but these were not statistically significant. As expected, relative IgE mRNA expression showed the opposite trend and was higher

Table 1: Characteristics of the Study Population stratified by school

Factor	RURAL N = 48	URBAN		TOTAL N = 151
		Urban Low SES N = 47	Urban High SES N = 56	
Age*, mean (range), years	9.1 (6 - 13)	8.8 (6 - 12)	8.7 (5 - 13)	8.9 (5 – 13)
Gender, Male	22 / 48 (45.8%)	26 / 47 (55.3%)	30 / 56 (53.6%)	78 / 151 (51.7%)
Parasitic Infections n / N (%)				
S. haematobium positive a,b	26 / 48 (54.2%)	0 / 47 (0.0%)	0 / 56 (0.0%)	26 / 151 (17.2%)
Intestinal helminth positive** a,b,c	18 / 47 (38.3%)	5 / 37 (13.5%)	1 / 50 (2.0%)	24 / 134 (17.9%)
Malaria infection Positive *** a,b	24 / 46 (52.2%)	0 / 35 (0.0%)	1 / 55 (1.8%)	25 / 136 (18.4%)
Serology				
CRP (mg/ml), geometric mean (95%CI)§ a,b	3.0 (2.0 - 4.6)	0.6 (0.4 – 0.8)	0.9 (0.6 – 1.2)	1.1 (0.9 – 1.4)
Total IgE (IU/ml), geometric mean (95%CI) §§ a,b	5061.6 (3419.0 – 7491.3)	228.6 (157.4 – 332.9)	234.10 (159.6 – 343.3)	526.6 (383.7 – 722.8)

Abbreviations: CRP, C-reactive protein; CI, confidence interval; IgE, Immunoglobulin E; SES, socioeconomic status.
Results of Mann–Whitney and Pearson's $\chi 2$ Tests comparing variables of interest by area.
a: if P<0.05 for rural versus urban low SES.
b: if P<0.05 for rural versus urban high SES.
c: if P<0.05 for urban low SES versus urban high SES.
* Missing ages for 3 rural participants.
** Intestinal helminths detected were Hookworm, *Ascaris lumbricoides*, *Trichuris trichiura*; missing intestinal helminth information for 1 rural, 10 urban low SES and 6 urban high participants.
*** Missing malaria information for 2 rural, 12 urban low SES and 1 urban high-SES participants.
§ Missing CRP values for 4 rural, 1 urban low SES and 3 urban high SES participants.
§§ Missing total IgE values for 11 rural and 1 urban low SES participants.

among *S. haematobium* infected children compared to uninfected children. With respect to intestinal helminths, as shown in Figure S2 (Supplementary material), there was a tendency towards lower expression levels of some of the TLR signalling genes among intestinal helminth infected participants compared to uninfected but this did not reach statistical significance. This tendency was observed specifically for NOD-2 (p = 0.09) and SOCS-3 (p=0.07). Malaria infection was not significantly associated with the expression levels of the genes investigated (Figure S3 – Supplementary material).

IL-10 gene expression in the rural area was not associated with current parasitic infection

After adjusting for age and gender, no significant associations were observed between IL-10 gene expression in the rural area and *S. haematobium* infection (p=0.292), having any intestinal helminth (p=0.967) or malaria infection (p=0.728).

Urban High vs. Urban Low				+	+	+		+	+			+		
Urban High vs. Rural	+		+	+	+	+		+		+	+		+	+
Urban Low vs. Rural	+											+		+

Figure 1: Relative gene expression profiles in urban high SES, urban low SES and rural schoolchildren
The relative gene expression profiles in each area expressed as z-scores with standard errors. Analysis of covariance models with individual mRNA expression levels as outcomes adjusted for age and gender were used to generate estimated marginal mean expression levels in each area.
+ = if $p < 0.05$ for the Bonferroni post-hoc pairwise comparison of estimated marginal mean gene expression levels between the 3 areas.

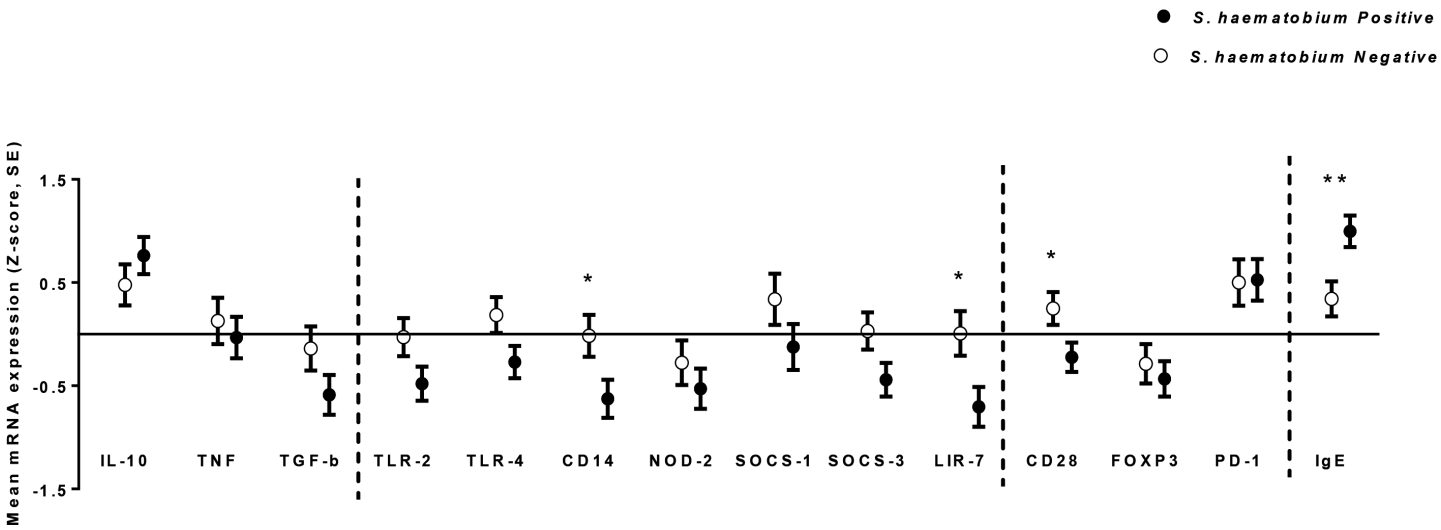


Figure 2: Relative expression profile in rural children stratified by *S.haematobium* infection status

The relative gene expression profiles in the rural area expressed as z-scores with standard errors stratified by *S. haematobium* infection status (positive versus negative). Analysis of covariance models with individual mRNA expression levels as outcomes adjusted for age, gender, intestinal helminth and malaria infection were used to generate estimated marginal means.

** P <0.01, * P <0.05 for analysis of covariance model test of between subject effects.

Area differences in IL-10 gene expression are not fully accounted for by IL-10 polymorphisms

Given the striking differences observed in relative IL-10 mRNA expression between rural and urban schoolchildren, we examined whether underlying variations in genetic polymorphisms could account for these differences.

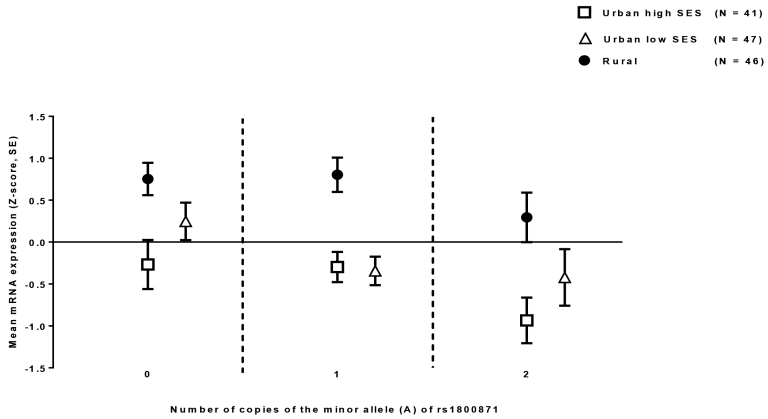
The minor allele frequencies (MAFs) for selected *IL-10* SNPs genotyped are shown in Table S3 (Supplementary material) and were compared to the Yoruba in Ibadan, Nigeria (YRI) genotype data from HapMap (www.hapmap.org) as well to data from a population study conducted in Northern Ghana [24]. As shown in Table S3, MAFs for the polymorphisms examined were similar across the three populations.

Table S4 details the MAFs for the *IL-10* SNPs stratified by area as well as the results of comparative between-area Pearson's χ^2 tests. Some significant differences in MAFs between the three areas were observed. Specifically, the MAF for rs3024496 was significantly higher in the rural school compared to both urban schools and the MAF for rs1878672 significantly lower in the urban high SES school compared to the other schools. In addition, the MAF for rs1800890 was significantly lower in the urban high SES school compared to rural school but not the urban low SES school.

Associations between *IL-10* SNPs and IL-10 mRNA expression were also examined for subjects with data for both parameters (N=134) and the results are shown in Table S5 (Supplementary material). After adjusting for area, only the marker rs1800871 was significantly associated with IL-10 mRNA ($p < 0.05$) with increasing copies of the minor allele of this SNP corresponding to decreasing IL-10 mRNA expression (shown in Figure S4 – Supplementary material). Figure 3 shows IL-10 mRNA expression in rural, urban low SES and urban high SES children according to the number of copies of the minor allele of *IL-10* SNP rs1800871. For rural children with two copies of the minor allele of rs1800871, IL-10 mRNA was higher compared to urban low SES and urban high SES children. When we examined the association between area and IL-10 mRNA after adjusting for the SNP rs1800871 and demographic factors, living in the rural area was still strongly associated with elevated IL-10 mRNA ($p = 1.87 \times 10^{-7}$).

Area differences in IL-10 gene expression are not explained by differences in ethnicity

Given that a few significant area differences in *IL-10* SNP frequencies were observed which may reflect variations in underlying genetics between the areas, we analyzed whether differences in reported ethnicity among the three groups could explain IL-10 gene expression variability. Information on ethnicity was available for 85 out of 151 children. For this subset, there was no association between reported ethnicity and IL-10 mRNA expression ($p=0.469$) after adjusting for age, gender and area. In this adjusted model, rural area was still strongly associated with elevated IL-10 mRNA expression ($p= 0.001$).



Urban High vs. Urban Low			
Urban High vs. Rural	+	+	+
Urban Low vs. Rural		+	

Figure 3: Relative IL-10 gene expression and copies of IL-10 SNP rs1800871

The relative IL-10 gene expression in urban high SES, urban low SES and rural children stratified by genotypes of *IL-10* SNP rs1800871.

+ = if p <0.05 for the Bonferroni post-hoc pairwise comparison of estimated marginal mean IL-10 gene expression levels between the 3 areas.

Area differences in TLR-2 and TLR-4 gene expression profiles are not explained by TLR-2 and TLR-4 polymorphisms

The minor allele frequencies for *TLR-2* and *TLR-4* SNPs in our population compared to the YRI genotype data from HapMap as well as to a population in Northern Ghana [25] are shown in Table S6 (Supplementary material). The MAFs across these three populations were comparable overall.

The area differences in minor allele frequencies for *TLR-2 and TLR-4* SNPs are shown in Table S7 (Supplementary material). For the *TLR-2* marker rs3804100, there were significant area differences but the overall MAF was low in our population. A significant difference was also observed for the marker rs4696480 which was significantly higher in the urban high SES school compared to urban low SES school. For *TLR-4* markers, the MAF of rs2737190 was significantly different between rural and urban high SES children (p=0.048) but the overall MAF of this SNP was low in our population. For rs10759932, there was a significant difference in MAF between rural and urban high SES children (p=0.036). A total of 142 children had data for TLR polymorphisms and TLR mRNA. None of the *TLR-2* and *TLR-4* markers were significantly associated with TLR-2 and TLR-4 mRNA expression, respectively (Table S8 - Supplementary material). In addition, adjusting for the *TLR-2* and *TLR-4* SNPs did not make a difference to observed area differences in mean TLR-2 and TLR-4 mRNA expression levels, respectively.

Discussion

In our study, we observed marked differences in the gene expression profiles of Ghanaian children attending schools in rural, urban low SES and urban high SES areas in the Greater Accra Region of Southern Ghana. Our study demonstrates that environmental determinants associated with specific geographical locations and lifestyle, have a strong impact on shaping immune gene expression profiles. Similar observations were made by Idaghadour *et al.* [14, 15] who used whole-genome expression arrays and found a genome-wide expression signature of regional population differences in Morocco. However, in our investigation, we further examined the effects of specific infections and socioeconomic differences on the gene expression patterns of our study population.

Higher gene expression levels of IgE, IL-10 and PD-1 were seen in the peripheral blood of rural compared to urban children in our study population. *S. haematobium* infection could account for the urban-rural difference observed in IgE gene expression as this infection in the rural area was strongly associated with higher IgE mRNA expression. This would be expected since high serum IgE protein is an established hallmark of schistosomiasis infection [29]. In addition, a publication by Hartgers *et al*. validated *ex vivo* mRNA expression in rural Ghana by showing a strong correlation between IgE mRNA and serum IgE [21].

Interestingly, current helminth infections did not account for the markedly elevated IL-10 and PD-1 mRNA levels among rural compared to urban children. For IL-10 this was somewhat unexpected since chronic helminth infections are characterized by an anti-inflammatory environment marked by elevated IL-10 and TGF-β [19]. Given this anomaly, we went on to examine whether differences in ethnicity or genetic polymorphisms may explain elevated IL-10 mRNA among rural compared to urban children. In a subset, no association was observed between ethnicity and IL-10 gene expression. Regarding genetic variants, while the minor allele of one particular marker rs1800871 was strongly associated with decreased expression of IL-10 mRNA, after controlling for the effects of this SNP, IL-10 mRNA expression in the rural area was still elevated compared to the two urban areas. This suggests that although underlying genetics plays a role in gene expression profiles, environmental factors appear to have a dominant influence. The lack of association between helminth infection and IL-10 mRNA expression might be due to the fact that in our study we have looked at current helminth infection whereas past infections could also have shaped the regulatory network. An additional factor to note is that the method used to diagnose helminth infections, might not have been sensitive enough to detect all infected subjects. Moreover, post-transcriptional regulation of the IL-10 gene has been reported [30] and therefore IL-10 gene expression might be different from the production of the protein. However, the possibility that factors other than helminth infections, such as environmental mycobacteria [31] play an important role in expanding IL-10 gene expression, cannot be ruled out and would need further investigation.

The level of IL-10 mRNA expression was also positively correlated with PD-1 gene expression. PD-1 protein is thought to have regulatory functions inhibiting T-cell proliferation and cytokine production [32]. In addition, PD-1 receptor is a well-established marker of 'T-cell exhaustion' which is the progressive loss of T-cell function under conditions of antigen persistence following chronic infections such as viral infections [33]. The role of IL-10 in T-cell exhaustion has also been demonstrated in the murine lymphocytic choriomeningitis virus system [34]. The correlation between IL-10 mRNA and PD-1 mRNA in our rural environment could indicate T-cell exhaustion in this area resulting from a chronic persistent infection.

We observed that the expression of TGF-β and FOXP3 was higher in the peripheral blood of children in both urban areas compared to their rural counterparts. A significant positive correlation between TGF-β and FOXP3 gene expression levels was also observed. The higher expression of these genes in the urban areas relative to the helminth-endemic rural area was unexpected since both TGF-β and FOXP3 are thought to be up-regulated as part of the immune regulatory network associated with chronic infections [35]. In addition, regulatory T cells expressing FOXP3 driven by TGF-β have been implicated in the suppression of host immunity during chronic helminth infection [36, 37]. It would be important to examine whether these molecules show higher expression in urban areas at the protein level as well.

The other set of genes with elevated expression in urban high SES subjects compared to the other two groups were some of those involved in TLR signalling. Specifically, TLR-2, TLR-4, CD14, NOD-2, SOC-3 and LIR-7. Of interest, studies in European farmers have indicated that higher exposure to microorganisms might be associated with higher expression of receptors such as TLR-2 and CD14 [38]. Here, we observed that *S. haematobium* infection in the rural area could explain the lower expression of CD14 and LIR-7 genes among the rural children and in comparison to the two urban groups. Other studies have shown that *Schistosoma* egg antigen down-regulates the expression of genes involved in the TLR signalling pathway [39] and we have previously reported lower TLR-2 mRNA expression associated with current *S. haematobium* infection in rural Ghanaian children [21]. The contrast between the effects of an environment rich in microorganisms or parasites on the expression of pattern recognition receptors (PRRs) in Europe versus in Africa might be explained by the very different types and burden of microorganisms and parasites that are present in these environments. Thus, in rural Ghana, exposures might lead to down-regulation whereas in central European farms, to the up-regulation of PRRs. Moreover, specific lifestyle factors may have a suppressive effect on the expression of genes involved in interactions with PAMPs within our population. None of the *TLR-2* or *TLR-4* SNPs examined was significantly associated with increased TLR-2 or TLR-4 mRNA expression respectively. Polymorphisms of *TLR* genes are of particular importance given the key role of TLRs as PRRs in host defence mechanisms against microbial pathogens [40].

Overall, we observed that not only were the rural and urban areas different but that there were also significant differences within the urban area. Changes along a gradient

from rural to urban low SES to urban high SES implicate factors that are likely to reflect exposure to pathogens. However, factors that segregate into urban (irrespective of SES) and rural area are likely to reflect characteristics of urbanization for example, changes in diet or pollution.

One limitation of our current study was that it was conducted with relatively small numbers of subjects in each study area. A larger sample size may have meant greater statistical power in detecting area differences for some of the genes as evident in some of the borderline significant observations. However, post-hoc power analysis based on mean IL-10 mRNA levels in urban compared to rural children showed the study to be sufficiently powered. Another weakness of our study was that there was only a single sample collected for parasitic infections. Therefore, there is the possibility that if infections were missed, we underestimated the prevalence of our parasitic infections. Although the focus of our genetic polymorphism selection was on variants of importance in Caucasian populations, studies conducted in Northern Ghana that included most of the genetic variants at the IL-10 gene locus examined in our study, demonstrated the functional importance of these variants in a Ghanaian population [24, 41]. Specifically, the minor allele of rs1800871 that was negatively associated with IL-10 mRNA in our study was negatively associated with *ex vivo* IL-10 cytokine production in response to stimulation with *E. coli* lipopolysaccharide and *Saccharomyces cerevisiae* zymosan in these two studies [24, 41]. However, the possibility still exists that IL-10 polymorphisms that were not examined in our study, may contribute to the observed rural versus urban differences in the expression of IL-10 mRNA.

Common to all mRNA studies of whole blood, our study suffers from the fact that mRNA expression might not be directly related to protein expression levels. Although we used IgE as a positive control, showing that mRNA expression was paralleled by protein levels, this might not be the case for all genes examined. In addition, the expression of the mRNA is in whole blood and does not reveal any cell-specific profiles which might be important when considering their function in determining disease profiles. An additional weakness of our investigation is that differential blood cell counts were not assessed and differences in cellular composition may play a role in the expression patterns observed.

Despite the limitations, our study demonstrates that contrasting environments shaped by urbanization and associated characteristics contribute significantly to gene expression profiles among children. Future studies are needed to identify specific factors that activate particular immunological pathways and to understand the functional consequences of the differential gene expression profiles observed in terms of disease patterns and susceptibility.

Acknowledgements

The authors would like to thank Dr. Anita van den Biggelaar for advice and feedback on the manuscript. We would also like to thank Ms. Mercy Geyi and Mr. Jonas Asigbee for

assistance in parasitological sampling, Dr. Marjolijn Duijvestein for assisting in mRNA measurements and Dr. Martin Depner for polymorphism genotyping. We are most indebted to the study participants and their families, teachers, school authorities and community leaders. This work was supported in part by the Netherlands Foundation for the Advancement of Tropical Research (grant WB 93-443), the Netherlands Organisation for Scientific Research, ZonMW TOP program (grant 912-03-048) and the European Commission (GA2LEN-FOOD-CT-2004-506378, GLOFAL-FOOD-CT-2005-517812 & EUROPREVALL-FOOD-CT-2005-514000).

Conflict of interest

The authors declare no conflict of interest.

References

1. WHO, Global status report on noncommunicable diseases 2010. Geneva: World Health Organization, 2011.
2. Rook GW, Hygiene Hypothesis and Autoimmune Diseases. Clinical Reviews in Allergy and Immunology 2012;42: 5-15.
3. Isunju JB, Schwartz K, Schouten MA, Johnson WP, van Dijk MP, Socio-economic aspects of improved sanitation in slums: A review. Public Health 2011;125: 368-76.
4. Alirol E, Getaz L, Stoll B, Chappuis F, Loutan L, Urbanisation and infectious diseases in a globalised world. The Lancet Infectious Diseases 2011;11: 131-41.
5. Ibrahim MM, Damasceno A, Hypertension in developing countries. The Lancet;380: 611-19.
6. Swinburn BA, Sacks G, Hall KD, McPherson K, Finegood DT, Moodie ML, Gortmaker SL, The global obesity pandemic: shaped by global drivers and local environments. The Lancet 2011;378: 804-14.
7. Dalal S, Beunza JJ, Volmink J, Adebamowo C, Bajunirwe F, Njelekela M, Mozaffarian D, Fawzi W, Willett W, Adami H-O, Holmes MD, Non-communicable diseases in sub-Saharan Africa: what we know now. International Journal of Epidemiology 2011;40: 885-901.
8. Popkin BM, Adair LS, Ng SW, Global nutrition transition and the pandemic of obesity in developing countries. Nutrition Reviews 2012;70: 3-21.
9. WHO, Increasing access to health workers in remote and rural areas through improved retention. World Health Organization, Geneva: WHO, 2010.
10. Fezeu L, Balkau B, Kengne A-P, Sobngwi E, Mbanya J-C, Metabolic syndrome in a sub-Saharan African setting: Central obesity may be the key determinant. Atherosclerosis 2007;193: 70-76.
11. Delisle H, Ntandou-Bouzitou G, Agueh V, Sodjinou R, Fayomi B, Urbanisation, nutrition transition and cardiometabolic risk: the Benin study. British Journal of Nutrition 2012;107: 1534-44.
12. Zeller T, Wild P, Szymczak S, Rotival M, Schillert A, Castagne R, Maouche S, Germain M, Lackner K, Rossmann H, Eleftheriadis M, Sinning CR, Schnabel RB, Lubos E, Mennerich D, Rust W, Perret C, Proust C, Nicaud V, Loscalzo J, Hübner N, Tregouet D, Münzel T, Ziegler A, Tiret L, Blankenberg S, Cambien F, Genetics and Beyond -- The Transcriptome of Human Monocytes and Disease Susceptibility. PLOS ONE 2010;5: e10693.
13. Cookson W, Liang L, Abecasis G, Moffatt M, Lathrop M, Mapping complex disease traits with global gene expression. Nature Reviews Genetics 2009;10: 184-94.
14. Idaghdour Y, Storey JD, Jadallah SJ, Gibson G, A Genome-Wide Gene Expression Signature of Environmental Geography in Leukocytes of Moroccan Amazighs. PLOS Genetics 2008;4: e1000052.
15. Idaghdour Y, Czika W, Shianna K, Lee S, Visscher P, Martin H, Miclaus K, Jadallah

S, Goldstein D, Wolfinger R, Gibson G, Geographical genomics of human leukocyte gene expression variation in southern Morocco. Nat Genet 2010;42: 62-7.

16. Koukounari A, Webster JP, Donnelly CA, Bray BC, Naples J, Bosompem K, Shiff C, Sensitivities and Specificities of Diagnostic Tests and Infection Prevalence of Schistosoma haematobium Estimated from Data on Adults in Villages Northwest of Accra, Ghana. The American Journal of Tropical Medicine and Hygiene 2009;80: 435-41.
17. Medzhitov R, Toll-like receptors and innate immunity. Nature Reviews Immunology 2001;1: 135-45.
18. Hussaarts L, van der Vlugt LEPM, Yazdanbakhsh M, Smits HH, Regulatory B-cell induction by helminths: Implications for allergic disease. Journal of Allergy and Clinical Immunology 2011;128: 733-39.
19. Maizels RM, Smith KA, Chapter 3 - Regulatory T Cells in Infection. In: Rudensky A, Sakaguchi S eds. Advances in Immunology: Academic Press, 2011:73-136.
20. Beddow V, ACCRA: Sanitation Status IWA WaterWiki 2010
21. Hartgers FC, Obeng BB, Kruize YCM, Duijvestein M, de Breij A, Amoah A, Larbi IA, van Ree R, Wilson MD, Rodrigues LC, Boakye DA, Yazdanbakhsh M, Lower Expression of TLR2 and SOCS-3 Is Associated with Schistosoma haematobium Infection and with Lower Risk for Allergic Reactivity in Children Living in a Rural Area in Ghana. PLOS Neglected Tropical Diseases 2008;2: e227.
22. Kormann MSD, Depner M, Hartl D, Klopp N, Illig T, Adamski J, Vogelberg C, Weiland SK, von Mutius E, Kabesch M, Toll-like receptor heterodimer variants protect from childhood asthma. Journal of Allergy and Clinical Immunology 2008;122: 86-92.e8.
23. Kormann MSD, Carr D, Klopp N, Illig T, Leupold W, Fritzsch C, Weiland SK, von Mutius E, Kabesch M, G-Protein–coupled Receptor Polymorphisms Are Associated with Asthma in a Large German Population. American Journal of Respiratory and Critical Care Medicine 2005;171: 1358-62.
24. Kuningas M, May L, Tamm R, van Bodegom D, van den Biggelaar AH, Meij JJ, Frolich M, Ziem JB, Suchiman HE, Metspalu A, Slagboom PE, Westendorp RG, Selection for genetic variation inducing pro-inflammatory responses under adverse environmental conditions in a Ghanaian population. PLOS ONE 2009;4: e7795.
25. May L, van Bodegom D, Frolich M, van Lieshout L, Slagboom PE, Westendorp RG, Kuningas M, Polymorphisms in TLR4 and TLR2 genes, cytokine production and survival in rural Ghana. European journal of human genetics : EJHG 2010;18: 490-5.
26. van den Biggelaar AH, Lopuhaa C, van Ree R, van der Zee JS, Jans J, Hoek A, Migombet B, Borrmann S, Luckner D, Kremsner PG, Yazdanbakhsh M, The prevalence of parasite infestation and house dust mite sensitization in Gabonese schoolchildren. International Archives of Allergy and Immunology 2001;126: 231-8.
27. Koenig W, Sund M, Frohlich M, Fischer HG, Lowel H, Doring A, Hutchinson WL, Pepys MB, C-Reactive protein, a sensitive marker of inflammation, predicts future risk of coronary heart disease in initially healthy middle-aged men: results from the MONICA (Monitoring Trends and Determinants in Cardiovascular Disease) Augsburg Cohort Study, 1984 to 1992. Circulation 1999;99: 237-42.
28. Barrett JC, Fry B, Maller J, Daly MJ, Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005;21: 263-65.
29. Capron M, Capron A, Immunoglobulin E and effector cells in schistosomiasis. Science 1994;264: 1876-7.
30. Powell MJ, Thompson SA, Tone Y, Waldmann H, Tone M, Posttranscriptional regulation of IL-10 gene expression through sequences in the 3'-untranslated region. Journal of Immunology (Baltimore, Md : 1950) 2000;165: 292-6.
31. Rook GA, Regulation of the immune system by biodiversity from the natural environment: an ecosystem service essential to health. Proceedings of the National Academy of Sciences of the United States of America 2013;110: 18360-7.
32. van Riet E, Everts B, Retra K, Phylipsen M, van Hellemond J, Tielens A, van der Kleij D, Hartgers F, Yazdanbakhsh M, Combined TLR2 and TLR4 ligation in the context of bacterial or helminth extracts in human monocyte derived dendritic cells: molecular correlates for Th1/Th2 polarization. BMC Immunology 2009;10: 9.

33. Yi JS, Cox MA, Zajac AJ, T-cell exhaustion: characteristics, causes and conversion. Immunology 2010;129: 474-81.
34. Rodriguez-Garcia M, Porichis F, de Jong OG, Levi K, Diefenbach TJ, Lifson JD, Freeman GJ, Walker BD, Kaufmann DE, Kavanagh DG, Expression of PD-L1 and PD-L2 on human macrophages is up-regulated by HIV-1 and differentially modulated by IL-10. Journal of Leukocyte Biology 2011;89: 507-15.
35. Sakaguchi S, Miyara M, Costantino CM, Hafler DA, FOXP3+ regulatory T cells in the human immune system. Nature Reviews Immunology 2010;10: 490-500.
36. Grainger JR, Smith KA, Hewitson JP, McSorley HJ, Harcus Y, Filbey KJ, Finney CA, Greenwood EJ, Knox DP, Wilson MS, Belkaid Y, Rudensky AY, Maizels RM, Helminth secretions induce de novo T cell Foxp3 expression and regulatory function through the TGF-beta pathway. The Journal of Experimental Medicine 2010;207: 2331-41.
37. Xu L, Kitani A, Strober W, Molecular mechanisms regulating TGF-beta-induced Foxp3 expression. Mucosal Immunology 2010;3: 230-8.
38. Lauener RP, Birchler T, Adamski J, Braun-Fahrländer C, Bufe A, Herz U, von Mutius E, Nowak D, Riedler J, Waser M, Sennhauser FH, Expression of CD14 and Toll-like receptor 2 in farmers' and nonfarmers' children. The Lancet 2002;360: 465-66.
39. Kane CM, Cervi L, Sun J, McKee AS, Masek KS, Shapira S, Hunter CA, Pearce EJ, Helminth Antigens Modulate TLR-Initiated Dendritic Cell Activation. The Journal of Immunology 2004;173: 7454-61.
40. Kawai T, Akira S, Toll-like receptors and their crosstalk with other innate receptors in infection and immunity. Immunity 2011;34: 637-50.
41. Boef AGC, May L, van Bodegom D, Kuningas M, Eriksson UK, Westendorp RGJ, The influence of genetic variation on innate immune activation in an environment with high infectious pressure. Genes and Immunity 2012;13: 103-08.

Table S1: Description of genes*

	GENE		
CATEGORY	**ABBREVIATION**	**FULL NAME**	**FUNCTION OF PROTEIN**
Cytokines	IL-10	Interleukin 10	Cytokine with multiple effects in immune regulation and inflammation.
	TNF	Tumor necrosis factor	Multifunctional pro-inflammatory cytokine
	TGF-*b*	Transforming growth factor beta	Immunosuppressive cytokine with multiple regulatory functions
Toll Like Receptor Signalling	TLR-2	Toll-like receptor 2	Key receptor in pathogen recognition and activation of innate immunity.
	TLR-4	Toll-like receptor 4	Key receptor in pathogen recognition and activation of innate immunity.
	CD14	Cluster of differentiation 14	Preferentially expressed surface antigen that cooperates with other proteins to mediate the innate immune response to bacterial lipopolysaccharide
	NOD-2	Nucleotide-binding oligomerization domain containing 2	Intracellular pattern recognition receptor that recognizes bacterial muramyl dipeptide
	SOCS-1	Suppressor of cytokine signalling 1	Negative regulator of cytokines that signals through the JAK/STAT3 pathway
	SOCS-3	Suppressor of cytokine signalling 3	Negative regulator of cytokines that signals through the JAK/STAT pathway
T-cell activation & polarization	LIR-7	Leukocyte immunoglobulin-like receptor	Receptor with immunosuppressive properties but also implicated in activation
	CD28	Cluster of Differentiation 28	Co-stimulatory molecule expressed on T-cells that provides signals for T-cell activation
	FOXP3	Forkhead box P3	Master regulator in the function of regulatory T-cells
	PD-1	Programmed cell death 1	Receptor with inhibitory functions known as a potent regulator of immune responses
Antibody	IgE	Immunoglobulin E	Antibody key in allergic responses and helminth infections

*Adapted from van Riet E *et al.,* Combined TLR2 and TLR4 ligation in the context of bacterial or helminth extracts in human monocyte derived dendritic cells: molecular correlates for Th1/Th2 polarization. *BMC Immunology 2009*; 10(1):9.

Table S2: Results of Bonferroni pair-wise comparisons of between area differences in gene expression levels

	Between Area Bonferroni Pair-wise Comparisons									ANCOVA Test of Between-Subject Effects for Area and Gene Expression (Outcome)
	Urban High SES (I) vs Urban Low SES (J)			Urban High SES (I) vs Rural (J)			Urban Low (I) vs Rural (J)			
GENE	Mean difference (I – J)	Std. Error	p-value	Mean difference (I – J)	Std. Error	p-value	Mean difference (I – J)	Std. Error	p-value	p-value
IL-10	-0.34	0.17	0.152	-1.12	0.17	**5.73×10^{-9}**	-0.78	0.18	**1.08×10^{-4}**	**9.53×10^{-9}**
TNF	0.45	0.19	0.069	0.20	0.20	0.968	-0.25	0.21	0.671	0.074
TGF-b	0.40	0.19	0.108	0.79	0.19	**2.27×10^{-4}**	0.39	0.20	0.172	**3.69×10^{-4}**
TLR-2	0.89	0.18	**9.21×10^{-6}**	0.86	0.19	**2.33×10^{-5}**	-0.03	0.19	1.000	**7.45×10^{-7}**
TLR-4	0.95	0.18	**2.60×10^{-6}**	0.53	0.19	**0.015**	-0.41	0.19	0.106	**4.58×10^{-6}**
CD14	0.70	0.17	**3.07×10^{-4}**	0.92	0.18	**2.32×10^{-6}**	0.22	0.18	0.704	**1.09×10^{-6}**
NOD-2	0.63	0.19	**0.003**	0.77	0.19	**2.05×10^{-4}**	0.14	0.20	1.000	**9.01×10^{-5}**
SOCS-1	0.22	0.20	0.805	0.03	0.20	1.000	-0.19	0.21	1.000	0.504
SOCS-3	0.87	0.19	**2.10×10^{-5}**	0.73	0.19	**4.82×10^{-4}**	-0.13	0.20	1.000	**7.20×10^{-6}**
LIR-7	0.90	0.18	**4.62×10^{-6}**	0.99	0.18	**7.80×10^{-7}**	0.09	0.19	1.000	**5.52×10^{-8}**
CD28	0.19	0.20	1.000	0.11	0.20	1.000	-0.08	0.21	1.000	0.634
FOXP3	-0.01	0.19	1.000	0.56	0.20	**0.016**	0.58	0.21	**0.018**	**0.007**
PD-1	-0.45	0.18	**0.049**	0.89	0.19	**1.34×10^{-5}**	-0.45	0.20	0.071	**2.48×10^{-5}**
IgE	0.01	0.18	1.000	0.95	0.18	**2.20×10^{-6}**	-0.963	0.19	**3.91×10^{-6}**	**2.29×10^{-7}**

P-values <0.05 are shown in bold.

Table S3: *IL-10* Polymorphisms (N=330)

Marker	Chromosome Position	Gene Location	Alleles*	Minor Allele Frequency			Hardy Weinberg p-value (current study)
				Current Study Population	Northern Ghana Study Population **	Yoruba in Ibadan (YRI) ***	
rs3024498	1: 206941529	Exon	T/C	0.085	0.083	0.092	0.052
rs3024496	1: 206941864	Exon	A/G	0.399	0.425	0.371	0.142
rs1878672	1: 206943713	Intron	G/C	0.233	0.244	0.235	0.209
rs1800871	1: 206946634	Promoter	G/A	0.398	0.470	0.466	0.261
rs1800893	1: 206947167	Promoter	C/T	0.324	0.284	0.303	**0.001**
rs1800890	1:206949365	Promoter	A/T	0.220	0.201	0.204	0.894
rs12122923	1:206951397	Promoter	C/T	0.123	n.a.	0.131	0.252
rs10494879	1:206952204	Promoter	C/G	0.285	0.284	0.255	0.201

*Minor allele underlined.
IL10 Minor allele Frequencies from a study conducted in a rural community in Northern Ghana (Kuningas M *et al*. Selection for genetic variation inducing pro-inflammatory responses under adverse environmental conditions in a Ghanaian population. *PLOS ONE* 2009; **4(11)**:** e7795), N= 4336.
***Minor allele Frequencies from **HapMap Database Release Number 28 PhaseII+III, August 2010 dbSNP b126.**
n.a.: Information not available.
P-values <0.05 are shown in bold.

Table S4: Minor allele frequencies of *IL-10* Polymorphisms stratified by area (N=330)

Marker	Minor Allele	Minor allele frequencies by Area			Pearson's χ^2 Test p-value for comparisons between areas		
		Urban High SES (N=109)	Urban Low SES (N=123)	Rural (N= 98)	Urban High SES vs. Urban Low SES	Urban High SES vs. Rural	Urban Low SES vs. Rural
rs3024498	C	0.084	0.087	0.082	0.922	0.999	0.925
rs3024496	G	0.305	0.392	0.505	0.058	**4.97×10^{-5}**	**0.018**
rs1878672	C	0.158	0.258	0.276	**0.012**	**0.005**	0.676
rs1800871	A	0.435	0.409	0.347	0.595	0.083	0.190
rs1800893	T	0.319	0.323	0.331	0.933	0.804	0.860
rs1800890	T	0.165	0.240	0.253	0.053	**0.032**	0.756
rs12122923	T	0.106	0.103	0.165	0.925	0.088	0.058
rs10494879	G	0.225	0.307	0.321	0.052	0.035	0.769

P-values <0.05 are shown in bold.

Table S5: *IL-10* Polymorphisms and IL-10 mRNA production (adjusted for age, gender and area)

IL-10 Marker name	Estimate (β)	Standard Error (SE)	IL-10 mRNA Linear Trend (p-value)
rs3024498	0.09	0.20	0.638
rs3024496	0.05	0.10	0.611
rs1878672	0.17	0.12	0.164
rs1800871	-0.28	0.11	**0.013**
rs1800893	0.20	0.11	0.084
rs1800890	0.14	0.13	0.277
rs12122923	0.04	0.15	0.776
rs10494879	0.16	0.12	0.172

P-values <0.05 are shown in bold.

Table S6: *TLR* Polymorphisms (N=318)

					Minor Allele Frequency			Hardy Weinberg p-value (current study)
Gene	Marker	Chromosome Position	Gene Location	Alleles*	Current Study Population	Northern Ghana Study Population **	Yoruba in Ibadan (YRI) ***	
TLR-2	rs3804099	4:154624656	Exon	C/T	0.374	0.380	0.364	0.131
	rs3804100	4:154625409	Exon	T/C	0.038	0.032	0.054	0.484
	rs4696480	4:154607126	Promoter	T/A	0.358	n.a.	0.000	0.120
TLR-4	rs4986790	9:120475302	Exon	A/G	0.094	0.075	0.041	0.392
	rs2737190	9:120464181	Promoter	G/A	0.094	n.a.	0.143	0.884
	rs10759932	9:120465144	Promoter	T/C	0.312	0.267	0.258	0.060
	rs4986791	9:120475602	Exon	C/T	0.009	0.012	0.022	0.865

*Minor allele underlined.
****** *TLR-2* and *TLR-4* gene minor allele frequencies from a study conducted in a rural community in Northern Ghana (May L *et al*. Polymorphisms in TLR-4 and TLR-2 genes, cytokine production and survival in rural Ghana. *European journal of human genetics* 2010; **18**(4)**:** 490-5), N= 4292.
******* Minor allele Frequencies from **HapMap Database Release Number 28 PhaseII+III, August 2010 dbSNP b126.**
n.a.: Information not available.

Table S7 Minor allele frequencies of TLR-2 and TLR-4 SNPs stratified by Area (N=318)

			Minor allele frequencies by Area			Pearson's χ^2 Test p-value for comparisons between areas		
Gene	Marker	Minor Allele	Urban High SES (N=115)	Urban Low SES (N=113)	Rural (N= 90)	Urban High SES vs. Urban Low SES	Urban High SES vs. Rural	Urban Low SES vs. Rural
TLR-2	rs3804100	T	0.065	0.027	0.017	**0.049**	**0.017**	0.502
	rs3804099	C	0.382	0.332	0.417	0.269	0.472	0.079
	rs4696480	A	0.400	0.308	0.367	**0.041**	0.491	0.214
TLR-4	rs4986790	G	0.089	0.089	0.106	0.985	0.574	0.582
	rs2737190	A	0.128	0.080	0.068	0.096	**0.048**	0.646
	rs10759932	C	0.271	0.306	0.369	0.409	**0.036**	0.186
	rs4986791	T	0.009	0.009	0.011	1.000	0.819	0.819

P-values <0.05 are shown in bold

Table S8: TLR Polymorphisms and TLR mRNA production (adjusted for age, gender and area)

Gene	Marker	Estimate (β)	Standard Error (SE)	TLR-2 mRNA Linear Trend p-value
TLR-2	rs3804099	0.084	0.123	0.496
	rs3804100	-0.456	0.292	0.120
	rs4696480	-0.032	0.119	0.790
Gene	**Marker**	**Estimate (β)**	**Standard Error (SE)**	**TLR-4 mRNA Linear Trend (p-value)**
TLR-4	rs4986790	-0.07	0.190	0.715
	rs2737190	-0.37	0.210	0.080
	rs10759932	0.033	0.128	0.795
	rs4986791	1.183	1.009	0.243

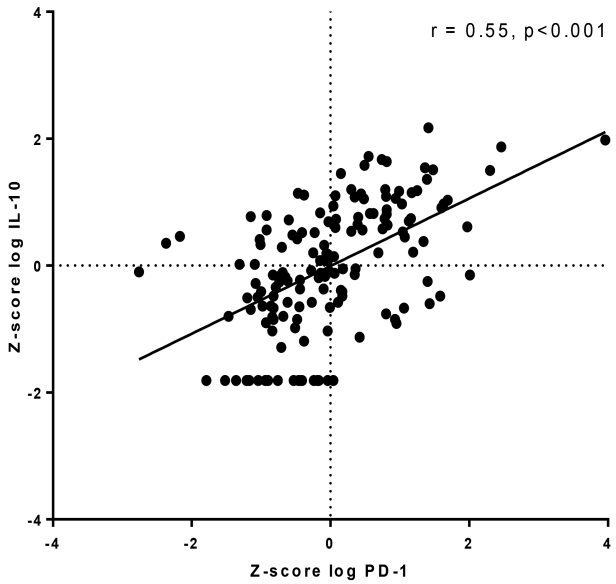


Figure S1: Correlation between IL-10 mRNA and PD-1 mRNA expressed as z-scores

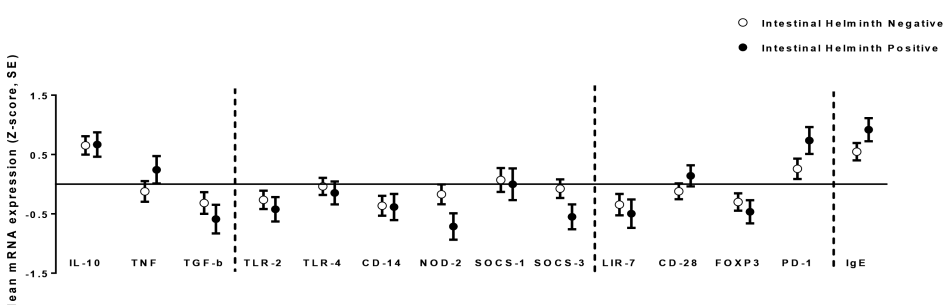


Figure S2: Relative gene expression profile in rural children stratified by intestinal helminth infection status (N=47)

The relative gene expression profiles in the rural area expressed as z-scores with standard errors stratified by intestinal helminth infection status (positive versus negative). Analysis of covariance models with individual mRNA expression levels as outcomes adjusted for age, gender and other parasitic infections were used to generate estimated marginal mean expression levels.

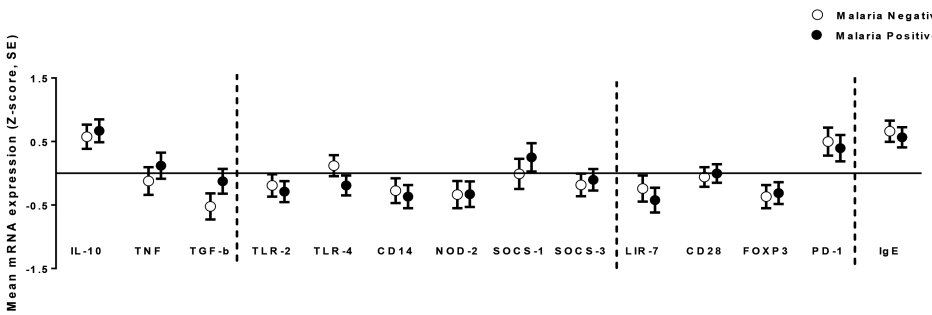


Figure S3: Relative gene expression profile in rural children stratified by malaria infection status (N=47)

The relative gene expression profiles in the rural area expressed as z-scores with standard errors stratified by malaria infection status (positive versus negative). Analysis of covariance models with individual mRNA expression levels as outcomes adjusted for age, gender and other parasitic infections were used to generate estimated marginal means.

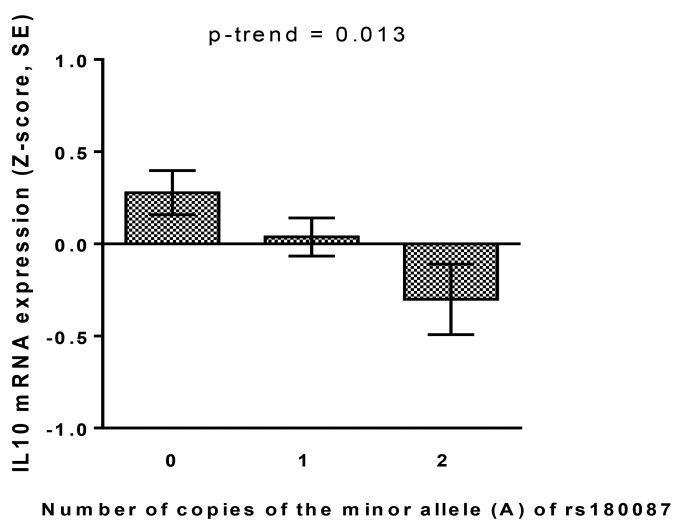


Figure S4: IL10 mRNA expression according to number of copies of marker rs1800871

The correlation between increasing copies of the minor allele of IL-10 SNP rs1800871 and IL-10 mRNA (expressed as a z-score with standard errors). An analysis of covariance model with IL-10 mRNA expression level as an outcome adjusted for age, gender and area was used to generate estimated marginal mean levels according to number of copies of the IL-10 SNP.

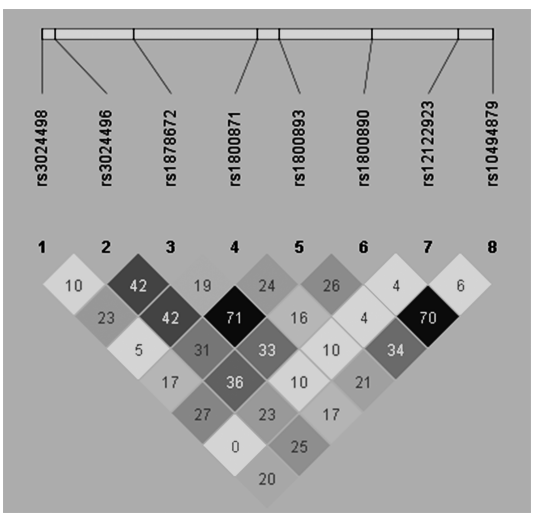
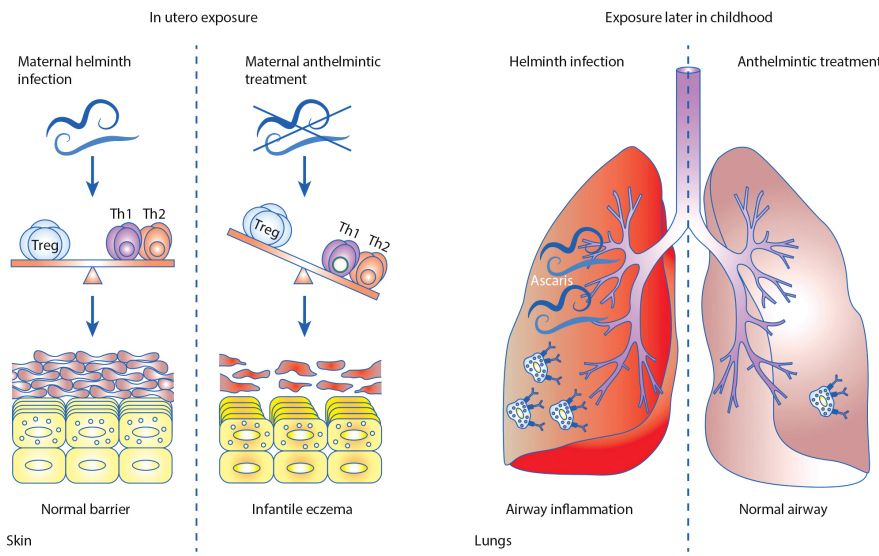


Figure S5: Linkage disequilibrium plot of IL-10 SNPs

The figure shows the linkage disequilibrium plot of IL-10 SNPs genotyped in the study and their relative positions. The increasing strength of correlation between SNPs is indicated from white to dark grey.



Discovering how anthelmintic treatment influences allergic symptoms in early and late childhood
Source: Pediatric Allergy and Immunology 2014 May; 25 (3)