

SHORT REPORT

# The frequency of an IL-18-associated haplotype in Africans

Simon R Thompson<sup>1</sup>, Steve E Humphries<sup>\*1,2</sup>, Mark G Thomas<sup>3,4</sup>, Rosemary Ekong<sup>5</sup>, Ayele Tarekegn<sup>4,7</sup>, Endeshaw Bekele<sup>6</sup>, Olivia Creemer<sup>5</sup>, Neil Bradman<sup>5</sup> and Krishna R Veeramah<sup>8</sup>

Variation within the gene for the proinflammatory cytokine interleukin (IL)-18 has been associated with inter-individual differences in levels of free protein and disease risk. We investigated the frequency of function-associated *IL18* gene haplotypes in an extensive sample ( $n = 2357$ ) of African populations from across the continent. A previously identified five tagging SNP (single-nucleotide polymorphism) haplotype (here designated hGATA), known to be associated with lower levels of IL-18, was observed at a frequency of 27% in a British population of recent European ancestry, but was found at low frequency (<8%) or completely absent in African populations. Potentially protective variants may, as a consequence, be found at low frequency in African individuals and may confer a difference in disease risk.

*European Journal of Human Genetics* (2013) 21, 465–468; doi:10.1038/ejhg.2012.184; published online 29 August 2012

**Keywords:** IL-18; atherosclerosis; genetic diversity; Africa

## INTRODUCTION

Interleukin (IL)-18 is a pleiotropic cytokine involved in both innate and adaptive immune responses.<sup>1,2</sup> Variation in the *IL18* gene has been shown to influence circulating levels of IL-18,<sup>3–6</sup> is associated with risk for various inflammatory diseases including cardiovascular disease (CVD),<sup>7–9</sup> and affects production of IL-18 by monocytes.<sup>10,11</sup> Though the distribution of genetic variation in *IL18* in populations of recent European origin has been examined, little is known about populations more globally. We have therefore investigated the distribution across Africa by determining the frequency of potentially informative haplotypes and compared the results with populations from neighbouring regions and the UK.

## MATERIALS AND METHODS

### Subjects

DNA samples (buccal swabs) from males over 18 years old unrelated at the paternal grandfather level ( $n = 2572$ ) were selected from the Centre for Genetic Anthropology sample collection to represent a wide selection of ethnic groups distributed across and close to Africa (Table 1). Individual populations were additionally amalgamated into broader geographical groups to assess regional variation. Samples from the second Northwick Park Hospital Heart Study cohort (NPHSII,  $n = 2696$ ),<sup>12</sup> which consists of individuals of European biogeographic ancestry aged 51–60 years from nine general medical practices across the UK, were used to represent a population of recent north-west European/UK ancestry.

### Genotyping

All samples were typed for five *IL18* tagging single-nucleotide polymorphisms (tSNPs).<sup>13</sup> TaqMAN probes were designed for all five tSNPs (–9731 G>T (rs1946519); –5848 T>C (rs2043055); +4860 A>C (rs549908); +8855 T>A (rs360729); +11015 A>C (rs3882891)) and assayed on an ABI Prism 7000

(Life Technologies Corporation, Carlsbad, CA, USA) sequence detection system. All per-SNP genotyping rates were >95%. A subset of samples, homozygous at all of the above positions, were genotyped for three further SNPs (–8949 C>T (rs360717); +11640 A>G (rs5744292); +11990 T>C (rs4937100)). Genotyping of –8949 C>T and +11640 A>G was carried out by K Biosciences (Hoddesdon, UK). Genotyping of +11990 T>C was carried out using dideoxy sequencing as previously described.

### Statistical analysis

Data handling and plotting, principle co-ordinate analysis (PCO), and unweighted pair group method with arithmetic mean (UPGMA) clustering analysis were performed using R.<sup>14</sup> Arlequin v3.5.1.2<sup>15</sup> was used to: test for departures from Hardy–Weinberg equilibrium; estimate gene diversity ( $h$ ), assess pairwise differences in SNP and haplotype frequencies; and perform Mantel and AMOVA tests. PHASE v2.1<sup>16,17</sup> was used to estimate haplotype frequencies and impute missing data. 95% Confidence intervals (CIs) and significant differences for  $h$  were assessed using bootstrapping of populations within groups and permutation among groups, respectively (number of iterations = 10 000).

## RESULTS

### Single SNP analysis

Single SNP frequencies are shown in Supplementary Figure 1. Substantial variation in allele frequencies was observed across all populations with large ranges seen in minor allele frequency. Significant among-population variation was observed for all SNPs ( $P < 0.001$ ) as assessed by a frequency-based AMOVA framework. For three of the five SNPs, gene diversity (the probability of sampling two different alleles) was higher in European and North African populations (NSAH) than in sub-Saharan (SSAH) populations (mean

<sup>1</sup>UCL Centre for Cardiovascular Genetics, Rayne Institute, London, UK; <sup>2</sup>UCL Genetics Institute, University College London, London, UK; <sup>3</sup>Department of Evolutionary Biology, Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden; <sup>4</sup>Research Department of Genetics, Evolution and Environment, University College London, London, UK; <sup>5</sup>The Centre for Genetic Anthropology, UCL Genetics Institute, University College London, London, UK; <sup>6</sup>Addis Ababa University, Addis Ababa, Ethiopia; <sup>7</sup>Center of Human Genetic Diversity, Addis Ababa, Ethiopia; <sup>8</sup>Arizona Research Laboratories, Division of Biotechnology, University of Arizona, Tucson, AZ, USA  
\*Correspondence: Professor SE Humphries, UCL Centre for Cardiovascular Genetics, Rayne Institute, 5 University Street, London NW1 1TY, UK. Tel: +44 20 7679 6969; Fax: +44 20 7679 6211; E-mail: rmhaseh@ucl.ac.uk

Received 14 March 2011; revised 14 June 2012; accepted 18 July 2012; published online 29 August 2012

**Table 1 Summary of populations used, number of individuals (*n*), acronym and location**

Population		Collection location		
code	Description	Latitude (°)	Longitude (°)	n
<b>NSAH</b>				
EUR				
AYL	NPHSII Aylesbury	51.816	-0.808	113
CAM	NPHSII Camberley	51.335	-0.743	373
CHE	NPHSII Chesterfield	53.235	-1.422	209
HAL	NPHSII Halesworth	52.343	1.501	424
HAR	NPHSII Harefield	51.603	-0.481	247
NMY	NPHSII North Mymms	51.721	-0.205	292
PRK	NPHSII Parkstone	50.72	-1.945	397
CAR	NPHSII Carnoustie	56.503	-2.702	341
STA	NPHSII St Andrews	56.341	-2.798	300
ANT				
ANT	Anatolian Turks	39.929	32.853	95
ARB				
ARB	Mixed Arabians	15.409	44.242	37
YEH	Yemeni from Hadramaut	14.544	49.127	83
NAF				
ALG	Mixed Algerians	35.083	-2.233	182
BBR	Moroccan Berbers	33.588	-5.165	87
<b>SSAH</b>				
WAF				
FUL	Fulani from Cameroon	6.477	11.551	63
MAM	Mambilla from Cameroon	6.477	11.551	96
SAR	Shewa Arabs	12.076	15.032	90
BUL	Bulsa from Ghana	10.726	-1.279	90
IGB	Igbo from Nigeria	4.957	8.314	100
MJK	Manjak from Senegal	14.687	-17.452	85
WEA	Mixed West African populations	—	—	144
CEA				
EAM	Ethiopian Amhara	11.683	38.15	262
EAN	Ethiopian Anuak	8.243	34.593	95
EOR	Ethiopian Oromo	7.664	36.831	94
NSD	Northern Sudanese	19.17	30.474	146
SSD	Southern Sudanese	6.227	31.567	125
CEA	Mixed East African populations	—	—	26
SEA				
CHW	Chewa from Malawi	-13.983	33.774	92
TBK	Tumbuka from Malawi	-11.465	34.023	60
YAO	Yao from Malawi	-14.478	35.264	56
SEN	Mixed Mozambiquans	15.409	44.242	83
BAN	Bantu speakers from South Africa	-26.144	28.05	107
LMB	South African Lemba	-20.48	29.914	82
SEA	Mixed Southeast African populations	—	—	192

*h*: -9731 (NSAH = 0.476, SSAH = 0.446); -5848 (0.478, 0.495); +4860 (0.408, 0.251); +8855 (0.407, 0.282); +11015 (0.483, 0.498), and for all five SNPs mean *h* values for NSAH and SSAH were significantly different with non-overlapping 95% CIs (Supplementary Table 1).

### Haplotype frequency analysis

In total, 19 haplotypes were inferred from the five SNPs, with 9 haplotypes found at a frequency greater than 2% in at least one population (Figure 1 and Supplementary Table 2). hGCATA was the modal haplotype in all populations (frequency range 29.0–56.6%), and only two other haplotypes were identified in every population (hTTATC and hTTCAC, frequency ranges of 6.7–28.9% and 6.6–36.8%, respectively). Haplotype diversity was significantly ( $P < 0.01$ ) higher in NSAH ( $h = 0.733$ , 95% CI = 0.718–0.747) than SSAH ( $h = 0.703$ , 95% CI = 0.687–0.718, Supplementary Table 1), although the number of haplotypes were fewer. The higher diversity in NSAH is most likely due to the dominance of GCATA in SSAH (around 50% in many SSAH populations) and the resultant paucity of equally frequent haplotypes. AMOVA analysis indicated that there is significant among-population variation in haplotype frequency ( $F_{ST} = 2.2\%$ ,  $P < 0.001$ ). A haplotype associated with lower IL-18 protein levels, hGTATA, was found at a frequency of 26.6% in the UK populations, but was absent or at low frequency in sub-Saharan Africa (median frequency of hGTATA was 0.0% across sub-Saharan populations).

### Geographic distribution of IL18 variation

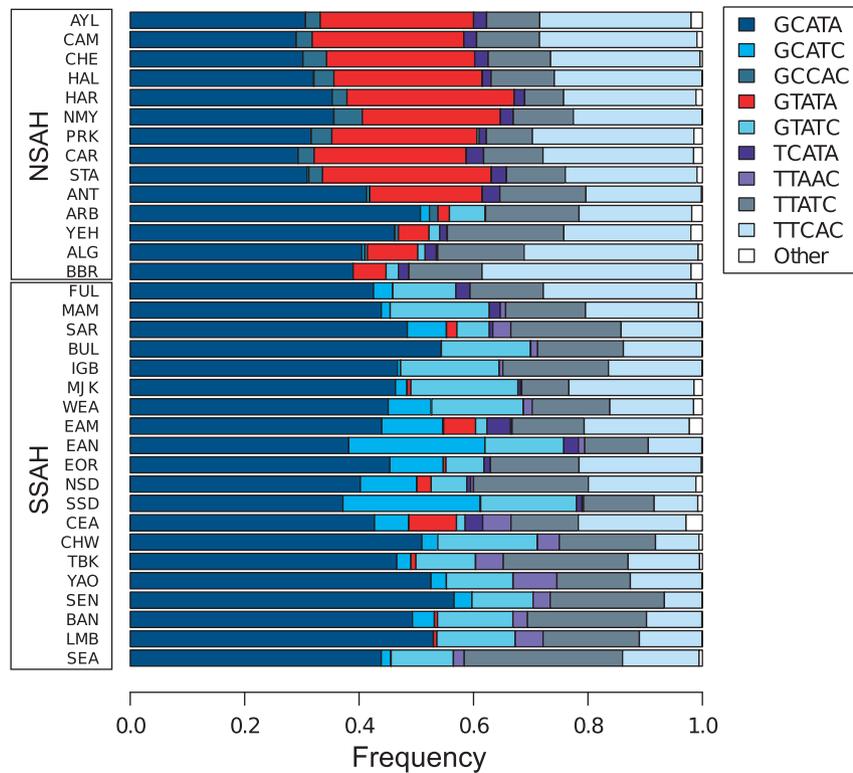
PCO of the population pairwise haplotype-based  $F_{ST}$  matrix showed the African populations forming a wide scatter while the European populations formed a more distinct cluster (Figure 2). UPGMA clustering analysis with  $K = 3$  identified clear population clusters. One (*Cluster\_1*) consisted of all European, North African and Anatolian Turkish groups, and one (*Cluster\_2*) contained both ARB populations and all SSAH populations except SSD and EAN, which together made up the last cluster (*Cluster\_3*) (see Supplementary Table 2 for haplotype frequencies within each cluster). The final cluster appears to be the result of a higher than average hGCATC frequency compared to other SSAH populations. Interestingly, the Anuak of Ethiopia (EAN) are believed to descend from a larger group from South Sudan (SSD). A Mantel test showed a significant ( $R^2 = 0.197$ ,  $P < 0.001$ ) correlation between pairwise genetic and geographic (great circle) distances.

### Relationship to published haplotypes

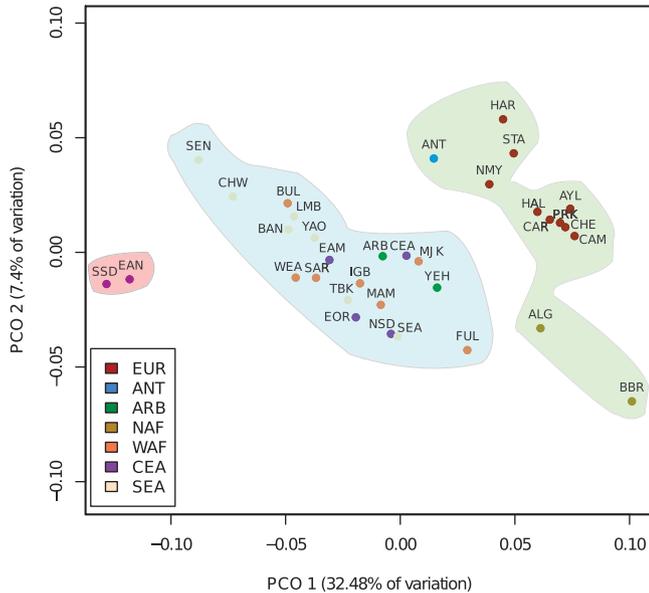
An IL18 haplotype (htirGCAGT) composed of five SNP alleles, of which two (-9731 G and +4860 A) are shared with hGTATA (Supplementary Table 3), has been shown to be associated with lower IL-18 protein<sup>8</sup> and mRNA levels<sup>11</sup> along with decreased CVD risk.<sup>8</sup> Given both htirGCAGT and hGTATA have been associated with lower levels of IL-18, we investigated whether the two haplotypes might exist on similar or the same genomic sequences. The three uncharacterised SNPs from htirGCAGT were genotyped in a set ( $n = 18$ ) of hGTATA homozygotes from the UK and a subset ( $n = 87$ ) of African samples, homozygous for various haplotypes. All UK samples were also homozygous for htirGCAGT, while in the African samples, six of the eight chromosomes known to carry hGTATA carried htirGCAGT (Supplementary Table 4). The two remaining chromosomes carried htirGCAAT and htirGCAGC. None of the African samples homozygous for non-hGTATA haplotypes carried htirGCAGT.

### DISCUSSION

The results from this study indicate that, although all studied SNP alleles were found at appreciable frequencies in all populations, the distribution of IL18 genetic variation differs greatly between Africa and Europe. The most striking example of this difference is the



**Figure 1** Common haplotype frequencies by population. See Table 1 for population codes.



**Figure 2** PCO plot of  $F_{ST}$  matrix derived from estimated *IL18* haplotypes in all populations. UPGMA clustering results for  $K=3$  are overlaid on PCO plot. *Cluster\_1*=green region, *Cluster\_2*=blue region, *Cluster\_3*=red region. Subsequent values of  $K$  were less informative, with the fourth and fifth clusters consisting of single populations within NSAH. See Table 1 for population codes.

here. Although a full exploration of the possible evolutionary mechanisms of this observation are beyond the scope this report, it is interesting to note that studies of other interleukin genes such as *IL-10* have suggested that natural selection may be acting on pro-inflammatory haplotypes within the peoples of sub-Saharan Africa and Caucasians.<sup>18</sup>

In addition, contrary to our expectations, we observed significantly higher gene diversity indices in the European populations when compared with the African populations at three tSNPs and at the haplotype level. It is likely that this is a result of the tSNPs used here being ascertained in European samples, and that with analysis of full resequencing data, we would observe higher genetic diversity in the African populations. Soon-to-be abundant data for *IL18* derived from large whole-exome sequencing projects will be particularly useful for understanding patterns of European and African *IL18* diversity.

We also found evidence that hGTATA is tagging the same gene-wide sequence as htirGCAGT, which has previously been found to be associated with lower levels of both IL-18 and CVD risk in European individuals,<sup>8</sup> which strongly suggests that the same potentially functional polymorphism(s) could be found on the background of both haplotypes. Although there was less correlation between the two haplotypes in African individuals, htirGCAGT was found only on chromosomes carrying hGTATA. Resequencing of this genomic region in association studies recording IL-18 levels in individuals homozygous for htirGCAGT/hGTATA in a search for causative polymorphisms could prove rewarding. IL-18 is an important inflammatory regulator, and as well as CVD has also been associated with autoimmunity,<sup>19</sup> appetite regulation/energy homeostasis<sup>20,21</sup> and protection from *Plasmodium falciparum* malaria.<sup>22–24</sup> As such, it is likely that functional genetic variation within *IL18* will influence disease susceptibility.<sup>7</sup>

complete absence of the second most common European haplotype, hGTATA, which is known to be associated with lower circulating IL-18 levels, in the vast majority of the sub-Saharan samples studied

We hope this relatively simple survey of genetic variation will motivate genotype:phenotype studies involving IL-18 that will focus specifically on cohorts of recent African ancestry (such studies are currently largely absent from the literature). African Americans, for example, are known to have high rates of CVD and our results suggest that such studies may be very beneficial for this group.

### CONFLICT OF INTEREST

Dr Neil Bradman is Chairman of the Centre for Genetic Anthropology (TCGA) and an Honorary Lecturer in the Research Department of Genetics, Evolution and Environment at University College London. He is also joint chairman of the London and City Group of Companies and has extensive business and financial interests including involvement in biotechnology ventures and educational material used by researchers in biomedicine and the life sciences. Nevertheless, he does not have any specific commercial interest in the subject matter of this study. The research has been funded in part by a charitable trust of which Dr Bradman is a trustee. The charitable trust has no intellectual property with respect to this research.

### ACKNOWLEDGEMENTS

SRT was supported by a British Heart Foundation studentship FS/04/039, and SEH is funded by the British Heart Foundation RG2008/08. We acknowledge the contribution of the late Professor George Miller (1939–2006) who was the PI on the NPHSII study. The NPHSII study was supported by the Medical Research Council, the US National Institutes of Health (NHIBI 33014) and the Du Pont Pharma. Sample analysis was supported by a grant to the Centre for Genetic Anthropology by Melford Charitable Trust. We also thank all staff and volunteers who contributed to the sample collection.

- 1 Dinarello CA: Interleukin-18. *Methods* 1999; **19**: 121–132.
- 2 Dinarello CA, Fantuzzi G: Interleukin-18 and host defense against infection. *J Infect Dis* 2003; **187**(Suppl 2): S370–S384.
- 3 Thompson SR, McCaskie PA, Beilby JP *et al*: IL18 haplotypes are associated with serum IL-18 concentrations in a population-based study and a cohort of individuals with premature coronary heart disease. *Clin Chem* 2007; **53**: 2078–2085.
- 4 Thompson SR, Novick D, Stock CJ *et al*: Free interleukin (IL)-18 levels, and the impact of IL18 and IL18BP genetic variation, in CHD patients and healthy men. *Arterioscler Thromb Vasc Biol* 2007; **27**: 2743–2749.

- 5 He M, Cornelis MC, Kraft P *et al*: Genome-wide association study identifies variants at the IL18-BCO2 locus associated with interleukin-18 levels. *Arterioscler Thromb Vasc Biol* 2010; **30**: 885–890.
- 6 Presta I, Andreozzi F, Succurro E *et al*: IL-18 gene polymorphism and metabolic syndrome. *Nutr Metab Cardiovasc Dis* 2009; **19**: e5–e6.
- 7 Thompson SR, Humphries SE: Interleukin-18 genetics and inflammatory disease susceptibility. *Genes Immun* 2007; **8**: 91–99.
- 8 Tiret L, Godefroy T, Lubos E *et al*: Genetic analysis of the interleukin-18 system highlights the role of the interleukin-18 gene in cardiovascular disease. *Circulation* 2005; **112**: 643–650.
- 9 Liu W, Tang Q, Jiang H *et al*: Promoter polymorphism of interleukin-18 in angiographically proven coronary artery disease. *Angiology* 2009; **60**: 180–185.
- 10 Arimitsu J, Hirano T, Higa S *et al*: IL-18 gene polymorphisms affect IL-18 production capability by monocytes. *Biochem Biophys Res Commun* 2006; **342**: 1413–1416.
- 11 Barboux S, Poirier O, Godefroy T *et al*: Differential haplotypic expression of the interleukin-18 gene. *Eur J Hum Genet* 2007; **15**: 856–863.
- 12 Miller GJ, Bauer KA, Barzegar S, Cooper JA, Rosenberg RD: Increased activation of the haemostatic system in men at high risk of fatal coronary heart disease. *Thromb Haemost* 1996; **75**: 767–771.
- 13 Thompson SR, Sanders J, Stephens JW, Miller GJ, Humphries SE: A common interleukin 18 haplotype is associated with higher body mass index in subjects with diabetes and coronary heart disease. *Metabolism* 2007; **56**: 662–669.
- 14 R Development Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing, 2011.
- 15 Excoffier L, Lischer HEL: Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 2010; **10**: 564–567. PMID: 21565059.
- 16 Stephens M, Smith NJ, Donnelly P: A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 2001; **68**: 978–989.
- 17 Stephens M, Scheet P: Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *Am J Hum Genet* 2005; **76**: 449–462.
- 18 Kuningas M, May L, Tamm R *et al*: Selection for genetic variation inducing pro-inflammatory responses under adverse environmental conditions in a Ghanaian population. *PLoS One* 2009; **4**: e7795.
- 19 Boraschi D, Dinarello CA: IL-18 in autoimmunity: review. *Eur Cytokine Netw* 2006; **17**: 224–252.
- 20 Netea MG, Joosten LA, Lewis E *et al*: Deficiency of interleukin-18 in mice leads to hyperphagia, obesity and insulin resistance. *Nat Med* 2006; **12**: 650–656.
- 21 Zorrilla EP, Sanchez-Alavez M, Sugama S *et al*: Interleukin-18 controls energy homeostasis by suppressing appetite and feed efficiency. *Proc Natl Acad Sci USA* 2007; **104**: 11097–11102.
- 22 Cramer JP, Lepenies B, Kamena F *et al*: MyD88/IL-18-dependent pathways rather than TLRs control early parasitaemia in non-lethal Plasmodium yoelii infection. *Microbes Infect* 2008; **10**: 1259–1265.
- 23 Kojima S, Nagamine Y, Hayano M, Looareesuwan S, Nakanishi K: A potential role of interleukin 18 in severe falciparum malaria. *Acta Trop* 2004; **89**: 279–284.
- 24 Torcia MG, Santarlasci V, Cosmi L *et al*: Functional deficit of T regulatory cells in Fulani, an ethnic group with low susceptibility to Plasmodium falciparum malaria. *Proc Natl Acad Sci USA* 2008; **105**: 646–651.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (<http://www.nature.com/ejhg>)