Widespread seasonal gene expression reveals annual differences in human immunity and physiology

Xaquin Castro Dopico1,*, Marina Evangelou1,*, Ricardo C. Ferreira1, Hui Guo1, Marcin L. Pekalski1, Deborah J. Smyth1, Nicholas Cooper1, Oliver S. Burren1, Anthony J. Fulford2, Branwen J. Hennig2, Andrew M. Prentice2, Anette-G. Ziegler3, Ezio Bonifacio4, Chris Wallace1,5,** & John A. Todd1,**

Seasonal variations are rarely considered a contributing component to human tissue function or health, although many diseases and physiological process display annual periodicities. Here we find more than 4,000 protein-coding mRNAs in white blood cells and adipose tissue to have seasonal expression profiles, with inverted patterns observed between Europe and Oceania. We also find the cellular composition of blood to vary by season, and these changes, which differ between the United Kingdom and The Gambia, could explain the gene expression periodicity. With regards to tissue function, the immune system has a profound pro-inflammatory transcriptomic profile during European winter, with increased levels of soluble IL-6 receptor and C-reactive protein, risk biomarkers for cardiovascular, psychiatric and autoimmune diseases that have peak incidences in winter. Circannual rhythms thus require further exploration as contributors to various aspects of human physiology and disease.

1 JDRF/Wellcome Trust Diabetes and Inflammation Laboratory, Department of Medical Genetics, NIHR Cambridge Biomedical Research Centre, Cambridge Institute for Medical Research, University of Cambridge, Wellcome Trust/MRC Building, Cambridge Biomedical Campus, Cambridge CB2 0XY, UK. 2 MRC International Nutrition Group at MRC Unit The Gambia & London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, UK. 3 Institute of Diabetes Research, Helmholtz Zentrum München, Neuherberg, Forschergruppe Diabetes, Klinikum rechts der Isar, Technische Universität München, Ingolstaedter Landstr. 1, D 85764 Neuherberg, Germany. 4 CRTD—DFG Research Center for Regenerative Therapies Dresden, Paul Langerhans Institute Dresden, Medical Faculty, Technische Universität Dresden, Fetscherstrasse, 01307 Dresden, Germany. 5 MRC Biostatistics Unit, Cambridge Institute of Public Health, Forvie Site, Robinson Way, Cambridge Biomedical Campus, Cambridge CB2 0SR, UK. * These authors contributed equally to this work. ** These authors jointly supervised this work. Correspondence and requests for materials should be addressed to X.C.D. (email: xaquin@gmail.com) or to J.A.T. (email: john.todd@cimr.cam.ac.uk).
Periodic seasonal changes have influenced all life forms, as exemplified by seasonal physiology and behaviours across plant and animal species. For example, reptile graft rejection and level of gonadal hormones in squirrel monkeys display seasonal variation. In humans, many complex polygenic diseases, including cardiovascular, autoimmune and psychiatric illnesses, have established seasonal patterns of incidence and disease activity. Infectious disease seasonality is well established in humans, and it has been proposed that an inborn physiological rhythm underlies the seasonality of diagnoses of infectious diseases and their pathologies, but direct evidence of such a system is lacking.

Various biological processes show seasonal variation in humans, including ones with important immunological roles, such as vitamin D metabolism. The loss of skin pigmentation in humans, including ones with important immunological roles, is an example of the evolutionary adaption of humans to different environments. Yet, how seasons might more broadly impact the underlying molecular details of human physiology is unknown. Along these lines, we hypothesized that the anti-inflammatory circadian transcription factor, ARNTL (BMAL1), would display seasonal gene expression differences as daylight entrains circadian rhythms in mammals. Tissue-specific molecular clocks control a diverse range of cellular processes, influencing the immune response.

From ethnically and geographically diverse populations, we analysed mRNA expression levels in peripheral blood mononuclear cells and adipose tissue biopsies, full blood count data, and the circulating levels of inflammatory protein biomarkers.

Results
Seasonal ARNTL expression in the immune system. We first analysed ARNTL expression in peripheral blood mononuclear cells (PBMCs) from children (454 samples from 109 individuals) enrolled into the BABYDIET cohort from Germany (Supplementary Table 1). ARNTL mRNA showed seasonal variation in expression (ANOVA, $\chi^2 = 1.04 \times 10^{-2}$), peaking in the summer months of June, July and August (Fig. 1a). The difference between the winter low and summer high in ARNTL expression was 1.5097-fold. Vitamin D receptor (VDR) expression was also higher in the summer months (Fig. 1a). The housekeeping genes, B2M and GAPDH, often used as standards in gene expression analyses, did not show seasonal variation (Fig. 1a). ARNTL showed the same seasonal expression profile independently of whether blood was drawn during morning or afternoon clinic visits (Fig. 1b), suggesting that diurnal oscillations are not responsible for the seasonal differences in ARNTL expression.

We then sought evidence for components of the circadian clock. Seasonal variation was found in 9 of the 16 clock genes tested: ARNTL, CLOCK, CRY1, CSNK1D, CSNK1E, NR1D2, RORA, TIMELESS and NFIL3 (which controls circadian Th17 cell development in mice) (Fig. 1c). Seven genes (CRY2, PER3, RORB, Npas2, PER1, PER2 and NR1D1) did not show evidence for seasonal effects (Supplementary Table 3). Novel components of the human circadian clock, as well as clock-targeted genes and pathways, are likely to be present among the genes whose expression correlated with ARNTL (Supplementary Table 2). Interestingly, the glucocorticoid receptor (NR3C1) had a strong positive correlation with ARNTL (Spearman $\rho = 0.819$), with lowest expression in the winter (ANOVA, $\chi^2 = 5.05 \times 10^{-19}$) (Fig. 1d). Glucocorticoids have anti-inflammatory properties and SCN-controlled hormones are thought to be essential molecules for maintaining the synchronicity of peripheral biological clocks. In contrast to NR3C1, receptors for the prostaglandins (PTGDR, PTGIR and PTGER4), leukotrienes (CYSLTR1) and oxoecosanoids (OXER1) were more highly expressed in the winter in Germany. Receptors for adipsinectin (ADIPOR1), estradiol (ESR2) and antidiuretic hormone (CUL5) were more highly expressed in the summer (Fig. 1d). Other hormone receptors did not show any seasonal variation in this data set.

Widespread seasonal gene expression in the immune system. Strikingly, we found ~23% of the genome (5,136 unique genes out of 22,822 genes tested) to show significant seasonal differences in expression in the BABYDIET data set (Fig. 2a and Supplementary Table 3). Among the seasonal genes, two distinct anti-phasic patterns of gene expression were evident: 3,11 genes (2,922 unique probes) had increased expression in the summer (defined as June, July and August, mean fold change = 1.2572) while 2,826 genes (3,436 unique probes) were upregulated in the winter (defined as December, January, February, mean fold change = 1.3150) (Fig. 2c, Supplementary Fig. 1), demonstrating that different transcriptional landscapes are present in the peripheral immune system during different seasons.

The daily variables of mean ambient temperature and mean sunlight hours both served as linear predictors of seasonality (Supplementary Fig. 2), suggestive of human environmental adaptation.

We replicated the observation in two independent data sets. First, in a collection of PBMCs isolated from autoimmune type 1 diabetes (T1D) patients (236 samples) from the United Kingdom, 1,697 genes were found to exhibit seasonal expression (Fig. 3a and Supplementary Table 4). The majority of seasonally associated transcripts could again be identified as having summer or winter expression profiles, with seasonal patterns matching those identified in the BABYDIET data set (Fig. 3b). This data set demonstrated seasonal gene expression in adults, adding to the observations made in samples from children (Supplementary Fig. 3 and Fig. 2).

Secondly, we analysed gene expression data from a collection of PBMCs from adult (18–83 years old, mean age 45) asthmatic individuals from diverse ethnic groups across Australia, United Kingdom/Ireland, United States and Iceland. We separated the entire cohort into distinct geographical locations and observed seasonal gene expression in each (Fig. 3c and Supplementary Tables 5–8). Seasonal genes identified in the BABYDIET cohort maintained their seasonal tropisms in the asthmatic patients (Fig. 3d). Most interestingly, in the Australian data set, the previously defined summer genes (increased in expression during the Northern hemisphere summer) were more highly expressed during the Southern hemisphere summer, spanning December, January and February (Fig. 3d); clearly illustrated, for example, by ARNTL expression (Supplementary Fig. 6B). The pattern of seasonal gene expression in samples from Iceland was unique (Supplementary Fig. 4).

Seasonal gene expression was not altered in samples from children with self-reported infections (Supplementary Fig. 5), and the type-1 interferon response gene, SIGLEC13, was not seasonal. Finally, recruitment into the asthma cohort was dependent on participants being free from infectious diseases. Nevertheless, the relationship between seasonal infections and diseases, and these seasonal gene expression patterns, remains to be fully described.

Common seasonal genes in the immune system. One hundred and forty-seven genes showed common seasonality in the BABYDIET, T1D, Australia, USA and UK/Ireland datasets (Supplementary Table 9). These 147 genes had similar seasonal expression patterns in each cohort (Supplementary Fig. 6).
Figure 1 | Seasonal mRNA expression in the peripheral human immune system. Relative expression profiles of seasonal genes (fitted values of the cosinor model). (a) ARNTL expression was increased in the summer months of June, July and August (ANOVA, $\chi^2_{23} = 1.04 \times 10^{-23}$), compared with the winter months of November through February ($1.5097$-fold difference between February and August ($n = 109$ individuals)). Similarly, the nuclear vitamin D receptor (VDR) shows peak expression in June through August (ANOVA, $\chi^2_{26} = 1.62 \times 10^{-06}$). The housekeeping genes, B2M and GAPDH, did not have seasonal expression profiles. (b) Seasonal ARNTL expression in PBMCs independent of the circadian phase. Similar seasonal ARNTL expression profiles were observed regardless of whether blood samples were collected during morning (BABYDIET, $n = 109$ individuals) or afternoon clinic visits (T1D cohort, $n = 236$ individuals). (c) In the BABYDIET data set, nine known components of the circadian clock had seasonal expression profiles in the peripheral immune system, as did certain hormone, leukotriene and prostaglandin receptors. (d) The receptors for the anti-inflammatory glucocorticoids (NR3C1) and the pro-inflammatory prostaglandins (PTGDR, PTGIR and PTGER4) and leukotrienes (CYSLTR1) had opposing seasonal expression profiles.
Notably, in the Icelandic cohort, the common seasonal genes did not share the same expression pattern (Supplementary Fig. 4). This could be due to near-24-h daylight during summer if seasonal human physiology is regulated by changes in the annual photoperiod\textsuperscript{36}. ARNTL was found to be a common seasonal gene (Fisher’s method, $\chi^2_{10}, P = 6.73 \times 10^{-57}$), with increased summer expression in each PBMC data set, except Iceland. The gene with the strongest seasonal profile common to all data sets (excluding Iceland) was C14orf159 (winter expressed, Fisher’s method, $\chi^2_{10}, P = 3.93 \times 10^{-66}$). The mitochondrial protein, UPF0317, encoded by C14orf159 (whose expression is regulated by oestrogen receptor alpha\textsuperscript{37}) is highly conserved in chordates, although its function in humans is largely unknown.

**Seasonal cellular remodelling of the human immune system.** As PBMCs represent several specialized haematopoietic lineages, we sought to determine whether seasonal gene expression resulted from annual changes in the cellular composition of blood. In support of this, we found the expression of seasonal genes to correlate strongly with the expression of 13 genes known to mark different immune cell types present in PBMCs\textsuperscript{38} (Fig. 4a). Furthermore, by analysing full blood count (FBC) data from 7,343 healthy adult donors enrolled in the Cambridge BioResource (United Kingdom), we found the total number of white blood cells (ANOVA, F-test, $P = 1.75 \times 10^{-10}$), lymphocytes (ANOVA, F-test, $P = 2.11 \times 10^{-11}$), monocytes (ANOVA, F-test, $P = 9.14 \times 10^{-30}$), basophils (ANOVA, F-test, $P = 2.74 \times 10^{-6}$), eosinophils (ANOVA, F-test, $P = 0.00235$), neutrophils (ANOVA, F-test, $P = 6.13 \times 10^{-27}$) and platelets (ANOVA, F-test, $P = 2.02 \times 10^{-12}$) to exhibit seasonality in the peripheral circulation, as did the mean corpuscular volume (MCV) (ANOVA, F-test, $P = 1.32 \times 10^{-21}$) and mean corpuscular...
haemoglobin (MCH) (ANOVA, F-test, $P = 1.73 \times 10^{-15}$) of erythrocytes (Fig. 4b). Our results are in agreement with a study that reported seasonal red blood cell and platelet gene expression.

In a more equatorial cohort, comprising 4,200 healthy individuals from The Gambia (West Africa), we observed seasonal variation in the number of total white blood cells (F-test, $P = 0.011$), lymphocytes (F-test, $P = 1.40 \times 10^{-05}$), monocytes (F-test, $P = 8.71 \times 10^{-16}$) and platelets (F-test, $P = 2.07 \times 10^{-18}$) (Fig. 4c), but not granulocytes. We also observed striking seasonal variation in red blood cell numbers (F-test, $P = 8.43 \times 10^{-30}$) and their mean corpuscular haemoglobin (F-test, $P = 4.07 \times 10^{-30}$) (Supplementary Fig. 7). The seasonal patterns in The Gambia were completely distinct to those observed in the UK cohort. In The Gambian cohort, the numbers of all seasonal cell types peaked during the rainy season.

Figure 3 | Seasonal gene expression in geographically distinct cohorts. (a) Seasonality was also observed in PBMCs collected from T1D patients in the United Kingdom ($n = 236$ individuals). A total of 1,697 genes were seasonal in this data set. (b) The previously defined summer and winter genes from the BABYDIET data set maintained their seasonal expression patterns in the T1D samples. (c) PBMCs from asthmatic patients collected from different countries also showed seasonal gene expression. In the United Kingdom/Ireland ($n = 26$ asthmatic individuals; 85 PBMC samples), 791 genes were seasonal, while 1,257 and 409 genes were seasonal in Australia ($n = 26$ individuals; 85 samples) and United States ($n = 37$ individuals; 123 samples), respectively. (d) Summer and winter BABYDIET genes maintained their seasonal expression patterns in the asthmatic PBMC samples, with their patterns inverted in Australia.
Figure 4 | Seasonal changes in the cellular composition of human peripheral blood. (a) Expression levels of 13 genes that have been used to identify different blood cell types among total PBMCs were strongly correlated (positively and negatively) with seasonal genes identified in the BABYDIET data set. In comparison, non-seasonal genes were less correlated with these marker genes, although exceptions exist: CTLA-4 expression also correlated with seasonal genes identified in the BABYDIET data set. (b) Indeed, by analysing full blood count data obtained from 7,343 healthy adult donors enrolled in the Cambridge BioResource, we found the cellular composition, and other haematological parameters of blood to vary by season. HCT was the only response that did not show seasonal variation. (c) Distinct seasonal variation in cell counts was observed in a cohort of 4,200 healthy adults and children from The Gambia. EOS, eosinophils; LYM, lymphocytes; NEU, neutrophils; PLT, platelets; RBC, red blood cells; WBC, total white blood cells; BAS, basophils; HGB, haemoglobin; MCH, mean corpuscular haemoglobin; MCV, mean corpuscular volume; MON, monocytes; HCT, haematocrit.
Seasonal differences in human immunity. To address whether immunological function varies seasonally, as suggested by the transcriptomic and cell count data, we generated modules of co-regulated seasonal mRNAs identified in the BABYDIET data set (Fig. 5a and Supplementary Fig. 8 and Supplementary Table 12). Among the seven winter-expressed modules we identified, we found pro-inflammatory processes to be more frequent, compared with the identified summer-expressed modules. B-cell receptor (BCR) signalling (Hypergeometric test, \( P = 3.39 \times 10^{-16} \)), FcR-gamma-associated processes (Hypergeometric test, \( P = 4.45 \times 10^{-6} \)), lysosomes (Hypergeometric test, \( P = 2.96 \times 10^{-5} \)), chemokine signalling (Hypergeometric test, \( P = 3.56 \times 10^{-5} \)) and phagosomes (Hypergeometric test, \( P = 5.97 \times 10^{-5} \)) were all strongly associated with winter-expressed modules. In contrast, RNA transport (Hypergeometric test, \( P = 1.70 \times 10^{-7} \)), RNA degradation (Hypergeometric test, \( P = 1.02 \times 10^{-5} \)), ubiquitin-mediated proteolysis (Hypergeometric test, \( P = 0.0002 \)), circadian rhythms in mammals (Hypergeometric test, \( P = 0.0011 \)) and splicosome (Hypergeometric test, \( P = 0.0014 \)) were the most-associated pathways with summer-expressed modules, suggesting that a more inflammatory status of the immune system predominates in winter (Fig. 5b and Supplementary Fig. 8).

In further support of this, we found the concentration of sIL-6R protein to be increased in winter in samples from BABYDIET and BABYDIAB (a related collection\(^{42}\), ANOVA, \( \chi^2_3; P = 2.74 \times 10^{-11} \)), in complete agreement with the increased winter expression of IL6R mRNA in BABYDIET samples (ANOVA, \( \chi^2_3; P = 9.33 \times 10^{-12} \)) (Fig. 5c). sIL-6R is an important orchestrator of leukocyte recruitment\(^{43}\) and trans-presents IL-6 to cells expressing gp130 in the absence of the cell-surface IL-6R\(^{44}\), endowing IL-6 with a broader spectrum of influence. Indeed, a coding variant in IL6R that alters circulating sIL-6R concentration is associated with impaired IL-6 signalling and the protection from cardiovascular disease, rheumatoid arthritis and T1D\(^{45}\). Interestingly, early-stage inflammation in rheumatoid arthritis (a disease treated with anti-IL-6 receptor reagents\(^{46}\)) has been shown to either resolve or progress to erosive disease, and a predictor of this outcome is the season when disease symptoms first present\(^{47}\). T1D also has seasonal trends in diagnoses\(^9\) and autoantibody positivity\(^{48}\), suggesting that seasonal environments impact on autoimmune disease pathologies. Furthermore, we found the circulating level of the acute-phase complement activator, C-reactive protein\(^{49}\), to be increased during winter months (Fig. 5d).

These gene expression data also suggest that the quality of a vaccine response may be influenced by season. We found the expression of TLR7 (ANOVA, \( \chi^2_3; P = 4.22 \times 10^{-16} \)), TLR8 (ANOVA, \( \chi^2_3; P = 3.70 \times 10^{-10} \)) and DDX58 (encoding the viral RNA receptor RIG-I, ANOVA, \( \chi^2_3; P = 9.65 \times 10^{-20} \)) to have increased in expression in winter months in the BABYDIET data set: the increased expression of these genes correlated with protective immunity in response to the Yellow Fever vaccine (YF-17D)\(^{50}\). TNFRSF17 also showed seasonal variation in the BABYDIET data set (ANOVA, \( \chi^2_3; P = 1.30 \times 10^{-12} \)), and its induction is shown to be predictive of an antibody response after trivial influenza vaccine\(^{51,52}\). Furthermore, OAS1 (ANOVA, \( \chi^2_3; P = 1.43 \times 10^{-20} \)), OAS2 (ANOVA, \( \chi^2_3; P = 2.85 \times 10^{-16} \)), STAT2 (ANOVA, \( \chi^2_3; P = 1.76 \times 10^{-18} \)), POU2AF1 (ANOVA, \( \chi^2_3; P = 3.27 \times 10^{-15} \)) and CD27 (ANOVA, \( \chi^2_3; P = 3.29 \times 10^{-16} \)) were also seasonal and their expression in PBMCs was correlated with increased anti-DT IgG responses after the meningococcal vaccine (MCV4)\(^{53}\). The antibody responses to rabies, typhoid and pneumococcal vaccines are influenced by the month of vaccine administration\(^{50}\).

Seasonal gene expression in subcutaneous adipose tissue. Given the remarkable seasonality of the peripheral immune system and the correlations we found with multiple health-associated phenotypes, we anticipated that tissues throughout the body would display extensive seasonality of gene expression. We were able to analyse gene expression data from a collection of subcutaneous adipose tissue samples obtained from 825 healthy female donors enrolled in the TwinsUK cohort\(^{53}\).

We found 4,027 genes to be seasonally expressed (Fig. 6, Supplementary Table 13), including IL6ST (gp130) (ANOVA, \( \chi^2_3; P = 2.55 \times 10^{-8} \)) and IL6R (ANOVA, \( \chi^2_3; P = 1.49 \times 10^{-8} \)); adipose tissue can produce IL-6\(^{54}\). One thousand, two hundred and thirteen genes were common to both adipose tissue and BABYDIET data sets (Supplementary Table 14), suggesting that common genetic mechanisms regulate seasonality.

In adipose tissue, as in PBMCs, metabolic pathways were among the most associated seasonal pathways (Supplementary Fig. 9). Such seasonal metabolic programmes may have been selected for due to annual differences in temperature and diet. Adipose tissue seasonality has important implications for immunology, obesity and metabolic disease research; for example, PPARG, targeted by thiazolidinediones as a current treatment of type 2 diabetes, was found to be seasonal in adipose tissue (ANOVA, \( \chi^2_3; P = 3.75 \times 10^{-9} \)).

Discussion

Ecological changes alter the types and dynamics of inter- and intra-organism biological processes, and it follows that such changes will be manifested as seasonal transcriptional signatures within the immune systems of different organisms, which adapt to their environment. Studies of the function, dynamics and variability of the immune system are undergoing a long-awaited renaissance partly owing to the development and application of new phenotyping technologies\(^{55,56}\). Nevertheless, to date, no study to our knowledge has taken into account the variability we have observed in the immune system according to season, which could, for example, increase the differences in some immune phenotypes between twins or other family members if blood samples were collected at different times of year. We observed seasonal differences in expression across a large number of genes in mixed populations of human peripheral white blood cells from geographically and ethnically diverse locations, and, remarkably, seasonal genes displayed opposing patterns in the Southern and Northern hemispheres. Fewer seasonal genes were identified in Icelandic donors, and common seasonal genes had a less similar seasonal pattern in this data set. If a seasonal photoperiodic clock exists in humans, the impact of living at higher latitudes requires further exploration.

These periodically changing transcriptional landscapes in PBMCs, which appear to be predominantly driven by annual changes in the cellular composition of blood, are likely to influence various aspects of the human immune response. Indeed, the increased winter expression of co-regulated pro-inflammatory gene modules, the functionally important increased concentration of sIL-6R and CRP in the blood, and the observation that a loss of BMI1 (ARNL1 was reduced in winter) promotes inflammation in mice\(^{28}\), strongly suggests that the immune system is more pro-inflammatory in Europeans during the northern hemisphere.
Figure 5 | Inflammatory responses predominate the immune system in Europe. (a) Co-regulated seasonal gene modules were generated to analyse differences in immune function by season: eight winter modules and three summer modules were generated. (b) Two modules of seasonally co-regulated genes from the BABYDIET data set are shown as examples. A module consisting of genes involved in B-cell receptor signalling, (including CR2, BLNK, BTK, FCGR2B, CD72, CD79B) was more highly expressed in the winter, as was a module associated with metabolic processes. In contrast, a RNA-processing module (containing RANBP2, EIF3J, NUP54, DDX20, STRAP, NUPL1, PAIP1) was more highly expressed in the summer. (c) IL6R mRNA expression was increased in the winter, in BABYDIET samples (ANOVA, $\chi^2 = 9.33 \times 10^{-12}$), as was observed for the circulating level of sIL-6R protein in the serum of BABYDIET/DIAB children (ANOVA, $\chi^2 = 2.74 \times 10^{-11}$). (d) The circulating levels of C-reactive protein displayed seasonal variation in a cohort of 3,412 donors diagnosed as hypertensive but not conventionally dyslipidemic. ASCOT enrolled participants in Ireland, Denmark, Finland, Iceland, Norway, Sweden and the UK (two measurements per donor), with increased levels present during winter HSCRP - high sensitivity C-reactive protein.
winter. In mice, Arntl-BMAL1 controls the diurnal variation of circulating and tissue-resident inflammatory monocyte numbers, although how ARNTL controls human immune function is not known. We note that, in Europeans, total monocyte numbers in blood are increased during winter, when ARNTL expression is the lowest. Notably, acute-phase proteins including CRP are induced by IL-6, which can be produced by macrophages and adipocytes. This entire network could be a major factor in the higher frequency of cardiovascular disease-associated deaths in winter, when increased risk is associated with excessive inflammation, IL-6 and monocytes. Furthermore, increased IL-6 signalling is associated with increased risk of rheumatoid arthritis and type 1 diabetes, which peaks in incidence during the European winter. Increased IL-6 signalling and elevated CRP levels have also been associated with neuropsychiatric symptoms in children and adults. Thus, modulation of IL-6 signalling according to season could be considered as a therapeutic strategy in various disease contexts. Whether a seasonal human immune system contributes to host-mediated pathology and morbidity after infection remains to be determined, but the correlations we report suggest this might be the case.

The breadth and functional characteristics of the seasonal gene expression we observed suggest that it has been evolutionarily selected for. During European winters, the thresholds required to trigger an immune response may be lower as a direct consequence of our co-evolution with infectious organisms and increased inter-species competition during winter, especially as humans migrated out of Africa to colder, more seasonally pronounced latitudes. In our European cohorts, winter was associated with increased monocytes and inflammation, while FBC data from the more-equatorial Gambian cohort exhibited distinct seasonal variation in cell numbers. In this data set, seasonal peaks in cell numbers correlated with the rainy season (June to October), during which time the infectious disease burden is at its highest levels.

Regardless of any particular causal factor driving these differences, which are likely many, our results demonstrate that different human populations independently vary the cellular composition of their immune system by season, suggestive of distinct environmental adaptations. Furthermore, although our data suggest that cell-type numbers contribute the majority of seasonal gene expression in PBMCs, future studies of seasonal phenotypic differences within purified immune cell subsets are likely to reveal an additional layer of complexity in the human immune system.

The origin and likely diverse mechanisms maintaining seasonal variation remain to be established: daylight and ambient temperature are candidate environmental cues that could coordinate seasonal hormonal phenotypes and cell-fate decisions in haematopoeitic and stem cells. Indeed, diurnal entrainment of the human circadian clock requires daylight changes, demonstrating that humans sense and process photoperiodic cues to co-ordinate physiology.

The environmental perturbation of our molecular clocks is thought to be deleterious to health, which may help explaining the increasing complex disease burden in industrialized countries and populations at extreme latitudes, where clock dysregulation or chronodisruption may be more frequent. In seasonally-breeding mammals, circadian melatonin production cues reproduction in response to changes in the annual photoperiod. In the arctic mammal, Rangifer tarandus, daily melatonin rhythms are acutely responsive to the night-day phase but not the circadian phase, demonstrating species-specific adaptation to the unique night-day cycles present at extreme latitudes: the ability of humans to properly function in such environments is not well understood. Furthermore, a circannual molecular clock was recently shown to control seasonal reproduction in hamsters, independently of melatonin and sex steroids, yet using the same neuroendocrine reproductive pathway. Human genetic variation in the ARNTL gene region has been associated with age of menarche, which is also seasonal.

The widespread seasonal gene expression observed in subcutaneous adipose demonstrates seasonality across different human tissues.

Regardless of the mechanisms causing and maintaining these and other seasonal variations, our results provide a plausible mechanism to explain part of the seasonality of human disease. These data provide a fundamental shift in how we conceptualize immunity in humans, and we propose that seasonal changes be more broadly considered as major determinants of human physiology.

Methods

Study subjects and human samples. All samples and information were collected with written and signed informed consent. One hundred and nine children genetically predisposed to T1D were enrolled in the BABYDIET study. The BABYDIET study is an intensively monitored dietary intervention study testing the potential effect of delayed gluten exposure on the development of islet autoimmunity in children at increased risk for diabetes in Germany. Children younger than 3 months with at least one first-degree relative with T1D and one of three specific T1D-associated HLA genotypes were recruited between 2000 and 2006 (participation rate: 88.8 %) and randomized to exposure to dietary gluten from age 6 months or from age 12 months. After inclusion, children were followed in three monthly intervals until the age of 3 years and yearly thereafter for efficacy (participation rate: 88.8 %) and randomized to exposure to dietary gluten from age 6 months or from age 12 months. After inclusion, children were followed in three monthly intervals until the age of 3 years and yearly thereafter for efficacy (participation rate: 88.8 %).
Self-reported infections in BABYDIET cohort. At each visit, parents of BABYDIET children completed a detailed questionnaire on their children’s history of infections, fever and medication. Specifically, they were asked about fever, infectious symptoms (such as diarrhea, vomiting, constipation and allergies) and the name of administered pharmaceutical agents or their active ingredient with starting date and duration of infections and medication. Infectious disease was defined as an acute event according to the ICD-710 Code or by a symptom indicating an infectious genesis. Infectious events were assigned to a specific time interval by their date of onset, and infectious events that could be matched to microarray samples were included for analysis, as described. Other disease events such as allergies or accidents were not considered as infectious diseases.

Soluble IL-6 receptor ELISA. Circulating sIL-6R concentrations were measured in BABYDIET and BABYDIAB serum samples using a highly sensitive non-isotopic time-resolved fluorescence ELISA assay based on the dissociation-enhanced lanthanide fluorescent immunoassay technology (DELFIA; PerkinElmer), as described. Test samples were diluted 1:20 in PBS + 10% FBS and measured in duplicate on 384-well Maxisorp microtiter plates (Nunc), coated with 1 μg ml⁻¹ monoclonal anti-human IL-6R antibody (clone 175706; RD Systems). Detection was performed using a biotinylated anti-C1D12 monoclonal antibody (clone M182, BD Biosciences) diluted to a final concentration of 100 ng ml⁻¹ in PBS + 10% FBS and a Europium-Streptavidin detection solution (PerkinElmer), diluted in PBS + 0.05% tween, 1% BSA, 7 μg ml⁻¹ DTPA to a final concentration of 0.05 μg ml⁻¹. Quantification of test samples was obtained by fitting the readings to a human recombinant IL-6Rz (BD systems) serial dilution standard curve plotted in quadruplicate on each plate. Data for 782 unique individuals existed from 722 families.

Cambridge BioResource full blood count data (UK cohort). Full blood count data were obtained from the Cambridge BioResource. BioResource volunteers are subjected to a full blood count on the day of blood sample collection using Beckman Coulter LH780, Beckman Coulter DXH800 5 part diff analyser or a Sysmex 5 part diff analyser. The available months of bleed were from February to November (no FBC data was available for December) and took the numeric values 2 to 11, respectively. Responses measured included counts for basophils, eosinophils, lymphocytes, monocytes, neutrophils, platelets, erythrocytes and total white blood cells. HCT (haematocrit), HGB (haemoglobin concentration), MCH (mean corpuscular haemoglobin) and MCV (mean corpuscular volume) were also analysed.

C-reactive protein. The level of CRP in the peripheral circulation was measured in 3,412 donors (two samples per donor) collected as part of the ASCOT study. Treatment with Atorvastatin did not remove the seasonal variation in this parameter. Age and sex were included as covariates, while a random intercept was added for the individual identifiers.

Statistical analysis of the data sets. Cosinor models with a period of 1 year were fitted to test the effect of season on gene expression. The general formula of the fitted model is given by:

\[ Y_{ij} = a + b \cos(2\pi t_{ij}) + c \sin(2\pi t_{ij}) + d (\text{fixed covariates}) + (\text{random intercepts}) + \epsilon_{ij} \]

where \( Y_{ij} \) represents the log2 expression of gene \( j \) for individual \( i \) recorded at time \( t_{ij} \) with \( \epsilon_{ij} \) computed as the calendar day of the bleed divided by the number of days within the equivalent year.

The fixed covariates and random intercepts terms were data-set specific. For the analysis of the BABYDIET and TID data sets we added age at bleed and gender as fixed effects covariates, whereas only gender was added as a covariate in the analysis of the asthma PBMC microarray dataset (age was not available). The identity of each subject of the BABYDIET and of the asthma data sets were modelled as a random intercept in the corresponding models. For the adipose tissue data set we modelled age at bleed as a fixed covariate and added family identity and an indicator whether the twin was monozygotic or dizygotic as random intercepts. Gender and age at bleed were treated as fixed effects covariates in the analysis of the soluble IL-6 receptor data, and family identity was included as a random effect.

Climatic data for modelling seasonal gene expression. Historical raw data for the mean daily temperature, as well as the total daily hours of sunlight in Munich (Germany), were obtained from the Integrated Climate Data Centre at the University of Hamburg (http://idc.zmaw.de/dwd_station.html?&l=1). For the 1957 TID PBMC data donors from all around United Kingdom we downloaded the maximum and minimum temperature data from seven stations across United Kingdom (Armagh, Camborne, Eddalemuir, Lerwick, Stornaway airport and Valley) from the National Climatic Data Centre, USA (http://www.ncdc.noaa.gov/cdo-web/search) and averaged readings across all stations.

For the analysis of the asthma cohort (ArrayExpress: E-GEOD-19301), the daily maximum and minimum temperature for relevant cities/regions in the United Kingdom (Central England UK station at Birmingham), United States (New Jersey, Seattle, Atlanta, New Haven), Iceland (Reykjavik), Ireland (Dublin) and Australia (Melbourne, Perth, Adelaide) were obtained from the National Climatic Data Centre, USA and The Digital Technology Group. The average temperature values were computed and used in subsequent analyses.
a random intercept. As only the month of bleed was available in the Cambridge Bioresource FBC data, we adjusted the cosinor model to depend on month instead of day; no other covariates were available and random intercepts were not required, as no individual was observed more than once. For the last two data sets, the response variable \( Y \) corresponds to IL-6R and to the tested FBC responses listed in the description of the data set. For analysis of CRP, age, sex and an age*sex interaction were included as fixed covariates, CRP was log transformed to remove right skew, and a random intercept was used to adjust for within individual repeated measures.

To examine whether the effect of season was significant we compared the fitted model in equation (1) with a model that did not include the effect of season. This alternative model is expressed by

\[
Y_{ik} = a + d (\text{fixed covariates}) + g (\text{random intercepts}) + \epsilon_{ik}
\]

(2)

The \( P \)-value for season was determined by comparing the two models for each gene using an analysis of variance test. Seasonal genes were classified as those with \( P \) values less than the data-set-specific Bonferroni correction threshold alpha = 0.05. For the BABYDIET and T1D data sets, we defined as seasonal the genes with \( P \) values less than the corresponding Bonferroni correction \( P \) value and with mean log expression greater than or equal to 0.6.

The relative estimated log2 expression of each seasonal gene for each data set was computed as

\[
Y_{ik} = b \cos(2\pi t_{ik}) + \sin(2\pi t_{ik})
\]

(3)

where \( b \) and \( c \) are the least squares estimates of \( b \) and \( c \) of the model in equation (1), respectively.

Furthermore, we tested whether temperature or sunlight hours could predict gene expression on the PBMC data sets. Temperature and sunlight were defined, respectively, as the average temperature and number of sunlight hours over the week preceding the date of bleed for each individual. For example the temperature model is given by

\[
y_{ik} = a + b \text{ temperature}_{ik} + d (\text{fixed covariates}) + g (\text{random intercepts}) + \epsilon_{ik}
\]

(4)

The three alternative models for the seasonal cosinor function, sunlight and temperature, each including only one of these predictors were fitted to log2 expression level for seasonal genes, as identified in each data set.

**Definition of winter and summer seasonal genes in BABYDIET.** Seasonal genes were classified as winter genes if the relative estimated log2 expression values of the genes were positive for all days of January, February and December and negative for all days of June, July and August. In contrast, summer seasonal genes were defined as those with positive relative estimated log2 expression for all days of June, July and August and negative for all days of January, February and December. The fold change for each summer and winter gene was computed as two computed to the power of the absolute difference of the estimated log2 expression between 15 January and 15 July (days 15 and 196 of a 365-day calendar year).

**Network and functional analysis of the seasonal genes identified in BABYDIET.** A weighted co-expression gene network of the seasonal genes identified in BABYDIET was constructed using the R package WGCNA\(^{17,18}\). For the construction of the network, individuals who sero-converted to T1D autoantibodies at any stage during the BABYDIET study were not included. A scale-free topology network was created based on the seasonal genes, where the correlation of their log2 expression of the genes within the network, individuals who sero-converted to T1D autoantibodies at any stage during the BABYDIET study were not included. A scale-free topology network was created based on the seasonal genes, where the correlation of their log2 gene expression of the PBMC data sets. Temperature and sunlight were defined, each including only one of these predictors were fitted to log2 expression level for seasonal genes, as identified in each data set.

**Analysis of the self-reported infections data of BABYDIET cohort.** The BABYDIET sample sets were divided into two categories, one that included all samples with no self-reported infections (57 samples) and one with all the samples with at least one reported infection (152 samples). A principal component analysis (PCA) was performed, and the first principal component from the analysis was used to summarize the gene expression of BABYDIET seasonal genes. Similarly, the gene expression of the genes within the black module (detected using network analysis of the seasonal genes identified in BABYDIET) was also summarized as the first component of a second PCA. The effect of infection on either of the two components was tested using analysis of variance. The black module was chosen as it contained genes associated with the response to Staphylococcus infection.

**Identification of common seasonal genes.** We wanted to explore whether any of the seasonal genes identified in the PBMC cohorts were shared between the five data sets (excluding Iceland). We compared the Bayesian information criterion (BIC) of the cosinor model (1) with the BIC of the model excluding the seasonality effect (2) for each of the genes from the two in-house data sets (BABYDIET and T1D) that had a \( P \) value <0.05/33297 in at least one of the two data sets. The common seasonal genes of the two in-house datasets were defined as genes whose BIC was smaller for (1) than (2) within each data set. We repeated the aforementioned steps to identify common seasonal genes in the asthma cohort. The intersection of the two lists from the five data sets were defined as common seasonal genes.

We further computed a combined \( P \) value for the association of each common seasonal gene by combining the \( P \) values of the five data sets using Fisher’s product \( P \) value method.

Common seasonal genes between the adipose tissue data set and the BABYDIET data set were defined as the genes that were found seasonal for both data sets.

**Seasonal analysis of The Gambian full blood count data.** Given the different seasonal climates present in West Africa compared to Europe, FBC parameters from The Gambia cohort were assessed through linear models that included sex, age (modelled through splines) and with seasonality modelled using three Fourier terms using STATA12.1. The significance of season was assessed using an F-test.

**Note added in proof.** Adaptive oscillations at balanced polymorphisms in Drosophila in response to acute and persistent changes in climate were reported while this work was under consideration (Bergland, A.O., Behrman, E.L., O’Brien, K.R., Schmidt P.S. & Petrov D.A. Proc. Nat. Genet. 10(111)e1004775 (2014)). Furthermore, seasonally-variable associations of three genes involved in glucose metabolism and circadian clock regulation, CRY1 (cryptochrome 1), CRY2 (cryptochrome 1) and MTNR1B (melatonin receptor 1B) have recently been reported in humans. (Renstrom, F., Koivula, R.W., Varga, T.Y., Hallmans, G., Malldin, H., Florez, J.C., Ha, F.B. & Franks, P.W. Diabetologia 10.1007/s00125-015-3533-8 (2015)).

**References**

1. Kohler, M., Marin-Moratalla, N., Jordana, X. & Aanes, R. Seasonal bone growth and physiology in endotherms shed light on dinosaur physiology. Nature 487, 358–361 (2012).

2. Yanovsky, M. J. & Kay, S. A. Molecular basis of seasonal time measurement in Arabidopsis. Nature 419, 308–312 (2002).

3. Ebling, F. J. On the value of seasonal mammals for identifying mechanisms underlying the control of food intake and body weight. Horm. Behav. 66, 56–65 (2014).

4. Affi, A., Mohamed, E. R. & Ridi, R. E. Seasonal conditions determine the occurrence of posttraumatic stress syndrome and preliminary findings with light therapy. J. Exp. Zool. 265, 439–468 (1993).

5. Mendozza, S. P., Lowe, E. L., Reenko, J. A. & Levine, S. Seasonal variations in gonadal hormones and social behavior in squirrel monkeys. Physiol. Behav. 20, 515–522 (1978).

6. Pell, J. P. & Cobbe, S. M. Seasonal variations in coronary heart disease. QJM 92, 689–696 (1999).

7. Pell, J. P., Sirel, J., Marsden, A. K. & Cobbe, S. M. Seasonal variations in out of hospital cardiopulmonary arrest. Heart 82, 680–683 (1999).

8. Ikuni, N. et al. What’s in season for rheumatoid arthritis patients? Seasonal fluctuations in disease activity. Rheumatology 46, 846–848 (2007).

9. Moltchanova, E. V., Schreier, N., Lammi, N. & Karvonen, M. Seasonal variation in hospital cardiopulmonary arrest. Heart 82, 680–683 (1999).

10. Rosenthal, N. E. et al. Seasonal affective disorder. A description of the syndrome and preliminary findings with light therapy. Arch. Gen. Psychiatry 41, 72–80 (1984).

11. Owens, N. & McGorry, P. D. Seasonality of symptom onset in first-episode schizophrenia. Psychol. Med. 33, 163–167 (2003).

12. Watson, C. G., Kucala, T. J., Kok, C. & Jacobs, L. Schizophrenic birth seasonality in relation to the incidence of infectious diseases and temperature extremes. Arch. Gen. Psychiatry 41, 85–90 (1984).

13. Fisman, D. N. Seasonality of infectious diseases. Annu. Rev. Public Health 28, 127–143 (2007).

14. Dowell, S. F. Seasonal variation in host susceptibility and cycles of certain infectious diseases. Emerg. Infect. Dis. 7, 369–374 (2001).

15. Kasahara, A. K., Singh, R. J. & Noymer, A. Vitamin D (25(OH)D) serum seasonality in the United States. PloS ONE 8, e65785 (2013).

16. Bunger, M. K. et al. Mop3 is an essential component of the master circadian pacemaker in mammals. Cell 103, 1009–1017 (2000).

17. Kondo, T., Kondratova, A. V., Gorbacheva, V. Y., Vykhnovets, O. V. & Antoch, M. P. Early aging and age-related pathologies in mice deficient in BMAL1, the core component of the circadian clock. Genes Dev. 20, 1868–1873 (2006).
18. Harmer, S. L. et al. Orchestrated transcription of key pathways in Arabidopsis by the circadian clock. Science 290, 2110–2113 (2000).

19. Meijer, J. H., van der Veen, E. A., Wignall, H. T. & Rijkers, G. J. Daily and seasonal adaptation of the circadian clock requires plasticity of the SCN neuronal network. Eur. J. Neurosci. 32, 2143–2151 (2010).

20. Dibner, C., Schibler, U. & Albrecht, U. The mammalian circadian timing system: organization and coordination of central and peripheral clocks. Annu. Rev. Physiol. 72, 517–549 (2010).

21. Azzi, A. et al. Circadian behavior is light-reprogrammed by plastic DNA methylation. Nat. Neurosci. 17, 377–382 (2014).

22. Dyar, K. A. et al. Muscle insulin sensitivity and glucose metabolism are controlled by the intrinsic muscle clock. Mol. Metabol. 3, 29–41 (2014).

23. Zhang, R., Lahens, N. F., Ballance, H. L., Hughes, M. E. & Hogness, J. B. A circadian gene expression atlas in mammals: Implications for biology and medicine. Proc. Natl Acad. Sci. USA 111, 16219–16224 (2014).

24. Shackelford, P. G. & Feigin, R. D. Periodicity of susceptibility to pneumococcal infection: influence of light and adenocortical secretions. Science 182, 285–287 (1973).

25. Curtis, A. M., Bellet, M. M., Sassone-Corsi, P. & O’Neill, L. A. Circadian clock alpha regulates expression of the breast cancer 1 associated ring domain 1 protein and heart. Nat. Rev. Endocrinol. 7, 705–714 (2011).

26. Silver, A. C., Arjona, A., Walker, W. E. & Fikrig, E. The circadian clock controls toll-like receptor 9-mediated innate and adaptive immunity. Immunity 36, 251–261 (2012).

27. Storch, K. F. et al. Extensive and divergent circadian gene expression in liver and heart. Nature 417, 78–83 (2002).

28. Nguyen, K. D. et al. Circadian gene Bmal1 regulates diurnal oscillations of inflammatory monocytes. Science 341, 1483–1488 (2013).

29. Beyerlein, A., Wehweck, F., Ziegler, A. G. & Pflueger, M. Respiratory infections for DR3/4 heterozygous and DR4/4 homozygous offspring of parents with autoimmunity in children genetically at-risk of type 1 diabetes. Cell 161, 387–403 (2015).

30. Ussal-Kaczmar, K., Mullen, T. E., Kaufmann, W. K. & Sancar, A. Coupling of human circadian and cell cycles by the timeless protein. Mol. Cell Biol. 25, 3109–3116 (2005).

31. Yu, X. et al. TH17 cell differentiation is regulated by the circadian clock. Science 342, 727–730 (2013).

32. Barnes, P. J. Anti-inflammatory actions of glucocorticoids: molecular mechanisms. Clin. Sci. 94, 557–592 (1998).

33. Lamia, K. A. et al. Cryptochromes mediate rhythmic repression of the glucocorticoid receptor. Nature 480, 552–556 (2011).

34. Bjornsdottir, U. S. et al. Circadian clock Bmal1 regulates diurnal oscillations of a type I interferon transcriptional signature precedes disease. Cell 170, 4176–4178 (2017).

35. Nakaya, H. I. et al. Systems biology approach predicts immunogenicity of the yellow fever vaccine in humans. Nat. Immunol. 10, 116–125 (2009).

36. Li, S. et al. Molecular signatures of antibody responses derived from a systems biology study of five human vaccines. Nat. Immunol. 15, 195–204 (2014).

37. Roederer, M. et al. The genetic architecture of the human immune system: a bioresource for autoimmunity and disease pathogenesis. Cell 161, 387–403 (2015).

38. Orru, V. et al. Genetic variants regulating immune cell levels in health and disease. Cell 155, 242–256 (2013).

39. Wium-Andersen, M., Ørsted, D., Nielsen, S. & Nordestgaard, B. Elevated c-reactive protein level, pathological distress, and depression in 73 131 individuals. JAMA Psychiatry 70, 176–184 (2013).

40. Khandaier, M. G., Pearson, R. M., Zimmitt, S., Lewis, G. & Jones, P. B. Association of serum interleukin 6 and c-reactive protein in childhood with depression and psychosis in young adult life: A population-based longitudinal study. JAMA Psychiatry 71, 1121–1128 (2014).

41. Mondes, M., Klauschen, F., Kuchroo, S. & Germain, R. N. A systems analysis identifies a feedforward inflammatory circuit leading to lethal influenza infection. Cell 154, 197–212 (2013).

42. Yamaguchi, Y. et al. Gastro enterovirus V1a and V1b receptors are resistant to jet lag. Science 342, 85–90 (2013).

43. Bach, J. F. The effect of infections on susceptibility to autoimmune and allergic diseases. N. Engl. J. Med. 347, 911–920 (2002).

44. Archer, S. N. et al. Limed sleep disrupts circadian regulation of the human transcriptome. Proc. Natl Acad. Sci. USA 111, E682–E691 (2014).

45. Tamarkin, L., Baird, C. J. & Almeida, O. F. Melatonin: a coordinating signal for mammalian reproduction? Science 227, 714–720 (1985).

46. Lu, W., Meng, Q. J., Tyler, N. J., Stokkan, K. A. & Loudon, A. S. A circadian clock is not required in an arctic mammal. Curr. Biol. 20, 533–537 (2010).

47. Saenz de Miera, C. et al. A circannual clock drives expression of genes central for seasonal reproduction. Curr. Biol. 24, 1500–1506 (2014).

48. Tanikawa, C. et al. Genome wide association study of age at menarche in the Japanese population. PLoS ONE 8, e63821 (2013).

49. Perry, J. R. et al. Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. Nature 514, 92–97 (2014).

50. Stettler, C. et al. Serum amyloid A, C-reactive protein, and retinal microvascular changes in hypertensive diabetic and nondiabetic individuals: an Anglo-Scandinavian Cardiac Outcomes Trial (ASCOT) substudy. Diabetes Care 32, 1098–1100 (2009).

51. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9, 559 (2008).

52. Wang, J., Duncan, D., Shi, Z. & Zhang, B. WEB-based GEne SeT Analysis Toolkit (WebGestalt): update 2013. Nucleic Acids Res. 41, W77–W83 (2013).
The Gambia for their participation, as well as our laboratory technicians, field assistants, nurses, data entry clerks and other staff at MRC Keneba, The Gambia, involved in the Keneba Biobank. Many thanks also to Kerry Jones, Sophie E Moore, Bai Lamin Dondeh, and Sophie Hawkesworth for their input in setting up the Keneba Biobank. The Gambian study providing data for analysis was supported by core funding MC-A760-5QX00 to the International Nutrition Group by the UK Medical Research Council (MRC) and the UK Department for the International Development (DFID) under the MRC/DFID Concordat agreement. This work was supported by the JDRF UK Centre for Diabetes-Genes, Autoimmunity and Prevention (D-GAP; 4-2007-1003), the JDRF (9-2011-253), the Wellcome Trust (WT061858/091157), the National Institute for Health Research Cambridge Biomedical Research Centre (CBRC) and the Medical Research Council (MRC) Cusrow Wadia Fund. The research leading to these results has received funding from the European Union’s 7th Framework Programme (FP7/2007–2013) under grant agreement no.241447 (NAIMIT). The Cambridge Institute for Medical Research (CIMR) is in receipt of a Wellcome Trust Strategic Award (WT100140). X.C.D. was a University of Cambridge/Wellcome Trust Infection and Immunity PhD student. R.C.F. is funded by a JDRF post-doctoral fellowship (3-2011-374). C.W. and H.G are funded by the Wellcome Trust (WT089989). The BABYDIET study was supported by grants from the Deutsche Forschungsgemeinschaft (DFG ZI-310/14-1 to-4), the JDRF (JDRF 17-2012-16 and 1-2006-665) and the German Center for Diabetes Research (DZD e.V.). E.B. is supported by the DFG Research Center and Cluster of Excellence—Center for Regenerative Therapies Dresden (FZ 111). We also thank Joanna Rorbach, Tomas Castro Dopico, Nikolas Pontikos, Charles Bell and Frank Waldron-Lynch for helpful discussions.

Author contributions
X.C.D. had the idea for the study. X.C.D., M.E., C.W. and J.A.T. designed experiments and analysed the data. M.E., H.G., N.C. and O.S.B. performed statistical analyses. C.W. supervised the statistical analyses. X.C.D. wrote the manuscript with contributions from M.E., E.B., A.-G.Z., C.W. and J.A.T.; R.C.F., M.L.P. and D.I.S. performed wet lab or in silico experiments and assisted with data processing. A.J.F., B.I.H. and A.M.P. were involved in the generation and analysis of blood count data from The Gambia. A.-G.Z. is the PI of the BABYDIET study and takes responsibility of the integrity of the data; A.-G.Z. and E.B. designed the BABYDIET protocol, provided samples and clinical data.

Additional information
Accession codes: Microarray data have been deposited in ArrayExpress under accession code E-MTAB-1724.

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Dopico, X. C. et al. Widespread seasonal gene expression reveals annual differences in human immunity and physiology. Nat. Commun. 6:7000 doi: 10.1038/ncomms8000 (2015).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/