


RESEARCH ARTICLE

Open Access



The combined impact of persistent infections and human genetic variation on C-reactive protein levels

Flavia Hodel^{1,2}, Olivier Naret^{1,2}, Clara Bonnet^{1,2}, Nicole Brenner³, Noemi Bender³, Tim Waterboer³, Pedro Marques-Vidal⁴, Peter Vollenweider⁴ and Jacques Fellay^{1,2,5*} 

Abstract

Multiple human pathogens establish chronic, sometimes life-long infections. Even if they are often latent, these infections can trigger some degree of local or systemic immune response, resulting in chronic low-grade inflammation. There remains an incomplete understanding of the potential contribution of both persistent infections and human genetic variation on chronic low-grade inflammation. We searched for potential associations between seropositivity for 13 persistent pathogens and the plasma levels of the inflammatory biomarker C-reactive protein (CRP), using data collected in the context of the UK Biobank and the CoLausPsyCoLaus Study, two large population-based cohorts. We performed backward stepwise regression starting with the following potential predictors: serostatus for each pathogen, polygenic risk score for CRP, and demographic and clinical factors known to be associated with CRP. We found evidence for an association between *Chlamydia trachomatis* (P -value = $5.04e-3$) and *Helicobacter pylori* (P -value = $8.63e-4$) seropositivity and higher plasma levels of CRP. We also found an association between pathogen burden and CRP levels (P -value = $4.12e-4$). These results improve our understanding of the relationship between persistent infections and chronic inflammation, an important determinant of long-term morbidity in humans.

Keywords: Human genomics, Persistent infections, Inflammation, C-reactive protein

Introduction

Inflammation is a complex and necessary response of the immune system to harmful stimuli such as tissue injury, infection, or exposure to toxins [1]. During the acute phase that is characterized by blood flow changes and increased blood vessel permeability, plasma proteins and leukocytes migrate from the circulation to the site of inflammation [2]. This immediate protective response usually enables the elimination of the initial cause of the cell injury and the restoration of homeostasis. However, when the acute response fails to clear tissue damage,

for example, because of prolonged exposure to stimuli, inflammation can become a chronic process [3]. A number of common diseases are at least partly caused by chronic inflammation, including coronary artery disease, type 2 diabetes, and some cancers [4]. Thus, although inflammation plays an important role in human defense against aggression, it also contributes to the pathophysiology of multiple diseases of major public health importance.

Diagnostic tests are capable of detecting the presence and intensity of systemic inflammation [5]. The most commonly used inflammatory biomarker is the acute-phase reactant C-reactive protein (CRP). This ring-shaped protein is produced by hepatocytes upon stimulation by pro-inflammatory cytokines such as interleukin (IL)-1b, IL-6, and TNF- α . Although CRP is

*Correspondence: jacques.fellay@epfl.ch

¹ Global Health Institute, School of Life Sciences, EPFL, Lausanne, Switzerland
Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

commonly used as a sensitive indicator of inflammation, the factors influencing its baseline plasma levels are only partially understood. Circulating amounts of CRP are positively associated with age, body mass index (BMI), and smoking and inversely with male sex and physical activity [6–8]. In addition, large-scale genomic analyses have found multiple associations with hs-CRP levels, mainly in the loci enriched in hepatic, immune, and metabolic pathways, such as CRP, LEPR, IL6R, GCKR, APOE, and HNF1A-AS1 [9–14]. Altogether, genetic variation explains up to 16% of the variance in plasma CRP levels [14].

To get a more comprehensive view of the factors influencing chronic inflammation in the general population, we used samples and data from the UK Biobank and the CoLaus|PsyCoLaus study to search for associations between baseline CRP levels and chronic infection by persistent/latent pathogens, after careful adjustment for all known demographic, clinical, and genomic influences. Indeed, some infectious agents causing long-term infections in humans have been shown to trigger some degree of local or systemic immune response, resulting in a chronic state of low-grade inflammation that may lead to deleterious health outcomes [15, 16].

Methods

Study cohorts

The UK Biobank is a population-based exploratory study of which the enrollment procedure has been outlined previously [17]. In brief, half a million men and women between the ages of 40 and 69 (45.6% male, mean age \pm SD: 56.5 ± 8.1) visited one of 22 UK Biobank screening centers in England, Scotland, and Wales between 2006 and 2010. The evaluation included a survey, a personal interview, and a number of physical measurements and blood. Urine and saliva samples were also collected for long-term storage. This research was undertaken with approved access to UK Biobank data under application number 50085 (PI: Fellay). All UK Biobank study participants gave informed consent at the time of recruitment. Ethical approval for the UK Biobank study was obtained from the North West Centre for Research Ethics Committee (11/NW/0382).

The CoLaus|PsyCoLaus study is a prospective population-based study initiated in 2003 in Lausanne, Switzerland (www.colaus-psycolaus.ch) [18]. It involves more than 6000 participants of European ancestry (47.5% male) initially aged 35 to 75 years (mean \pm SD: 51.1 ± 10.9), thus representing a sample of approximately 10% of the inhabitants of Lausanne. Individuals were randomly recruited from the general population and are monitored every 5 years regarding their lifestyle and health status. Detailed phenotypic information was obtained from each study

participant through questionnaires, physical assessment, and biological measurements of blood and urine markers. The institutional Ethics Committee of the University of Lausanne, which afterward became the Ethics Commission of Canton Vaud (www.cer-va.ch), approved the baseline CoLaus|PsyCoLaus study (reference 16/03, decisions of 13 January and 10 February 2003), and written consent was obtained from all participants.

DNA genotyping and quality checks

Genotyping and imputation of UK Biobank individuals have been fully described by Bycroft et al. [19]. Briefly, samples were genotyped on either the UK BiLEVE Axiom array (Affymetrix) or UK Biobank Axiom array (Applied Biosystems). Genotypes were phased using SHAPEIT3 and the 1000 Genome phase 3 dataset as a reference, then imputed using IMPUTE4 using the Haplotype Reference Consortium data, 1000 Genomes phase 3, and UK10K data as references [20–22]. Post-imputation quality checks resulted in a total number of 9,349,624 single nucleotide polymorphisms (SNPs) available for analyses. DNA samples from 5399 CoLaus|PsyCoLaus participants were genotyped for 799,653 SNPs using the BB2 GSK-customized Affymetrix Axiom Biobank array. Quality control procedures and imputation of genotypes have been previously described in Hodel et al. [23]. A total of 9,031,263 SNPs from the CoLaus|PsyCoLaus dataset were included for further analyses (flowchart of the inclusion/exclusion criteria are in Additional file 1: Fig. S1).

Measurement of inflammatory biomarkers

For the UK Biobank, non-fasting venous blood samples (~ 50 mL) were collected at recruitment. Blood samples were shipped at 4 °C to the central processing and archiving facility in Stockport. Serum high-sensitivity CRP (hs-CRP) concentrations were measured in participants by immunoturbidimetric assay on a Beckman Coulter AU5800. The manufacturer's analytical range was 0.08 to 80 mg/L. Ninety-five individuals with a hs-CRP level of 20 mg/L were removed from the analysis. For CoLaus|PsyCoLaus, venous blood samples (≥ 50 mL) were drawn in the fasting state and allowed to clot. Serum blood samples were kept at 80 °C before the assessment of cytokines and sent in dry ice to the laboratory. hs-CRP was assessed by immunoassay and latex HS (IMMULITE 1000–High, Diagnostic Products Corporation, LA, CA, USA). For quality control, repeated measurements were conducted on 80 subjects randomly drawn from the initial sample. Forty-seven individuals with hs-CRP levels above 20 mg/L were assigned a value

of 20 by the manufacturer and were therefore removed from the hs-CRP analyses as they are indicative of acute inflammation.

Serological analyses

To assess the humoral responses to a total of 56 antigens derived from 24 persistent infectious agents (45 antigens from 20 pathogens in UK Biobank, and 38 antigens from 18 pathogens in CoLaus|PsyCoLaus), serum samples were independently analyzed by the Infections and Cancer Epidemiology Division at the German Cancer Research Center (Deutsches Krebsforschungszentrum, DKFZ) in Heidelberg [24, 25]. Seroreactivity was measured at serum dilution 1:1000 using multiplex serology based on glutathione-S-transferase (GST) fusion capture immunosorbent assays combined with fluorescent bead technology. For each infectious agent tested, antibody responses were measured for one to six antigens and then expressed as a binary result (IgG positive or negative) based on predefined median fluorescence intensity (MFI) thresholds [26]. For our analysis, only antigens shared between the two cohorts were retained, resulting in a final combination of 27 antigens from 13 pathogens. To define the overall seropositivity against infectious agents when more than one antigen was used, we applied the pathogen-specific algorithms suggested by the manufacturer. Details of the methods on how the antigens were combined have been described previously [26].

Combining study cohorts

Upon completion of the genotyping and quality control (QC) analyses for each cohort, imputed datasets were matched on the strand, SNP ID, and genomic coordinates. Additional analyses and QC checkpoints were performed to ensure proper merging. This resulted in a dataset of 12,055 unique individuals of European ancestry and a total of 6,899,629 markers.

Polygenic risk score calculation for hs-CRP level

We carried out a polygenic risk score (PRS) analysis to investigate the relationship between human genetic variation and hs-CRP levels. A CRP-PRS was calculated for each study participant based on the risk effects of common SNPs derived from GWAS summary statistics of hs-CRP. As a baseline cohort, we referred to the GWAS summary statistics of the CHARGE cohort ($N=204,402$, heritability $h^2=6.5\%$) [10, 27]. These summary statistics were used to construct the CRP-PRS in our target cohort consisting of the merged UK Biobank and CoLaus|PsyCoLaus data using the clumping and thresholding method of the PRSice-2 v2.2.7 software [28]. We used a standardized method to obtain PRS, by

multiplying the dosage of risk alleles for each variant by the effect size in the GWAS and summing the scores across all of the selected variants. SNPs were clumped based on linkage disequilibrium (LD) ($r^2 \geq 0.1$) within a 250-kb window. Model estimates of the PRS effect were adjusted for sex, age, BMI, and the top 10 PCs. As an additional quality control, the distribution of PRS was checked in each cohort separately, to ensure that they followed a normal distribution.

Analyses of the determinants of hs-CRP levels

We used linear regression with backward selection to identify the factors significantly associated with hs-CRP plasma levels. Tested covariates included serostatus for each pathogen, polygenic risk score for CRP, age, sex, BMI, and the first 10 PCs of the genotyping data. P -value < 0.05 was considered statistically significant. The analysis was performed using the stepAIC function in R version 4.0.5 [29].

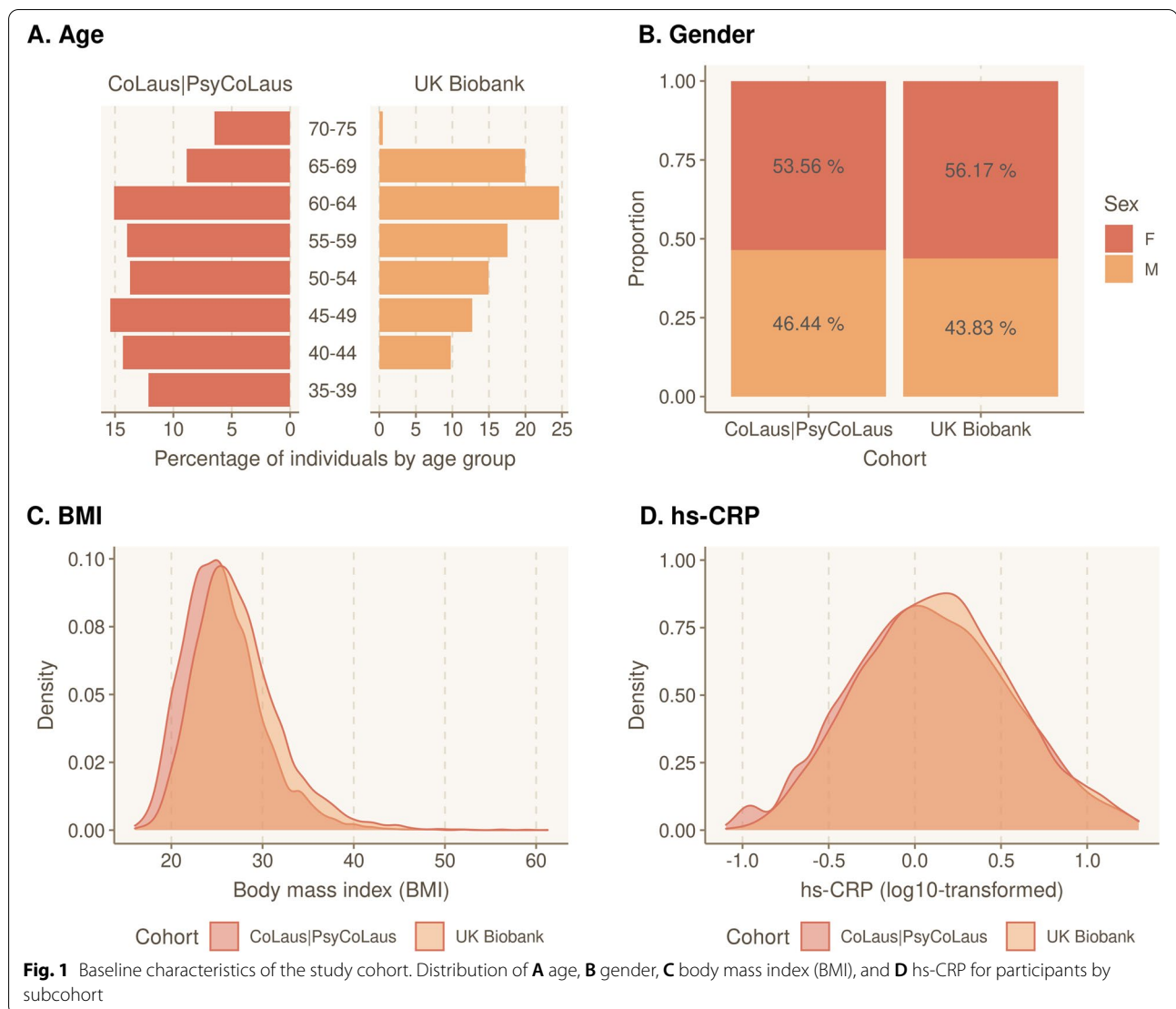
Results

Baseline characteristics of study participants

We studied a total of 12,055 individuals with available hs-CRP level, serological results, and genome-wide genotyping data from two independent population-based studies: the UK Biobank ($N=8371$) and the CoLaus|PsyCoLaus study ($N=3684$) (Additional file 1: Fig. S1). Participants ranged in age from 35 to 75 years (mean age \pm SD: 55.68 ± 9.07), with a majority of women (55.4%) and a mean BMI of 26.80 (\pm SD: 4.73). The hs-CRP level was measured in all participants. The median hs-CRP level was 1.30 mg/L (10th, 90th percentiles: 0.35 mg/L, 5.10 mg/L, respectively). Figure 1 shows the distributions of age, sex, BMI, and log₁₀-transformed hs-CRP in both cohorts. We observed a very comparable distribution of all relevant variables in the two cohorts, which were merged for downstream analyses. Additional file 2: Fig. S2 shows the associations of hs-CRP with demographic and clinical factors. Higher levels of hs-CRP associated with female sex, age, and increased BMI (P -values = $1.5e-3$, $3.4e-69$, and ≈ 0 , respectively).

The impact of genetic variation on hs-CRP levels

The filtered genetic variants from the two cohorts were combined (see the “Methods” section) to increase the sample size. To estimate the sample variation, and to control for potential population structure and genotyping bias, principal component analysis (PCA) was performed using the correlation matrix of the genotyping data. PCA plots for the first ten principal components (PC1–PC10) are shown in Additional file 3: Fig. S3A, annotated by the original cohort from which the sample was drawn. We



observed that samples from both subgroups (UK Biobank and CoLaus|PsyCoLaus) were segregated on the first PC (PC1) and eighth PC (PC8), but not on the other PCs. The top 10 PCs explained 61% of the total variance and were used throughout the study to correct for stratification (Additional file 3: Fig. S3B).

We computed a CRP-PRS to investigate the effect of multiple gene variants on hs-CRP levels. A total of 1809 SNPs were included at the best P -value threshold (P -value = $3.65e-3$). The PRS followed a normal distribution in the merged cohort, as well as in each subcohort separately (Additional file 4: Fig. S4). To describe the influence of common human genetic variation on plasma hs-CRP levels, we quantified the trait variance (R^2) explained by the derived PRS and covariates across individuals. We observed that the variance explained by the full model was 25.8%, with 21.5% attributed to

the demographic and clinical covariates and 4.3% to the CRP-PRS. The association between the CRP-PRS and hs-CRP levels was very strong (P -value = $6.58e-123$; Additional file 5: Fig. S5), with hs-CRP levels increasing by 0.48 [standard error (SE) 0.02] for each standard deviation increment in CRP-PRS.

Associations between persistent/latent infections and hs-CRP levels

We searched for associations between hs-CRP levels and serostatus for the following persistent or frequently recurring human pathogens: 10 viruses (BK virus (BKV), *Cytomegalovirus* (CMV), Epstein-Barr virus (EBV), Human Herpes Virus (HHV)-6, HHV-7, Herpes Simplex Virus (HSV)-1, HSV-2, JC virus (JCV), Kaposi's sarcoma-associated herpesvirus (KSHV), and Varicella

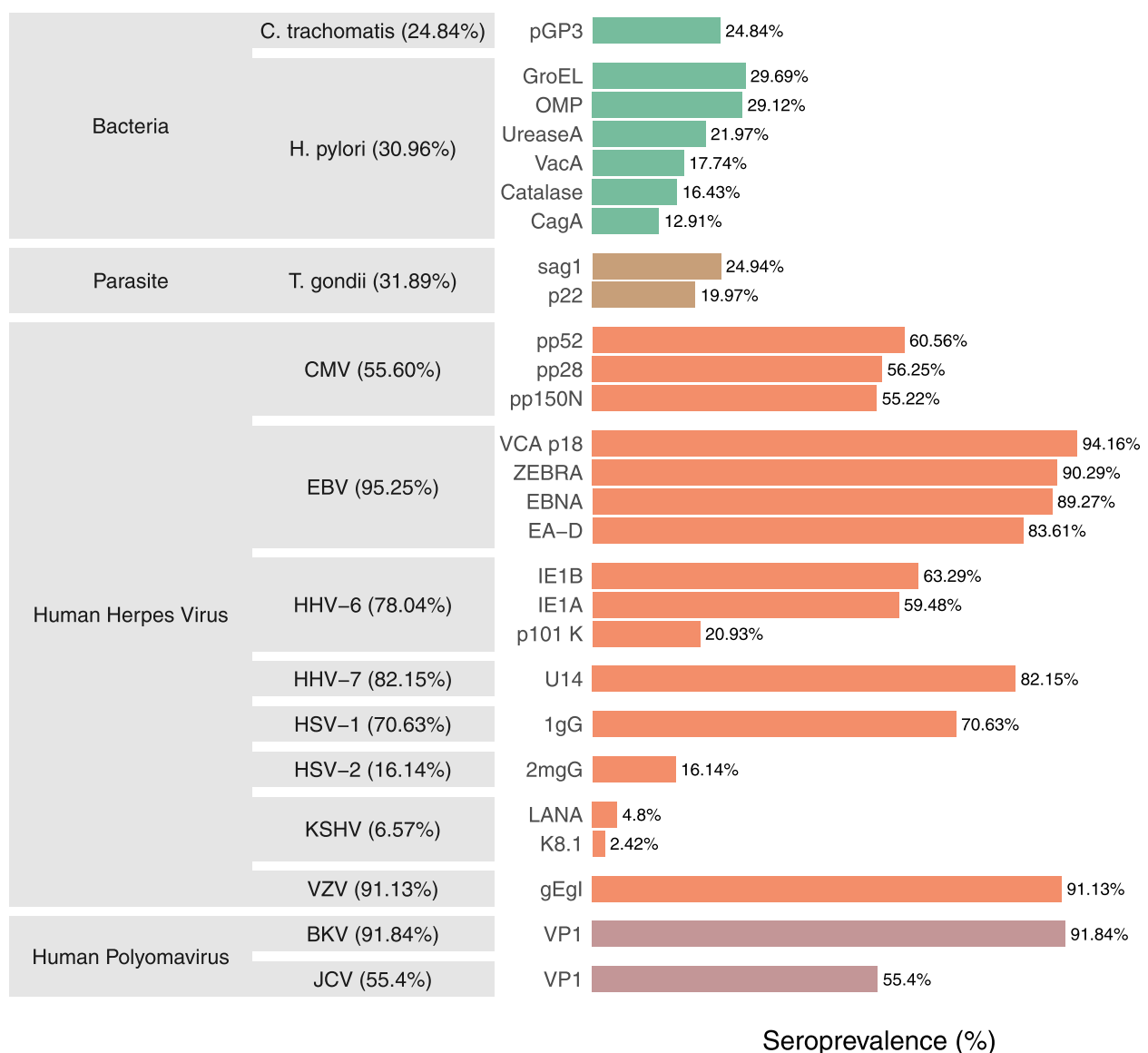


Fig. 2 Overall pathogen seropositivity and seroprevalence of tested antigens. List of the 13 pathogens and 27 antigens available from the combined study. The gray boxes indicate the pathogen on which the antigen protein is found, and the family to which the pathogen belongs. Percentages in parentheses after pathogen names indicate the overall seropositivity for the specified pathogen. The percentages on the right indicate the seroprevalence of antibodies against infectious disease antigens tested using the Multiplex Serology platform. For study-based figures, see Supplementary Figs. 6 and 7

zoster virus (VZV)); two bacteria (*Chlamydia trachomatis* (*C. trachomatis*) and *Helicobacter pylori* (*H. pylori*)); and one parasite (*Toxoplasma gondii* (*T. gondii*)) (Fig. 2). The overall seropositivity ranged from 6.57% (KSHV) to 95.25% (EBV). Cohort-separated seroprevalences are shown in Additional files 6 and 7: Figs. S6 and S7.

Using backward stepwise regression including all significantly identified persistent or frequently recurring human pathogens, adjusted for CRP-PRS, sex,

age, BMI, and the top 10 PCs, we observed significant associations of hs-CRP levels with seropositivity for *H. pylori* (P -value = 8.63×10^{-4}) and *C. trachomatis* (P -value = 5.04×10^{-3}) (Table 1). The final regression model including all significant factors explained 25.9% of the variance of hs-CRP levels. This explained that the fraction of the variance is similar to the value obtained without including the serological results (above), indicating that the impact of *H. pylori* and *C. trachomatis*

Table 1 Linear regression analysis results for hs-CRP

Characteristic	Beta	95% CI	P-value
<i>C. trachomatis</i>			0.005
Seronegative	–	–	
Seropositive	0.02	0.01, 0.04	
<i>H. pylori</i>			< 0.001
Seronegative	–	–	
Seropositive	0.03	0.01, 0.04	
Age	0.01	0.01, 0.01	< 0.001
BMI	0.04	0.04, 0.04	< 0.001
Sex			< 0.001
M	–	–	
F	0.07	0.06, 0.09	
CRP-PRS	0.10	0.09, 0.10	< 0.001

CI Confidence interval

of pathogens for which they were seropositive (range: 0–13). The number of individuals in each serological stratum ranged from 5 (index=0) to 2717 (index=7) and is presented in Fig. 3. We used a linear model to search for an association between pathogen burden and hs-CRP levels. hs-CRP levels were found to be significantly and positively associated with increasing pathogen burden (P -value = 4.12×10^{-4}) (Fig. 3).

Discussion

Mounting evidence suggests that exposure to multiple pathogens, even when they do not cause obvious disease, can affect the immune system and human health [18, 30, 31]. In an effort to better understand the variability of humoral immune response and inflammation patterns in response to pathogen exposure, we selected 27 antigens from 13 persistent infectious agents, which we evaluated

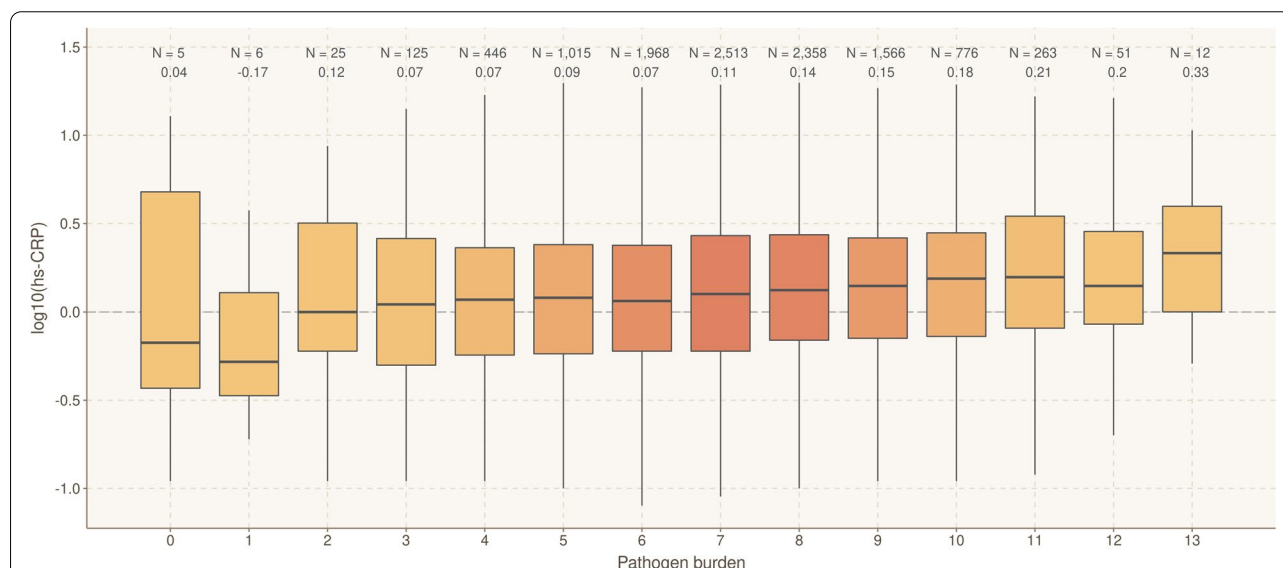


Fig. 3 Levels of hs-CRP by infectious burden. Boxplots showing the hs-CRP value for each pathogen load group. The black bold line within the boxplot indicates the median of the hs-CRP measurement. The boxes are colored by sample size. The sample size and median for each group are shown above the box

seropositivity on chronic inflammation, even if statistically significant, is likely to be minimal at the population level. We also investigated the interaction effect between the two identified pathogens on the hs-CRP level. No significant interaction was observed, suggesting a joint independent impact of *H. pylori* and *C. trachomatis*.

Pathogen burden associates with higher hs-CRP levels

We then checked if the overall burden of chronic infections contributes to increased hs-CRP levels. Study participants were stratified according to their overall seropositivity index, calculated by summing the number

using multiplex serology to detect specific immunoglobulin G levels in two well-characterized population-based cohorts.

We first investigated the relationship between common genetic variation and hs-CRP levels by calculating a PRS for all study participants. The PRS explained about 4% of the variation in hs-CRP levels, in agreement with previously published results [9]. We also found that BMI was the major non-genetic predictor of hs-CRP, with approximately 19% of the variance explained.

Next, we studied the impact of persistent infections on chronic inflammation after adjustment for known

influencing factors, including age, sex, BMI, and human genetic variability, as explored above. We observed an association between increased levels of hs-CRP and seropositivity for *C. trachomatis* and *H. pylori*. The two gram-negative bacteria *C. trachomatis* and *H. pylori* do not cause life-long, latent infections. Nevertheless, they are responsible for some of the most frequent chronic infections in humans.

H. pylori can colonize the gastric epithelium for long periods of time, leading to chronic inflammation of the gastric mucosa. Even if the majority of individuals infected with *H. pylori* have no symptoms, the bacterium has been causally linked with gastritis, gastric ulcer, and an increased risk of gastric cancer [32, 33]. Our results suggest a systemic impact of chronic *H. pylori* infection beyond the known local inflammatory effect on the gastric mucosa, confirming an observation made previously in a cross-sectional population-based study [34].

C. trachomatis causes genital and ocular infections. The ocular manifestation of the infection, trachoma, is the world's leading cause of preventable blindness and is endemic in many developing countries. This clinical presentation is however highly unlikely to contribute to the 25% seroprevalence of anti-chlamydia antibodies detected in the Swiss and UK cohorts included in our study. More relevant here, *C. trachomatis* is the etiological agent of human chlamydia urogenital tract infection, which is the most common bacterial sexually transmitted disease. Chronic or recurrent forms of the disease are frequently observed. To our knowledge, no study has examined the direct association between *C. trachomatis* infection and hs-CRP levels at the population level. However, studies conducted in the context of associations between *C. trachomatis* and tubal factor-related subfertility and preterm delivery have also shown elevated hs-CRP levels [30, 31, 35]. Altogether, these results confirm the role of chronic or recurrent bacterial infections in low-grade inflammation, reflected by a small but consistent increase in hs-CRP levels in seropositive individuals. In addition, we found an association between increased pathogen burden and hs-CRP levels by stratifying individuals according to their cumulative number of positive serological results. This indicates that latent infections might play an enhancing role in chronic low-grade inflammation, even if that effect is too small to be detected at the individual pathogen level.

Previous studies have shown that pro-inflammatory cytokines and chronic inflammation are associated with cellular aging (“inflammaging”) and a number of non-communicable diseases, including certain cancers, type 2 diabetes, and cardiovascular disease [3, 4, 36, 37]. It would therefore not be surprising to find that infections

also play a key role in these diseases and that the reactivation of these pathogens can contribute to the deterioration of the overall health of older individuals. Finally, CRP-PRS was also found to be significantly associated in the analysis including both genetics and serological results, confirming that human genetic variation plays a modulating role in systemic inflammation.

Our study has some limitations. Firstly, we cannot rule out the effects of other non-measured infections at the time of hs-CRP measurement that may have influenced the level of inflammatory biomarkers. Also, we did not adjust our models for all known influencing factors (e.g., smoking, anti-inflammatory or anti-infective drugs, or possible inflammatory diseases). However, participants in both studies were assumed to be in good overall health at the time of data collection, and the data were filtered before analysis to detect the levels indicative of acute infection. Secondly, some pathogens had relatively low or high seroprevalences and should be reexamined in a larger study. In particular, it will be interesting to repeat the analysis once serological data for all individuals in the UK Biobank are available. This will allow for greater reliability in terms of statistical power. Third, hs-CRP was the only inflammatory biomarker studied. Other pro-inflammatory cytokines such as IL-1 β , IL-6, and TNF- α are regulators of host responses to infection and positive mediators of inflammation. Consideration of these other biomarkers would give insight into more specific inflammatory pathways and provide a more comprehensive picture of the overall inflammatory status. Fourth, we only observed associations with the presence of chronic inflammation, and our study design does not allow us to infer any kind of causality. In particular, we cannot exclude the possibility that higher levels of inflammation are responsible for the reactivation of a pathogen, resulting in detectable seropositivity. [38, 39]

In conclusion, we found that seropositivity for *C. trachomatis* and *H. pylori* antigens is associated with increased levels of hs-CRP. Together with demographic, clinical, and genetic factors, persistent infections contribute to chronic low-grade inflammation, which can have deleterious long-term consequences on health.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12916-022-02607-7>.

Additional file 1: Fig. S1. Flowchart illustrating the inclusion/exclusion of individuals in the study. Orange boxes indicate the number of included antigens and pathogens.

Additional file 2: Fig. S2. Scatterplot and regression line (with 95% confidence intervals) to describe the relationship of hs-CRP with characteristics of study participants. Relationship between hs-CRP and A) age, B) sex, C) BMI and D) polygenic risk score (PRS). For linear regressions, linear regression equation, R-squared and P-value are shown.

Additional file 3: Fig. S3. Principal component analysis (PCA) of combined genotyping data. A) PCA plot of the first ten PCs of the genotyping data. Samples are colored by cohort. B) Histogram explaining the variance of each PC component. In the histogram, the variance explained by each eigenvalue is labeled on top.

Additional file 4: Fig. S4. Distribution of polygenic risk score (PRS) values. Density distribution of standardized PRS values by subcohort (CoLaus|PsyCoLaus and UKB) and across all participants (combined).

Additional file 5: Fig. S5. Polygenic risk score for hs-CRP (CRP-PRS) was significantly associated with hs-CRP levels. Scatter plots with linear regression line of polygenic risk scores predicting hs-CRP levels for individuals in the cohort. 95% confidence interval is showed in grey shade.

Additional file 6: Fig. S6. Seroprevalence of tested antigens in the CoLaus|PsyCoLaus. List of the 27 antigens available from the CoLaus|PsyCoLaus study that are shared with the UK Biobank. The percentages indicate the seroprevalence of antibodies against infectious disease antigens tested using Multiplex Serology platform. The grey boxes indicate the pathogen on which the antigen protein is found, and the family to which the pathogen belongs.

Additional file 7: Fig. S7. Seroprevalence of tested antigens in the UK Biobank. List of the 27 antigens available from the UK Biobank that are shared with the CoLaus|PsyCoLaus study. The percentages indicate the seroprevalence of antibodies against infectious disease antigens tested using Multiplex Serology platform. The grey boxes indicate the pathogen on which the antigen protein is found, and the family to which the pathogen belongs.

Acknowledgements

We thank the participants in the UK Biobank and CoLaus|PsyCoLaus study for their time and contribution to this study. We also thank all the clinical, academic, and administrative collaborators who helped with the participant recruitment, study coordination, data collection, and storage.

Authors' contributions

FH: conceptualization, methodology, software, formal analysis, visualization, and writing—original draft. ON: formal analysis. CB: formal analysis. NBR: resources. NBe: resources. TW: investigation and resources. PMV: resources, data curation, and investigation. PV: investigation. JF: Funding acquisition, project administration, supervision, and writing—original draft. All co-authors reviewed the manuscript. The authors read and approved the final manuscript.

Funding

This project was supported by the Swiss National Science Foundation (grant 31003A_175603 to JF). The CoLaus|PsyCoLaus study was and is supported by research grants from GlaxoSmithKline, the Faculty of Biology and Medicine of Lausanne, and the Swiss National Science Foundation (grants 3200B0_105993, 3200B0_118308, 33CSO_122661, 33CSO_139468, 33CSO_148401, and 33CSO_177535/1).

Availability of data and materials

The data of the CoLaus|PsyCoLaus study used in this article cannot be fully shared as they contain potentially sensitive personal information on participants. According to the Ethics Committee for Research of the Canton of Vaud, sharing these data would be a violation of Swiss legislation with respect to privacy protection. However, coded individual-level data that do not allow researchers to identify participants are available upon request to researchers who meet the criteria for data sharing of the CoLaus|PsyCoLaus Datacenter (CHUV, Lausanne, Switzerland). Any researcher affiliated to a public or private research institution who complies with the CoLaus|PsyCoLaus standards can submit a research application to research.colaus@chuv.ch or research.psycolaus@chuv.ch. Proposals requiring baseline data only will be evaluated by the baseline (local) Scientific Committee (SC) of the CoLaus and PsyCoLaus studies. Proposals requiring follow-up data will be evaluated by the follow-up (multicentric) SC of the CoLaus|PsyCoLaus cohort study. Detailed instructions for gaining access to the CoLaus|PsyCoLaus data used in this study are available at www.colaus-psycolaus.ch/professionals/how-to-collaborate. GWAS summary statistics results and the list of SNPs included in the CRP-PRS

calculation are available for download from Zenodo under: <https://zenodo.org/record/6517122>.

Declarations

Ethics approval and consent to participate

Ethical approval for the UK Biobank study was obtained from the North West Centre for Research Ethics Committee (11/NW/0382). The institutional Ethics Committee of the University of Lausanne, which afterward became the Ethics Commission of Canton Vaud (www.cer-va.ch) approved the baseline CoLaus|PsyCoLaus study (reference 16/03, decisions of 13 January and 10 February 2003).

Consent for publication

I confirm that written consent was obtained from all participants and the appropriate institutional forms have been archived.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Global Health Institute, School of Life Sciences, EPFL, Lausanne, Switzerland. ²Swiss Institute of Bioinformatics, Lausanne, Switzerland. ³Division of Infections and Cancer Epidemiology, German Cancer Research Center, Heidelberg, Germany. ⁴Department of Medicine, Internal Medicine, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland. ⁵Precision Medicine Unit, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland.

Received: 18 January 2022 Accepted: 13 October 2022
Published online: 01 November 2022

References

- Medzhitov R. The spectrum of inflammatory responses. *Science*. 2021;374(6571):1070–5.
- Ryan GB, Majno G. Acute inflammation: a review. *Am J Pathol*. 1977;86:183–276.
- Franceschi C, Campisi J. Chronic inflammation (inflammaging) and its potential contribution to age-associated diseases. *J Gerontol A Biomed Sci Med Sci*. 2014;69(Suppl_1):S4–9.
- Furman D, Campisi J, Verdin E, Carrera-Bastos P, Targ S, Franceschi C, Ferrucci L, Gilroy DW, Fasano A, Miller GW, et al. Chronic inflammation in the etiology of disease across the life span. *Nat Med*. 2019;25(12):1822–32.
- Feghali CA, Wright TM, et al. Cytokines in acute and chronic inflammation. *Front Biosci*. 1997;2(1):d12–26.
- De Martinis M, Franceschi C, Monti D, Ginaldi L. Inflammation markers predicting frailty and mortality in the elderly. *Exp Mol Pathol*. 2006;80(3):219–27.
- Ferrucci L, Corsi A, Lauretani F, Bandinelli S, Bartali B, Taub DD, Guralnik JM, Longo DL. The origins of age-related proinflammatory state. *Blood*. 2005;105(6):2294–9.
- Marques-Vidal P, Bochud M, Bastardot F, Lüscher T, Ferrero F, Gaspoz J-M, Paccaud F, Urwyler A, von Känel R, Hock C, et al. Levels and determinants of inflammatory biomarkers in a Swiss population-based sample (CoLaus study). *PLoS ONE*. 2011;6(6):e21002.
- Dehghan A, Dupuis J, Barbalic M, Bis JC, Eiriksdottir G, Lu C, Pellicka N, Wallaschowski H, Kettunen J, Henneman P, et al. Meta-analysis of genome-wide association studies in > 80 000 subjects identifies multiple loci for C-reactive protein levels. *Circulation*. 2011;123(7):731–8.
- Ligthart S, Vaez A, Vösa U, Stathopoulou MG, De Vries PS, Prins BP, Van der Most PJ, Tanaka T, Naderi E, Rose LM, et al. Genome analyses of > 200,000 individuals identify 58 loci for chronic inflammation and highlight pathways that link inflammation and complex disorders. *Am J Human Genet*. 2018;103(5):691–706.
- Prins BP, Kuchenbaecker KB, Bao Y, Smart M, Zabaneh D, Fatemifar G, Luan J, Wareham NJ, Scott RA, Perry JRB, et al. Genome-wide analysis of health-related biomarkers in the UK household longitudinal study reveals novel associations. *Sci Rep*. 2017;7(1):1–9.

12. Reiner AP, Barber MJ, Guan Y, Ridker PM, Lange LA, Chasman DI, Walston JD, Cooper GM, Jenny NS, Rieder MJ, et al. Polymorphisms of the HNF1A gene encoding hepatocyte nuclear factor-1 α are associated with C-reactive protein. *Am J Human Genet.* 2008;82(5):1193–201.
13. Ridker PM, Pare G, Parker A, Zee RYL, Danik JS, Buring JE, Kwiatkowski D, Cook NR, Miletich JP, Chasman DI. Loci related to metabolic-syndrome pathways including LEPR, HNF1A, IL6R, and GCKR associate with plasma C-reactive protein: the women's genome health study. *Am J Human Genet.* 2008;82(5):1185–92.
14. Said S, Pazoki R, Karhunen V, Vösa U, Ligthart S, Bodinier B, Koskeridis F, Welsh P, Alizadeh BZ, Chasman DI, et al. Genetic analysis of over half a million people characterises C-reactive protein loci. *Nat Commun.* 2022;13(1):1–10.
15. Thom DH, Grayston JT, Siscovick DS, Wang S-P, Weiss NS, Daling JR. Association of prior infection with Chlamydia pneumoniae and angiographically demonstrated coronary artery disease. *Jama.* 1992;268(1):68–72.
16. Zhu J, Quyyumi AA, Norman JE, Csako G, Epstein SE. Cytomegalovirus in the pathogenesis of atherosclerosis: the role of inflammation as reflected by elevated C-reactive protein levels. *J Am Coll Cardiol.* 1999;34(6):1738–43.
17. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* 2015;12(3):e1001779.
18. Firmann M, Mayor V, Marques Vidal P, Bochud M, Pécoud A, Hayoz D, Paccaud F, Preisig M, Song KS, Yuan X, et al. The CoLaus study: a population-based study to investigate the epidemiology and genetic determinants of cardiovascular risk factors and metabolic syndrome. *BMC Cardiovasc Disord.* 2008;8(1):1–11.
19. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J, et al. The UK biobank resource with deep phenotyping and genomic data. *Nature.* 2018;562(7726):203–9.
20. Delaneau O, Marchini J, Zagury J-F. A linear complexity phasing method for thousands of genomes. *Nat Methods.* 2012;9(2):179–81.
21. Loh P-R, Danecek P, Francesco Palamara P, Fuchsberger C, Reshef YA, Finucane HK, Schoenherr S, Forer L, McCarthy S, Abecasis GR, et al. Reference-based phasing using the haplotype reference consortium panel. *Nat Genet.* 2016;48(11):1443–8.
22. McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, Teumer A, Kang HM, Fuchsberger C, Danecek P, Sharp K, et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet.* 2016;48(10):1279.
23. Hodel F, Chong AY, Scepanovic P, Xu ZM, Naret O, Thorball CW, Rüeger S, Marques-Vidal P, Vollenweider P, Begemann M, et al. Human genomics of the humoral immune response against polyomaviruses. *Virus Evol.* 2021;7(2):veab058.
24. Waterboer T, Sehr P, Michael KM, Franceschi S, Nieland JD, Joos TO, Templin MF, Pawlita M. Multiplex human papillomavirus serology based on in situ-purified glutathione S-transferase fusion proteins. *Clin Chem.* 2005;51(10):1845–53.
25. Waterboer T, Sehr P, Pawlita M. Suppression of non-specific binding in serological Luminex assays. *J Immunol Methods.* 2006;309(1–2):200–4.
26. Mentzer AJ, Brenner N, Allen N, Littlejohns TJ, Chong AY, Cortes A, et al. Identification of host–pathogen–disease relationships using a scalable multiplex serology platform in UK Biobank. *Nat Commun.* 2022;13(1):1–12.
27. Psaty BM, O'donnell CJ, Gudnason V, Lunetta KL, Folsom AR, Rotter JJ, Uitterlinden AG, Harris TB, Witteman JCM, Boerwinkle E. Cohorts for heart and aging research in genomic epidemiology (CHARGE) consortium: design of prospective meta-analyses of genome-wide association studies from 5 cohorts. *Circ Cardiovasc Genet.* 2009;2(1):73–80.
28. Choi SW, Mak TS-H, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. *Nat Protoc.* 2020;15(9):2759–72.
29. R Core Team, et al. R: a language and environment for statistical computing. USA: SA; 2013.
30. Dubey P, Pandey K, Singh N, Bhagoliwal A, Sharma D. Role of maternal serum chlamydia trachomatis IgG antibodies and serum C-reactive protein in preterm labor. *Int J Reprod Contracept Obstet Gynecol.* 2014;3(1):195–8.
31. Den Hartog JE, Land JA, Stassen FRM, Kessels AGH, Bruggeman CA. Serological markers of persistent C. trachomatis infections in women with tubal factor subfertility. *Human Reprod.* 2005;20(4):986–90.
32. Sipponen P, Hyvärinen H. Role of helicobacter pylori in the pathogenesis of gastritis, peptic ulcer and gastric cancer. *Scand J Gastroenterol.* 1993;28(sup196):3–6.
33. Taylor-Robinson D, Thomas BJ. The role of chlamydia trachomatis in genital-tract and associated diseases. *J Clin Pathol.* 1980;33(3):205.
34. Jackson L, Britton J, Lewis SA, McKeever TM, Atherton J, Fullerton D, Fogarty AW. A population-based epidemiologic study of helicobacter pylori infection and its association with systemic inflammation. *Helicobacter.* 2009;14(5):460–5.
35. Karinen L, Pouta A, Bloigu A, Koskela P, Paldanius M, Leinonen M, Saikku P, Järvelin M-R, Hartikainen A-L. Serum C-reactive protein and Chlamydia trachomatis antibodies in preterm delivery. *Obstet Gynecol.* 2005;106(1):73–80.
36. Gavazzi G, Krause K-H. Ageing and infection. *Lancet Infect Dis.* 2002;2(11):659–66.
37. Khansari N, Shakiba Y, Mahmoudi M. Chronic inflammation and oxidative stress as a major cause of age-related diseases and cancer. *Recent Pat Inflammation Allergy Drug Discov.* 2009;3(1):73–80.
38. Hammer C, Begemann M, McLaren PJ, Bartha I, Michel A, Klose B, Schmitt C, Waterboer T, Pawlita M, Schulz TF, et al. Amino acid variation in HLA class II proteins is a major determinant of humoral response to common viruses. *Am J Human Genet.* 2015;97(5):738–43.
39. Scepanovic P, Alanio C, Hammer C, Hodel F, Bergstedt J, Patin E, Thorball CW, Chaturvedi N, Charbit B, Abel L, et al. Human genetic variants and age are the strongest predictors of humoral immune responses to common pathogens and vaccines. *Genome Med.* 2018;10(1):1–13.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

