



## Plasma proteins associated with circulating carotenoids in Nepalese school-aged children<sup>☆</sup>

Abdulkerim Eroglu<sup>a,b,\*</sup>, Kerry J. Schulze<sup>a,b</sup>, James Yager<sup>a,c</sup>, Robert N. Cole<sup>d,e</sup>, Parul Christian<sup>a,b</sup>, Bareng A.S. Nonyane<sup>a,b</sup>, Sun Eun Lee<sup>a,b</sup>, Lee S.F. Wu<sup>a,b</sup>, Subarna Khatry<sup>a,b</sup>, John Groopman<sup>a,c</sup>, Keith P. West Jr.<sup>a,b</sup>

<sup>a</sup> Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

<sup>b</sup> Center for Human Nutrition, Department of International Health, Baltimore, MD, USA

<sup>c</sup> Department of Environmental Health and Engineering, Baltimore, MD, USA

<sup>d</sup> Johns Hopkins School of Medicine, Mass Spectrometry and Proteomics Facility, Baltimore, MD, USA

<sup>e</sup> Department of Biological Chemistry, Baltimore, MD, USA

### ARTICLE INFO

#### Keywords:

Proteomics  
Plasma proteins  
β-cryptoxanthin  
Antioxidants  
Carotenoids  
Nepal

### ABSTRACT

Carotenoids are naturally occurring pigments that function as vitamin A precursors, antioxidants, anti-inflammatory agents or biomarkers of recent vegetable and fruit intake, and are thus important for population health and nutritional assessment. An assay approach that measures proteins could be more technologically feasible than chromatography, thus enabling more frequent carotenoid status assessment. We explored associations between proteomic biomarkers and concentrations of 6 common dietary carotenoids (α-carotene, β-carotene, lutein/zeaxanthin, β-cryptoxanthin, and lycopene) in plasma from 500 6–8 year old Nepalese children. Samples were depleted of 6 high-abundance proteins. Plasma proteins were quantified using tandem mass spectrometry and expressed as relative abundance. Linear mixed effects models were used to determine the carotenoid:protein associations, accepting a false discovery rate of  $q < 0.10$ . We quantified 982 plasma proteins in > 10% of all child samples. Among these, relative abundance of 4 were associated with β-carotene, 11 with lutein/zeaxanthin and 51 with β-cryptoxanthin. Carotenoid-associated proteins are notably involved in lipid and vitamin A transport, antioxidant function and anti-inflammatory processes. No protein biomarkers met criteria for association with α-carotene or lycopene. Plasma proteomics may offer an approach to assess functional biomarkers of carotenoid status, intake and biological function for public health application.

Original maternal micronutrient trial from which data were derived as a follow-up activity was registered at ClinicalTrials.gov: NCT00115271.

### 1. Introduction

Carotenoids are pigments occurring naturally in fruits and vegetables [1]. They are compounds synthesized from eight isoprenoid units, and more than 700 are found in nature [2–6]. Structurally, hydrocarbon carotenoids are referred to as carotenes and oxygenated carotenoids are termed xanthophylls. Among them, α-carotene, β-carotene, β-cryptoxanthin, lycopene, lutein, and zeaxanthin are the major dietary carotenoids found in human plasma [7]. While α-carotene, β-carotene and β-cryptoxanthin are provitamin A carotenoids, meaning

they can be metabolized to retinol, lycopene, lutein and zeaxanthin cannot be converted to vitamin A [8]. Provitamin A carotenoids may be particularly important dietary sources for maintaining vitamin A status in impoverished regions, such as in rural Southern Asia [9], where vitamin A deficiency (VAD) has been shown to affect 20–35% of young children, school-aged adolescents and women of reproductive age [10–12] and is widely associated with low intakes of carotenoid-rich foods such as dark green leaves, yellow and orange vegetables and fruits and egg [12]. Carotenoids also appear to have *in vivo* antioxidant [13] and immunoregulatory [14] properties that are thought to give

<sup>☆</sup> This study was supported by the Assessment of Micronutrient Status by Nutriproteomics Grant OPP 5241 from the Bill & Melinda Gates Foundation (Yiwu He, Senior Program Officer). The cohort study in Nepal from which plasma samples were obtained was supported by the Global Control of Micronutrient Deficiency Grant GH 614 from the Bill & Melinda Gates Foundation (Ellen Piwoz, Senior Program Officer). The original field trial in Nepal from 1999 to 2001 in which mothers of studied children were enrolled was supported by Micronutrients for Health Cooperative Agreement HRN-A-00-97-00015-00 between the Office of Health, Infectious Diseases and Nutrition, US Agency for International Development and the Center for Human Nutrition, Johns Hopkins Bloomberg School of Public Health. The Sight and Life Global Nutrition Research Institute also provided additional technical assistance.

\* Corresponding author. Johns Hopkins Bloomberg School of Public Health, 615 N. Wolfe St., Room W2607, Baltimore, MD, 21205, USA.

E-mail address: [aeroglul@jhu.edu](mailto:aeroglul@jhu.edu) (A. Eroglu).

rise to frequent associations between their dietary intake or circulating concentrations and reduced risks of cardiovascular disease [15], various cancers [16] and macular degeneration [17,18]. Thus, plasma carotenoid concentrations may comprise a class of *in vivo* biomarkers that both reflect a diverse and nutritious diet [19–21] and nutritional, antioxidant, anti-inflammatory health of populations.

However, as molecules largely detected by chromatographic methods [22], carotenoids represent a group of micronutrients that are rarely assessed in low-middle income countries, signaling a need to explore novel approaches for their assessment in populations. In exploring plasma proteomics as an approach to ascertain potential biomarkers of micronutrient, functional, and health status in an undernourished population of school-aged children in Nepal, we have revealed associations between clusters of circulating proteins and micronutrient status (vitamins A, E, D and K, copper and selenium) [23–26], inflammation [27], cognition [28] and anthropometry [29] in an undernourished population of school-aged children in Nepal. Findings to date suggest that plasma proteomics can identify proteins predictive of nutritional and health status that are candidate biomarkers with the potential to be measured by multi-analyte approaches for protein quantification. Missing from the emerging knowledge base is evidence of proteins that reflect plasma carotenoids, which could benefit from assays more readily conducted than conventional biochemical methods. The objective of this study was to explore the direction, strength and plausibility of association between plasma proteins and plasma carotenoid concentrations in a rural population of Nepalese school-aged children.

## 2. Methods

**Study cohort and field data collection.** We obtained plasma samples from 3305 children 6–8 years of age living in the southern plains district of Sarlahi, Nepal, born to mothers who had previously participated in a 5-arm antenatal micronutrient supplementation trial [30]. Following stratification by original maternal supplementation allocation group, 1000 samples were randomly selected (200 per original trial group) from children with multiple aliquots of plasma samples, complete data records from both the original trial and current follow-up study, and valid birth size measures for multiple biochemical assessments [31]. Of these, 500 samples were randomly chosen for proteomics analysis, maintaining original trial balance. Children from whom samples were selected have been described in detail previously and are typical of children in the region [23–29,31]. Follow-up data were collected on household socioeconomic characteristics, dietary frequencies and morbidity history for the previous 7 days and anthropometry (weight, height, mid-upper arm circumference), as reported earlier [24]. Weight-for-age Z-score (WAZ), height-for-age Zscore (HAZ) and body mass index (BMI)-for-age Z-score (BMIZ) were used to characterize nutritional status [32]. Venous blood was drawn from children following an overnight fast, which was processed into plasma aliquots and stored in liquid nitrogen in a field laboratory. Frozen plasma was transported in vapor phase liquid nitrogen shippers to the micronutrient analysis laboratory at Johns Hopkins University in Baltimore, Maryland, U.S.A. and stored at  $-80^{\circ}\text{C}$  until analysis. In both the original field trial and follow-up study, informed consents were obtained and protocols were approved by the Nepal Health Research Council in Kathmandu, Nepal, and the Institutional Review Board at Johns Hopkins Bloomberg School of Public Health in Baltimore, Maryland, USA.

**Plasma carotenoid analyses.** Plasma carotenoids including  $\beta$ -carotene, lutein and zeaxanthin,  $\beta$ -cryptoxanthin,  $\alpha$ -carotene, and lycopene were analyzed by HPLC (Waters 2795) with a quaternary gradient pump, autosampler, photodiode array detector and Empower 2 software following the procedure of Yamini et al. [33]. The peaks of lutein and zeaxanthin could not be distinguished as they are combined. Separation of carotenoids was achieved using an All sphere ODS-2, 5- $\mu\text{m}$ , 4.6-mm column (Alltech) and a Supelguard Discovery C18 2-cm  $\times$  4.0-

**Table 1**

Plasma carotenoid concentrations ( $\mu\text{mol/L}$ ) of children, 6–8 years of age, Sarlahi, Nepal.

Plasma carotenoids	n <sup>a</sup>	Median (IQR) <sup>b</sup>
$\beta$ -Carotene <sup>c</sup>	497	0.10 (0.06, 0.19)
Lutein/zeaxanthin	500	0.34 (0.25, 0.48)
$\beta$ -Cryptoxanthin	500	0.06 (0.04, 0.12)
$\alpha$ -Carotene	481	0.01 (0.01, 0.02)
Lycopene	171	0.02 (0.01, 0.03)

<sup>a</sup> n represents the number of samples with detectable concentration of each carotenoid.

<sup>b</sup> IQR = interquartile range.

<sup>c</sup> Cutoff for low plasma  $\beta$ -carotene is  $0.09\ \mu\text{mol/L}$ , below which 207 values were observed (41.6%).

mm guard column (Sigma-Aldrich). The assay was calibrated using the National Institute of Standards and Technology standard reference material SRM968d.

**Plasma proteomics analysis.** Mass spectrometric and proteomics procedures developed for this study have been reported elsewhere [23,34]. In brief, 500 plasma samples (40  $\mu\text{L}$ ) were immunoaffinity-depleted of six high abundance proteins (albumin, IgG, IgA, transferrin, haptoglobin and antitrypsin), which constitute 85–90% of total plasma protein content, using a Human-6 Multiple Affinity Removal System (MARS) LC column (Agilent Technologies). Protein extracts (100  $\mu\text{g}$  each) were TCA/acetone precipitated, trypsin digested, labeled by isobaric mass tags (iTRAQ 8-plex reagents), and then seven samples plus one pooled sample for quality control was fractionated by strong cation exchange (SCX) chromatography and analyzed on a LTQ Orbitrap Velos mass spectrometer (Thermo Scientific). MS/MS spectra were searched against the RefSeq 40 database using Mascot (Matrix Science) through Proteome Discoverer software (version 1.3, Thermo Scientific) to quantify proteins with respect to the within-iTRAQ medians of  $\log_2$  transformed and normalized reporter ion intensities derived from Proteome Discoverer. Data were obtained from 72 iTRAQ experiments with average  $589 \pm 65$  proteins quantified per iTRAQ experiment. A total of 4705 proteins were detected, with 982 quantified in  $> 10\%$  ( $n > 50$ ) of all samples [23] and 146 proteins measured in all 500 samples, representing the plasma proteome for this study.

### 2.1. Statistical analysis

Detailed information on estimation of protein relative abundance from reporter ion intensities within each iTRAQ experiment was published elsewhere [34]. We applied linear mixed effects models (LME) to determine the association between  $\log_2$  transformed plasma concentration of each carotenoid and the relative abundance of individual plasma proteins accounting for multiple iTRAQ experiments.

The expected values of  $\log_2$  carotenoid concentrations for each individual protein from the LME can be expressed as

$$E\{N_{rk}\} = b_0 + B_r + b_1 P_{rk}$$

where  $N_{rk}$  is the  $\log_2$ -transformed plasma concentration of each carotenoid,  $k$  is the index for each sample in each  $r$  iTRAQ experiment, and  $P_{rk}$  is the protein relative abundance estimate. The parameter  $b_0$  is the estimate of the intercept which is the overall mean concentration of each carotenoid;  $B_r$  is the random deviation of experiment  $r$  from this mean; and,  $b_1$  is the estimate of the slope of the nutrient:protein association. Statistical significance of a protein:nutrient association was assessed by a two-sided hypothesis test for  $b_1 = 0$ . For individual significant nutrient:protein correlations, a  $q$ -value, an adjusted  $p$ -value to control false discovery rate (FDR) was reported [35]. Protein:nutrient correlation ( $r$ ) and  $R^2$  were calculated based on the observed plasma carotenoid concentrations and their respective best linear unbiased predictions from the LME models [36].

retinol	.20 (n=481)	.18 (n=497)	.30 (n=500)	.34 (n=500)	.33 (n=171)	1 (n=500)
log <sub>2</sub> lycopene	.14 (n=168)	.00 (n=168)	.34 (n=171)	.52 (n=171)	1 (n=171)	retinol
log <sub>2</sub> lutein/zeaxanthin	.23 (n=481)	.49 (n=497)	.42 (n=500)	1 (n=500)	log <sub>2</sub> lycopene	
log <sub>2</sub> β-cryptoxanthin	.27 (n=481)	.37 (n=497)	1 (n=500)	log <sub>2</sub> lutein/zeaxanthin		
log <sub>2</sub> β-carotene	.03 (n=481)	1 (n=497)	log <sub>2</sub> β-cryptoxanthin			
log <sub>2</sub> α-carotene	1 (n=481)	log <sub>2</sub> β-carotene				

**Fig. 1. Correlations between plasma carotenoids in 6–8 year old children of rural Nepal.** Correlation coefficients were generated using all complete pairwise data. All of the correlations were statistically significant ( $p < 0.0001$ ) except log<sub>2</sub> β-carotene with log<sub>2</sub> α-carotene ( $p = 0.56$ ), log<sub>2</sub> β-carotene with log<sub>2</sub> lycopene ( $p = 0.96$ ) and log<sub>2</sub> α-carotene with log<sub>2</sub> lycopene ( $p = 0.0786$ ).

**Table 2**  
Plasma proteins associated with plasma log<sub>2</sub> β-carotene in 6–8 year old children of rural Nepal (n = 497).<sup>a</sup>

Gene Name	Gene Symbol	n <sup>b</sup>	r	R <sup>2</sup>	p	q	b <sub>1</sub> <sup>c</sup>	gi Accession Number <sup>d</sup>
Orosomucoid 1	ORM1	497	-0.7	0.49	6.07 × 10 <sup>-5</sup>	2.60 × 10 <sup>-2</sup>	-33.0	167857790
Apolipoprotein A-I	APOA1	497	0.7	0.49	6.28 × 10 <sup>-5</sup>	2.60 × 10 <sup>-2</sup>	94.1	4557321
Pyruvate kinase, muscle	PKM	55	-0.65	0.42	1.78 × 10 <sup>-4</sup>	4.94 × 10 <sup>-2</sup>	-67.7	33286422
TNFAIP3 interacting protein 1	TNIP1	385	-0.69	0.48	1.79 × 10 <sup>-4</sup>	4.94 × 10 <sup>-2</sup>	-35.6	116256481

<sup>a</sup> Four proteins quantified by mass spectrometry and estimated by linear mixed effects (LME) modelling in > 10% of the samples (50 < n ≤ 497) that are correlated with plasma log<sub>2</sub> β-carotene, subjected to a false discovery rate (FDR) cutoff of 10% ( $q < 0.10$ ), and listed in increasing order of q, defined as candidate protein biomarkers for a plasma β-carotene proteome.

<sup>b</sup> n represents the number of child plasma samples in which a protein was detected and quantified by iTRAQ MS.

<sup>c</sup> b<sub>1</sub> represents the percent change in plasma β-carotene, (in μmol/L) per 2-fold (100%) increase in protein relative abundance.

<sup>d</sup> GenInfo Identifier sequence number, as assigned to all nucleotide and protein sequences by the National Center for Biotechnology Information at the National Library of Medicine, National Institutes of Health, Bethesda, MD, USA.

**Table 3**  
Plasma proteins associated with plasma log<sub>2</sub> lutein/zeaxanthin in 6–8 year old children of rural Nepal (n = 500).<sup>a</sup>

Gene Name	Gene Symbol	n <sup>b</sup>	r	R <sup>2</sup>	p	q	b <sub>1</sub> <sup>c</sup>	gi Accession Number <sup>d</sup>
Apolipoprotein C-III	APOC3	500	0.59	0.34	9.15 × 10 <sup>-5</sup>	5.12 × 10 <sup>-2</sup>	25.4	4557323
Inter-alpha-trypsin inhibitor heavy chain 3	ITIH3	500	-0.58	0.34	1.21 × 10 <sup>-4</sup>	5.12 × 10 <sup>-2</sup>	-28.2	133925809
Heat shock 70 kDa protein 1A	HSPA1A	119	0.51	0.26	1.61 × 10 <sup>-4</sup>	5.43 × 10 <sup>-2</sup>	69.9	194248072
Collagen, type V, alpha 1	COL5A1	56	0.59	0.35	3.57 × 10 <sup>-4</sup>	8.40 × 10 <sup>-2</sup>	101.3	89276751
Minichromosome maintenance complex component 2	MCM2	119	-0.64	0.41	4.95 × 10 <sup>-4</sup>	8.40 × 10 <sup>-2</sup>	-31.8	33356547
TNFAIP3 interacting protein 1	TNIP1	388	-0.6	0.36	5.43 × 10 <sup>-4</sup>	8.40 × 10 <sup>-2</sup>	-22.4	116256481
CD14 molecule	CD14	500	-0.58	0.34	5.57 × 10 <sup>-4</sup>	8.40 × 10 <sup>-2</sup>	-37.7	4557417
Carnosine dipeptidase 1 (metallopeptidase M20 family)	CNDP1	500	0.58	0.34	6.12 × 10 <sup>-4</sup>	8.40 × 10 <sup>-2</sup>	16.8	21071039
Interferon-related developmental regulator 2	IFRD2	486	0.59	0.35	6.47 × 10 <sup>-4</sup>	8.40 × 10 <sup>-2</sup>	27.8	197333755
Apolipoprotein A-I	APOA1	500	0.58	0.34	6.48 × 10 <sup>-4</sup>	8.40 × 10 <sup>-2</sup>	43.3	4557321
Proteoglycan 4	PRG4	403	0.6	0.36	7.84 × 10 <sup>-4</sup>	9.44 × 10 <sup>-2</sup>	46.7	189181724

<sup>a</sup> Eleven proteins quantified by mass spectrometry and estimated by linear mixed effects (LME) modelling in > 10% of the samples (50 < n ≤ 500) that are correlated with plasma log<sub>2</sub> lutein/zeaxanthin, subjected to a false discovery rate (FDR) cutoff of 10% ( $q < 0.10$ ), and listed in increasing order of q, defined as candidate protein biomarkers for a plasma proteome of lutein/zeaxanthin.

<sup>b</sup> n represents the number of child plasma samples in which a protein was detected and quantified by iTRAQ MS.

<sup>c</sup> b<sub>1</sub> represents the percent change in plasma lutein/zeaxanthin, (in μmol/L) per 2-fold (100%) increase in protein relative abundance.

<sup>d</sup> GenInfo Identifier sequence number, as assigned to all nucleotide and protein sequences by the National Center for Biotechnology Information at the National Library of Medicine, National Institutes of Health, Bethesda, MD, USA.

We present a list of all proteins with an FDR less than 10% ( $q < 0.10$ ) in their associations with each plasma carotenoid, and their corresponding Human Genome Organization (HUGO) gene symbols [37], the number of samples with detected protein values (n), protein:nutrient correlation (r), the amount of variance in nutrient

concentration as explained by the protein (R<sup>2</sup>), p-value derived from testing the fixed effects slope of carotenoid concentration on protein abundance, chance-adjusted p-value (q), the slope (b<sub>1</sub>), denoting relative (%) change in carotenoids per 2-fold (100%) increase in relative abundance of each protein and GenInfo identifier (gi) accession number

**Table 4**  
Plasma proteins positively associated with plasma log<sub>2</sub> β-cryptoxanthin in 6–8 year old children of rural Nepal (n = 500).<sup>a</sup>

Gene Name	Gene Symbol	n <sup>b</sup>	r	R <sup>2</sup>	p	q	b <sub>1</sub> <sup>c</sup>	gi Accession Number <sup>d</sup>
Apolipoprotein A-I	<i>APOA1</i>	500	0.52	0.27	5.27 × 10 <sup>-9</sup>	7.95 × 10 <sup>-6</sup>	193.1	4557321
Interferon-related developmental regulator 2	<i>IFRD2</i>	486	0.51	0.26	1.33 × 10 <sup>-7</sup>	1.00 × 10 <sup>-4</sup>	99.4	197333755
Carnosine dipeptidase 1 (metallopeptidase M20 family)	<i>CNDP1</i>	500	0.50	0.25	5.41 × 10 <sup>-6</sup>	1.63 × 10 <sup>-3</sup>	44.0	21071039
Apolipoprotein C-III	<i>APOC3</i>	500	0.50	0.25	8.47 × 10 <sup>-6</sup>	2.13 × 10 <sup>-3</sup>	58.1	4557323
Phospholipid transfer protein	<i>PLTP</i>	472	0.49	0.24	1.09 × 10 <sup>-5</sup>	2.36 × 10 <sup>-3</sup>	133.7	5453914
PILR alpha associated neural protein	<i>PIANP</i>	360	0.49	0.24	1.85 × 10 <sup>-5</sup>	3.49 × 10 <sup>-3</sup>	191.8	24308547
Anthrax toxin receptor 2	<i>ANTXR2</i>	367	0.53	0.28	2.11 × 10 <sup>-5</sup>	3.53 × 10 <sup>-3</sup>	95.5	50513243
Selenoprotein P, plasma, 1	<i>SEPP1</i>	500	0.50	0.25	2.69 × 10 <sup>-5</sup>	4.06 × 10 <sup>-3</sup>	96.7	62530391
Clusterin	<i>CLU</i>	493	0.50	0.25	5.73 × 10 <sup>-5</sup>	7.20 × 10 <sup>-3</sup>	300.1	42740907
Insulin-like growth factor binding protein, acid labile subunit	<i>IGFALS</i>	500	0.49	0.24	8.82 × 10 <sup>-5</sup>	8.87 × 10 <sup>-3</sup>	71.2	4826772
Protein phosphatase, Mg <sup>2+</sup> /Mn <sup>2+</sup> dependent, 1 M	<i>PPM1M</i>	98	0.57	0.33	1.24 × 10 <sup>-4</sup>	1.17 × 10 <sup>-2</sup>	14.2	171460934
Apolipoprotein A-II	<i>APOA2</i>	500	0.49	0.24	1.52 × 10 <sup>-4</sup>	1.35 × 10 <sup>-2</sup>	90.4	4502149
Paraoxonase 1	<i>PON1</i>	500	0.49	0.24	2.03 × 10 <sup>-4</sup>	1.51 × 10 <sup>-2</sup>	61.5	19923106
Lymphatic vessel endothelial hyaluronan receptor 1	<i>LYVE1</i>	479	0.5	0.25	2.10 × 10 <sup>-4</sup>	1.51 × 10 <sup>-2</sup>	64.4	40549451
Retinol binding protein 4, plasma	<i>RBP4</i>	500	0.49	0.24	2.80 × 10 <sup>-4</sup>	1.69 × 10 <sup>-2</sup>	75.6	55743122
Eukaryotic translation initiation factor 2D	<i>EIF2D</i>	256	0.52	0.27	3.31 × 10 <sup>-4</sup>	1.82 × 10 <sup>-2</sup>	82.1	56699485
Apolipoprotein D	<i>APOD</i>	500	0.49	0.24	3.38 × 10 <sup>-4</sup>	1.82 × 10 <sup>-2</sup>	71.5	4502163
Kelch-like family member 34	<i>KLHL34</i>	403	0.51	0.26	3.90 × 10 <sup>-4</sup>	2.03 × 10 <sup>-2</sup>	90.3	23397572
Kruppel-like factor 17	<i>KLF17</i>	284	0.51	0.26	4.69 × 10 <sup>-4</sup>	2.36 × 10 <sup>-2</sup>	52.5	104294874
Prenylcysteine oxidase 1	<i>PCYOX1</i>	493	0.48	0.23	5.48 × 10 <sup>-4</sup>	2.43 × 10 <sup>-2</sup>	68.4	166795301
Gelsolin	<i>GSN</i>	493	0.49	0.24	5.93 × 10 <sup>-4</sup>	2.47 × 10 <sup>-2</sup>	92.5	4504165
Thrombospondin 4	<i>THBS4</i>	451	0.5	0.25	6.08 × 10 <sup>-4</sup>	2.47 × 10 <sup>-2</sup>	49.5	31543806
Apolipoprotein C-I	<i>APOC1</i>	500	0.49	0.24	6.63 × 10 <sup>-4</sup>	2.50 × 10 <sup>-2</sup>	27.8	4502157
Dipeptidyl-peptidase 4	<i>DPP4</i>	416	0.49	0.24	7.05 × 10 <sup>-4</sup>	2.60 × 10 <sup>-2</sup>	94.0	18765694
Interleukin 1 receptor accessory protein	<i>IL1RAP</i>	321	0.48	0.23	8.68 × 10 <sup>-4</sup>	3.05 × 10 <sup>-2</sup>	93.5	19882209
Peptidase inhibitor 16	<i>PI16</i>	500	0.49	0.24	1.01 × 10 <sup>-3</sup>	3.46 × 10 <sup>-2</sup>	64.3	70780384
Lumican	<i>LUM</i>	500	0.49	0.24	1.01 × 10 <sup>-3</sup>	3.59 × 10 <sup>-2</sup>	86.2	4505047
Insulin-like growth factor binding protein 3	<i>IGFBP3</i>	500	0.49	0.24	1.87 × 10 <sup>-3</sup>	5.46 × 10 <sup>-2</sup>	49.3	62243068
Cartilage oligomeric matrix protein	<i>COMP</i>	500	0.49	0.24	1.88 × 10 <sup>-3</sup>	5.46 × 10 <sup>-2</sup>	52.3	40217843
Glycosylphosphatidylinositol specific phospholipase D1	<i>GPLD1</i>	500	0.49	0.24	2.93 × 10 <sup>-3</sup>	7.76 × 10 <sup>-2</sup>	72.8	29171717
Nucleolar protein 12	<i>NOL12</i>	228	0.52	0.27	3.39 × 10 <sup>-3</sup>	8.53 × 10 <sup>-2</sup>	239.9	13236553

<sup>a</sup> Thirty-one proteins quantified by mass spectrometry and estimated by linear mixed effects (LME) modelling in > 10% of the samples that were positively correlated with plasma log<sub>2</sub> β-cryptoxanthin ( $p < 0.01$ ,  $q < 0.10$ ), listed in increasing order of  $q$ , defined as positively associated protein biomarkers of a plasma β-cryptoxanthin.

<sup>b</sup>  $n$  represents the number of child plasma samples in which a protein was detected and quantified by iTRAQ tandem MS (excludes subsequent imputations required for multivariable LME models).

<sup>c</sup>  $b_1$  represents the percent change in plasma β-cryptoxanthin (in μmol/L) per 2-fold (100%) increase in protein relative abundance.

<sup>d</sup> GenInfo Identifier sequence number, as assigned to all nucleotide and protein sequences by the National Center for Biotechnology Information at the National Library of Medicine, National Institutes of Health, Bethesda, MD, USA.

[37] is provided in the tables.

Datasets of plasma carotenoid concentrations and protein relative abundance presented in this study are available in [Supplementary Table 1](#). All analyses were conducted using open source software built under the R statistical computing environment [38].

### 3. Results

Nutritional status and demographic characteristics of study children (n = 500) are shown in [Supplementary Table 2](#). Study children were undernourished as reflected by low anthropometry Z-scores: 48.5%, 39.1%, and 16.1% of them were considered underweight, stunted, and thin (weight-for-age, height-for-age, and BMI-for-age Z-scores < -2, respectively), relative to the World Health Organization (WHO) reference population [32]. For β-carotene, 41.6% children had plasma concentrations < 0.09 μmol/L ([Table 1](#)), which is considered low [33]. While the plasma xanthophyll carotenoids, β-cryptoxanthin and lutein/zeaxanthin were detected among all children, plasma carotenes, α-carotene and lycopene, were more likely to be below detection limits. Values below the detection limit were not included in the distribution of values shown in [Table 1](#).

A correlation matrix for measured carotenoids is provided for all available pairwise data in [Fig. 1](#), with coefficients ranging from 0.03 ( $p = 0.56$ ) to 0.52 ( $p < 0.0001$ ). Strongest correlations were observed between lutein/zeaxanthin and lycopene ( $r = 0.52$ ,  $p < 0.0001$ ) and β-carotene ( $r = 0.49$ ,  $p < 0.0001$ ). Significant and comparable correlations also existed between β-cryptoxanthin and both lutein/

zeaxanthin ( $r = 0.42$ ,  $p < 0.0001$ ) and β-carotene ( $r = 0.37$ ,  $p < 0.0001$ ). The prevalence of vitamin A deficiency (plasma retinol < 0.7 μmol/L) was 8.8 %, as reported [23].

Of the 982 detected proteins, 4 were associated with plasma β-carotene, 11 with lutein/zeaxanthin, and 51 with β-cryptoxanthin, meeting a FDR threshold of 10% ( $q < 0.10$ ). No proteins met this criterion of association for plasma α-carotene or lycopene.

Among the 4 proteins associated with plasma log<sub>2</sub> β-carotene, only apolipoprotein-A1 (APOA1) was positively associated, while orosomucoid 1 (ORM1), pyruvate kinase muscle (PKM) and TNFAIP3 interacting protein 1 (TNIP1) were negatively associated with this carotenoid ([Table 2](#)).

Seven proteins, apolipoprotein-C3 (APOC3), heat shock 70 kDa protein 1A (HSPA1A), collagen type V alpha 1 (COL5A1), carnosine dipeptidase 1 (CNDP1), interferon-related developmental regulator (IFRD2), APOA1, and proteoglycan 4 (PRG4) were positively associated with plasma log<sub>2</sub> lutein/zeaxanthin. Four, including inter-alpha-trypsin inhibitor heavy chain 3 (ITH3), minichromosome maintenance complex 2 (MCM2), TNIP1, and CD14 molecule (CD14) protein were negatively associated with log<sub>2</sub> lutein/zeaxanthin ([Table 3](#)).

Among the 51 proteins associated with log<sub>2</sub> β-cryptoxanthin, 31 were positive correlates ([Table 4](#)), of which 10 met a more stringent FDR of < 1%, including APOA1, IFRD2, CNDP1, APOC3, phospholipid transfer protein (PLTP), PILR alpha associated neural protein (PIANP), anthrax toxin receptor 2 (ANTXR2), selenoprotein P plasma 1 (SEPP1), clusterin (CLU), and insulin-like growth factor binding protein, acid labile subunit (IGFALS). Other positively correlated plasma proteins



**Table 5**  
Plasma proteins negatively associated with plasma log<sub>2</sub> β-cryptoxanthin in 6–8 year old children of rural Nepal (n = 500).<sup>a</sup>

Gene Name	HUGO Gene Symbol	n <sup>b</sup>	R	R <sup>2</sup>	p	q	b <sub>1</sub> <sup>c</sup>	Accession Number <sup>d</sup>
Orosomucoid 1	ORM1	500	-0.50	0.25	1.60 × 10 <sup>-6</sup>	8.04 × 10 <sup>-4</sup>	-41.5	167857790
TNFAIP3 interacting protein 1	TNIP1	388	-0.50	0.25	2.18 × 10 <sup>-6</sup>	8.22 × 10 <sup>-4</sup>	-46.3	116256481
Mannosidase, alpha, class 1A, member 1	MAN1A1	500	-0.50	0.25	4.17 × 10 <sup>-5</sup>	5.71 × 10 <sup>-3</sup>	-66.0	24497519
Haptoglobin	HP	354	-0.53	0.28	6.94 × 10 <sup>-5</sup>	8.06 × 10 <sup>-3</sup>	-13.7	4826762
Alpha-1-B glycoprotein	A1BG	500	-0.50	0.25	8.04 × 10 <sup>-5</sup>	8.66 × 10 <sup>-3</sup>	-67.1	21071030
Haptoglobin-related protein	HPR	431	-0.52	0.27	1.83 × 10 <sup>-4</sup>	1.51 × 10 <sup>-2</sup>	-22.0	45580723
Leucine-rich alpha-2-glycoprotein 1	LRG1	500	-0.49	0.24	2.76 × 10 <sup>-4</sup>	1.69 × 10 <sup>-2</sup>	-37.0	16418467
Serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 3	SERPINA3	500	-0.49	0.24	5.42 × 10 <sup>-4</sup>	2.43 × 10 <sup>-2</sup>	-49.1	50659080
Orosomucoid 2	ORM2	500	-0.49	0.24	6.37 × 10 <sup>-4</sup>	2.43 × 10 <sup>-2</sup>	-41.3	4505529
Leucine rich repeat containing 47	LRRC47	70	-0.51	0.26	7.35 × 10 <sup>-4</sup>	2.64 × 10 <sup>-2</sup>	-75.0	24308207
Complement factor B	CFB	500	-0.49	0.24	1.08 × 10 <sup>-3</sup>	3.59 × 10 <sup>-2</sup>	-46.1	67782358
Beta-2-microglobulin	B2M	493	-0.49	0.24	1.28 × 10 <sup>-3</sup>	4.10 × 10 <sup>-2</sup>	-32.7	4757826
Ecotropic viral integration site 5	EVI5	271	-0.54	0.29	1.58 × 10 <sup>-3</sup>	4.95 × 10 <sup>-2</sup>	-34.6	68299759
Ubiquitin-conjugating enzyme E2L 3	UBE2L3	144	-0.33	0.11	1.68 × 10 <sup>-3</sup>	5.18 × 10 <sup>-2</sup>	-44.2	4507789
Component of oligomeric golgi complex 3	COG3	215	-0.31	0.09	1.76 × 10 <sup>-3</sup>	5.30 × 10 <sup>-2</sup>	-35.8	13899251
Lipopolysaccharide binding protein	LBP	500	-0.49	0.24	2.09 × 10 <sup>-3</sup>	5.73 × 10 <sup>-2</sup>	-28.4	31652249
Complement component 9	C9	500	-0.49	0.24	2.32 × 10 <sup>-3</sup>	6.26 × 10 <sup>-2</sup>	-38.6	4502511
Inter-alpha-trypsin inhibitor heavy chain family, member 4	ITI4	500	-0.49	0.24	3.12 × 10 <sup>-3</sup>	8.03 × 10 <sup>-2</sup>	-58.1	31542984
Mitogen-activated protein kinase kinase kinase 14	MAP3K14	307	-0.50	0.25	3.14 × 10 <sup>-3</sup>	8.03 × 10 <sup>-2</sup>	-28.4	115298645
Lysozyme	LYZ	493	-0.49	0.24	3.74 × 10 <sup>-3</sup>	9.25 × 10 <sup>-2</sup>	-34.6	4557894

<sup>a</sup> Twenty proteins quantified by mass spectrometry and estimated by linear mixed effects (LME) modelling in > 10% of the samples that were negatively correlated with plasma log<sub>2</sub> β-cryptoxanthin ( $p < 0.01$ ,  $q < 0.10$ ), listed in increasing order of  $q$ , defined as negatively associated protein biomarkers of a plasma β-cryptoxanthin.

<sup>b</sup>  $n$  represents the number of child plasma samples in which a protein was detected and quantified by iTRAQ MS (excludes subsequent imputations required for multivariable LME models).

<sup>c</sup>  $b_1$  represents the percent change in plasma β-cryptoxanthin (in μmol/L) per 2-fold (100%) increase in protein relative abundance.

<sup>d</sup> GenInfo Identifier sequence number, as assigned to all nucleotide and protein sequences by the National Center for Biotechnology Information at the National Library of Medicine, National Institutes of Health, Bethesda, MD, USA.

included retinol binding protein 4 (RBP4), paraoxonase 1 (PON1), prenylcysteine oxidase 1 (PCYOX1), lymphatic vessel endothelial hyaluronan receptor 1 (LYVE1), and insulin-like growth factor (IGFBP3).

Among 20 plasma proteins negatively associated with log<sub>2</sub> β-cryptoxanthin (Table 5), 5 met an FDR < 1%, including ORM1, TNIP1, mannosidase alpha class 1A (MAN1A1), haptoglobin (HP), and alpha-1-B glycoprotein (A1BG). Other negatively associated proteins included orosomucoid 2 (ORM2), complement factor B (CFB), complement 9 (C9), haptoglobin-related precursor (HPR), serine peptidase clade A, member 3 (SERPINA3), and lipopolysaccharide binding protein (LBP).

We examined the extent of correlation across proteins associated with log<sub>2</sub> β-cryptoxanthin, comprising the largest plasma carotenome, restricted to associations with FDR < 5% (Fig. 2). Within each of the pairs of proteins of the β-cryptoxanthin proteome, the correlation coefficients ( $r$ ) ranged from 0.28 to 0.96. We demonstrated that proteins positively and negatively associated with β-cryptoxanthin were also consistently correlated with each other in the expected directions given their associations with β-cryptoxanthin, with the exception of protein phosphatase, Mg<sup>2+</sup>/Mn<sup>2+</sup> dependent, 1 M (PPM1M) and leucine rich repeat containing 47 (LRRC47), which were also more weakly correlated with other proteins than most.

#### 4. Discussion

Provitamin A carotenoids play important roles as dietary precursors of vitamin A that may take on particular significance in impoverished regions, such as in rural Southern Asia, where vitamin A deficiency (VAD) is endemic among young children, adolescents and women of reproductive age [11]. Carotenoids also may have important antioxidant [13], immunological [14] or metabolic [39] functions and thus serve as indicators of general population health [40]. However, given their infrequent assessment, and strengthening evidence supporting the use of plasma proteomics for assessing population status with respect to other micronutrients [23–26], inflammation [27], cognition [28] and

growth [29] in this setting, we have revealed in this study protein biomarkers associated with circulating log<sub>2</sub>-normalized concentrations of six common dietary carotenoids which were plausible in their direction and strength of association.

We observed four proteins associated with β-carotene, eleven with lutein/zeaxanthin, and fifty-one with β-cryptoxanthin, all with a probability of false discovery below ten percent. APOA1, a major component of high density lipoprotein (HDL) in plasma [41], was positively associated with each of the three carotenoids, possibly reflecting shared lipoprotein transport or, co-existing antioxidant, anti-inflammatory and other metabolic functions [42]. On the other hand, TNIP1, an inhibitor of the pro-inflammatory transcription factor, NF-κB [43,44] was negatively correlated with all three carotenoids. We had also shown relative abundance of TNIP1 to be positively associated with the acute phase reactant, alpha-1-acid glycoprotein (AGP), or orosomucoid, in this population [27], explained by a negative feedback loop whereby TNIP1 is upregulated by inflammation in order to maintain immune homeostasis [44]. TNIP1 also functions as a retinoic acid receptor corepressor in the presence of its ligand [45].

Nearly all proteins negatively associated across proteomes of β-carotene, lutein/zeaxanthin, and β-cryptoxanthin were previously found to be positively correlated with inflammation markers AGP and C-reactive protein (CRP) [27]. Among these proteins, complement factor B (CFB) and complement 9 (C9) are involved in regulation of complement activation [46]; haptoglobin (HP) and haptoglobin-related precursor (HPR) are responsible for scavenging of heme iron from plasma in response to inflammation and oxidative stress in red blood cells [47,48], and these proteins were negatively associated with β-cryptoxanthin. Inflammatory proteins such as AGP isoforms of orosomucoid (ORM1) -inversely associated with β-carotene- and (ORM2) [49]. Serine peptidase inhibitors, serine peptidase clade A, member 3 (SERPINA3), also known as alpha-1-antichymotrypsin, which increases in the blood during the inflammatory response [50,51], and inter-alpha-trypsin inhibitor heavy chain H4 isoform (ITI4) as a type II

	APOA1	IFRD2	CNDP1	APOC3	PLTP	PIANP	ANTXR2	SEPP1	CLU	IGFALS	PPM1M	APOA2	PON1	LYVE1	RBP4	EIF2D	APOD	KLHL34	KLF17	PCYOX1	GSN	THBS4	APOC1	DPP4	IL1RAP	PI16	LUM	ORM1	TNIP1	MAN1A1	HP	A1BG	HPR	LRG1	SERPINA3	ORM2	LRRC47	CFB	B2M	EV15
B-Cryp	.52	.51	.5	.5	.49	.49	.53	.5	.5	.49	.57	.49	.49	.5	.49	.52	.49	.51	.51	.48	.49	.5	.49	.49	.48	.49	.49	.5	.5	.5	.53	.5	.52	.49	.49	.49	.51	.49	.49	.54
APOA1	.88	.59	.76	.72	.57	.77	.64	.68	.68	.53	.9	.73	.67	.76	.85	.79	.69	.82	.62	.71	.68	.78	.7	.63	.72	.73	.75	.73	.62	.71	.62	.69	.76	.75	.72	.55	.72	.68	.79	
IFRD2	.61	.72	.72	.71	.73	.68	.68	.67	.56	.85	.78	.69	.79	.88	.75	.66	.68	.66	.7	.67	.74	.7	.64	.72	.75	.68	.72	.65	.7	.64	.69	.77	.74	.72	.56	.73	.7	.83		
CNDP1	.66	.59	.62	.65	.67	.77	.58	.69	.68	.67	.72	.64	.69	.64	.68	.65	.68	.68	.71	.66	.64	.68	.67	.63	.62	.67	.67	.68	.7	.69	.69	.69	.47	.7	.71	.73				
APOC3	.68	.66	.72	.66	.69	.74	.65	.8	.69	.68	.81	.76	.75	.71	.78	.66	.71	.71	.79	.71	.64	.74	.72	.7	.69	.66	.72	.68	.72	.77	.74	.72	.53	.78	.73	.81				
PLTP	.62	.78	.69	.74	.66	.54	.72	.72	.77	.67	.72	.74	.74	.74	.73	.73	.75	.7	.74	.64	.73	.76	.73	.71	.67	.71	.71	.72	.72	.72	.72	.38	.73	.72	.72					
PIANP	.59	.66	.67	.65	.62	.65	.7	.62	.7	.63	.63	.59	.6	.6	.59	.66	.66	.68	.65	.64	.65	.59	.58	.64	.62	.63	.64	.7	.64	.65	.53	.7	.65	.74						
ANTXR2	.7	.74	.71	.54	.78	.75	.76	.71	.71	.72	.83	.75	.73	.81	.84	.74	.81	.6	.82	.82	.79	.74	.71	.81	.71	.79	.78	.79	.78	.47	.74	.73	.8							
SEPP1	.74	.71	.54	.74	.73	.73	.74	.68	.73	.7	.74	.71	.7	.7	.76	.73	.7	.72	.73	.67	.66	.73	.72	.72	.72	.74	.71	.71	.45	.73	.72	.79								
CLU	.78	.64	.74	.75	.78	.75	.71	.76	.75	.74	.73	.76	.73	.76	.77	.7	.76	.79	.73	.69	.74	.76	.69	.78	.73	.74	.75	.35	.74	.73	.76									
IGFALS	.62	.75	.74	.74	.81	.72	.72	.72	.72	.73	.75	.73	.75	.77	.73	.77	.78	.73	.71	.71	.75	.71	.77	.77	.75	.76	.51	.75	.75	.79										
PPM1M	.6	.61	.57	.62	.57	.5	.67	.63	.57	.6	.64	.68	.65	.61	.57	.57	.71	.69	.41	.67	.6	.6	.64	.72	.67	.49	.62	.49	.67											
APOA2	.78	.73	.84	.83	.82	.74	.85	.67	.76	.75	.85	.74	.69	.77	.79	.77	.77	.7	.79	.66	.77	.83	.82	.77	.48	.79	.77	.87												
PON1	.73	.76	.74	.75	.7	.73	.74	.73	.72	.76	.76	.71	.74	.75	.73	.73	.68	.75	.73	.74	.78	.78	.47	.76	.74	.81														
LYVE1	.75	.73	.74	.81	.7	.74	.8	.79	.75	.76	.73	.82	.84	.75	.72	.74	.78	.74	.79	.77	.79	.77	.3	.78	.75	.81														
RBP4	.79	.77	.74	.71	.72	.76	.74	.78	.74	.7	.79	.79	.73	.75	.73	.77	.7	.77	.84	.8	.77	.49	.83	.75	.87															
EIF2D	.76	.68	.67	.71	.74	.69	.75	.77	.74	.77	.77	.68	.72	.69	.72	.72	.7	.76	.74	.74	.49	.75	.73	.78																
APOD	.76	.8	.72	.78	.69	.82	.72	.7	.75	.75	.68	.61	.7	.68	.69	.72	.73	.71	.71	.53	.76	.75	.81																	
KLHL34	.75	.74	.96	.84	.74	.76	.69	.85	.85	.78	.74	.73	.77	.7	.79	.78	.81	.8	.44	.8	.76	.82																		
KLF17	.68	.74	.77	.83	.76	.7	.75	.75	.78	.71	.69	.75	.71	.73	.76	.77	.76	.55	.76	.69	.82																			
PCYOX1	.73	.71	.74	.76	.7	.74	.75	.73	.71	.7	.74	.74	.76	.74	.74	.76	.4	.74	.76	.73																				
GSN	.82	.75	.75	.69	.85	.85	.77	.73	.72	.78	.71	.8	.79	.8	.8	.44	.8	.75	.81																					
THBS4	.76	.76	.67	.87	.87	.8	.78	.72	.8	.73	.79	.8	.85	.81	.52	.8	.77	.82																						
APOC1	.73	.73	.78	.78	.74	.69	.71	.76	.73	.76	.79	.79	.76	.55	.8	.74	.84																							
DPP4	.73	.78	.8	.76	.75	.74	.81	.72	.78	.78	.76	.78	.48	.74	.77	.8																								
IL1RAP	.71	.73	.65	.63	.69	.69	.72	.71	.72	.68	.68	.28	.74	.72	.76																									
PI16	.89	.78	.77	.76	.82	.74	.8	.81	.83	.81	.49	.83	.75	.83																										
LUM	.8	.79	.76	.83	.73	.82	.82	.85	.83	.45	.82	.76	.85																											
ORM1	.95	.7	.84	.7	.81	.87	.88	.89	.57	.81	.69	.87																												
TNIP1	.68	.82	.69	.77	.87	.87	.88	.58	.79	.68	.87																													
MAN1A1	.72	.73	.75	.76	.75	.74	.47	.75	.74	.77																														
HP	.71	.87	.86	.84	.84	.54	.81	.7	.88																															
A1BG	.72	.72	.75	.75	.47	.76	.75	.77																																
HPR	.84	.83	.85	.55	.81	.76	.85																																	
LRG1	.89	.87	.63	.87	.77	.95																																		
SERPINA3	.88	.5	.86	.76	.88																																			
ORM2	.56	.84	.75	.88																																				
LRRC47	.76	.39	.73																																					
CFB	.77	.88																																						
B2M	.8																																							
EV15																																								

**Fig. 2.** Matrix of correlation coefficients (*r*) for pairs of LME-based estimates of relative abundance estimates and plasma log<sub>2</sub> β-cryptoxanthin concentration for proteins associated with β-cryptoxanthin, restricted to associations with *q* < 0.05 (*n* = 40), in 6–8 year old children of rural Nepal (*n* = 500). Blue color depicts proteins that share a positive correlation and red color a negative correlation with each other. Darker colors represent stronger association. Correlation coefficients (*r*) are presented *r* × 10<sup>2</sup> to improve visualization given in each cell. APOA2, Apolipoprotein A-II; APOC1, Apolipoprotein C-I; APOD, Apolipoprotein D; B2M, β-2-microglobulin; DPP4, Dipeptidyl-peptidase 4; EIF2D, Eukaryotic translation initiation factor 2D; EV15, ecotropic viral integration site 5; GSN, gelsolin; IL1RAP, Interleukin 1 receptor accessory protein; KLHL34, Kelch-like family member 34; KLF17, Kruppel-like factor 17; LRG1, leucine-rich α-2-glycoprotein 1; LUM, lumican; PI16, Peptidase inhibitor 16; THBS4, Thrombospondin 4.

acute-phase protein involved in the inflammatory response to trauma [52], were all negatively associated with plasma β-cryptoxanthin.

Somewhat surprisingly, the proteome for β-carotene, to our knowledge the most metabolically active carotenoid in human tissue and a specific vitamin A precursor, was quite limited in size (*n* = 4) and overlapped with that of β-cryptoxanthin, with the exception of PKM.

Despite a modest, albeit universally detectable, concentration of plasma β-cryptoxanthin in the bloodstream, its proteome was far more extensive than β-carotene's. Variation in carotenoid hydrophobicity [53] may offer one explanation for this difference. Being less hydrophobic than β-carotene, β-cryptoxanthin is more likely located on the lipoprotein surface than in the core, where β-carotene is transported, thus



allowing more extensive interactions with circulating proteins than possible with  $\beta$ -carotene. Secondly, as  $\beta$ -cryptoxanthin is known to be carried by HDL [54], it is notable that nearly half of the proteins found to be positively (ANTXR2, APOA1, APOA2, APOC1, APOC3, APOD, CLU, GPLD1, GSN, IGFALS, LUM, PCYOX1, PLTP, PON1 and RBP4) and negatively (CFB, C9, HP, ITIH4, LBP, ORM1, ORM2 and SERPINA3) associated with  $\beta$ -cryptoxanthin are known constituents of the HDL complex in human circulation [55].

Carotenoids exert their biological activity as antioxidants due to their extended conjugated carbon-carbon bonds [13]. The protective roles of carotenoids have been explored in blood plasma, where  $\beta$ -carotene, lutein, and zeaxanthin inhibited lipid peroxidation and hemoglobin oxidation but surprisingly lycopene and  $\beta$ -cryptoxanthin did not [56]. While an antioxidant function of  $\beta$ -cryptoxanthin has not been demonstrated in *in vivo* studies, we found it to be positively correlated with SEPP1, the major plasma carrier for selenium [57], an essential trace element that displays antioxidant activity by serving as an essential cofactor of glutathione peroxidase [58]. Plasma  $\beta$ -cryptoxanthin was also positively associated with PON1, an antioxidant/anti-inflammatory protein mostly synthesized by the liver and primarily associated with serum HDL [59]. To our knowledge, this is the first study demonstrating strong associations between antioxidant/anti-inflammatory PON1 and SEPP1 with plasma  $\beta$ -cryptoxanthin in a human population.

Plasma  $\beta$ -cryptoxanthin was positively correlated with relative abundance of IGFALS, IGFBP3, CNDP1, and cartilage oligomeric matrix protein (COMP), proteins that we have previously reported to be positively associated with child height and arm muscle mass in this population of school-aged Nepalese children [29], suggesting that  $\beta$ -cryptoxanthin nutriture, as reflected in plasma, is associated with general nutritional status, although mechanisms explaining this relationship remain unknown.

Lutein and zeaxanthin, measured together, were associated with an intermediate proteome of 11 proteins. While present in plasma, lutein and zeaxanthin are concentrated in the macula, the central region of the retina [60]. These macular carotenoids protect the retina from light-induced damage via filtering blue light [61]. Both lutein and zeaxanthin are effective antioxidants like other major carotenoids found in human plasma [62]. Lutein has been shown to protect against inflammation, by reducing the production of pro-inflammatory factors observed in retinal injury [63]. There was a positive correlation between plasma lutein/zeaxanthin and proteoglycan 4 (PRG4), a glycoprotein recently identified at the ocular surface where it functions as a lubricant [64] and its loss results in inflammation [65].

In summary, a plasma proteomics approach has revealed an extensive proteome that covaries with relative abundance of  $\beta$ -cryptoxanthin, despite its low circulating concentration in a generally undernourished rural population of Nepalese school-aged children. The number and diversity of plasma proteins associated with  $\beta$ -cryptoxanthin suggests involvement in vitamin A metabolism, lipid transport and immunoregulation. Moreover, for the first time, we speculate an *in vivo* antioxidant function of  $\beta$ -cryptoxanthin. Our findings suggest that plasma proteins could be measured in populations as surrogates for carotenoid intake or status, and help reveal protein:carotenoid functional relationships. More work is merited in this line of study to verify our findings and probe the implications of these novel findings for carotenoid assessment, metabolism and function and health.

## Acknowledgements

We thank the Johns Hopkins nutriproteomics research team including Margia Arguello, Raghobhama Chaerkady, Hongie Cui, Lauren R DeVine, Jaime Johnson, Robert O'Meally, Steven C LeClerg, Ashika Nanayakkara-Bind, Hee-Sool Rho, Sudeep Shrestha, and Fredrick Van Dyke for assistance with the nutriproteomics research and Nepal field activities. We thank Ingo Ruczinski for his proteomics modeling and

analytic guidance and C Conover Talbot Jr. for assistance with the Human Genome Organization gene annotation.

## Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.abb.2018.03.025>.

## References

- [1] G. Britton, S. Liaaen-Jensen, H. Pfander, Carotenoids: Natural Functions, vol. 4, Switzerland Springer Press, Birkhäuser, Basel, 2008.
- [2] S. Lu, L. Li, Carotenoid metabolism: biosynthesis, regulation, and beyond, *J. Integr. Plant Biol.* 50 (2008) 778–785.
- [3] J.A. Maresca, J.E. Graham, A.D. Bryant, The biochemical basis for structural diversity in the carotenoids of chlorophototrophic bacteria, *Photosynth. Res.* 97 (2008) 121–140.
- [4] S. Takaichi, M. Mochimaru, Carotenoids and carotenogenesis in cyanobacteria: unique ketocarotenoids and carotenoid glycosides, *Cell. Mol. Life Sci.* 64 (2007) 2607–2619.
- [5] Z.W. Ye, J.G. Jiang, G.H. Wu, Biosynthesis and regulation of carotenoids in *Dunaliella*: progresses and prospects, *Biotechnol. Adv.* 26 (2008) 352–360.
- [6] J.E. Romanchik, D.W. Morel, E.H. Harrison, Distributions of carotenoids and  $\alpha$ -tocopherol among lipoproteins do not change when human plasma is incubated *in vitro*, *J. Nutr.* 125 (1995) 2610–2617.
- [7] A. Eroglu, E.H. Harrison, Carotenoid metabolism in mammals, including man: formation, occurrence, and function of apocarotenoids, *JLR (J. Lipid Res.)* 54 (2013) 1719–1730.
- [8] S. Narayanasamy, J. Sun, E.R. Pavlovicz, A. Eroglu, E.C. Rush, B.D. Sunkel, C. Li, E.H. Harrison, R.W. Curley, Synthesis and biological effects of Apo-11-, Apo-13-, and Apo-15-lycopenoids, cleavage products of lycopene that function as retinoic acid receptor antagonists, *JLR (J. Lipid Res.)* 58 (2017) 1021–1029.
- [9] V. Singh, K.P. West Jr., Vitamin A deficiency and xerophthalmia among school-aged children in Southeastern Asia, *Eur. J. Clin. Nutr.* 58 (2004) 1342–1349.
- [10] World Health Organization, Global Prevalence of Vitamin A Deficiency in Populations at Risk 1995–2005. WHO Global Database on Vitamin A Deficiency, WHO, Geneva, Switzerland, 2009.
- [11] K.P. West Jr., Extent of vitamin A deficiency among preschool children and women of reproductive age, *J. Nutr.* 132 (2002) 2857–2866.
- [12] A. Sommer, K.P. West Jr., Vitamin A Deficiency: Health, Survival and Vision, Oxford University Press, New York, 1996.
- [13] G.W. Burton, K.U. Ingold, beta-Carotene: an unusual type of lipid antioxidant, *Science* 224 (1984) 569–573.
- [14] L. Rubin, A.C. Ross, C.B. Stephensen, T. Bohn, S.A. Tanumihardjo, Metabolic effects of inflammation on vitamin a and carotenoids in humans and animal models, *Adv Nutr* 8 (2017) 197–212.
- [15] G. Riccioni, N. D'Orazio, L. Speranza, E. Di Ilio, M. Glade, V. Bucciarelli, L. Scotti, F. Martini, A. Pennelli, T. Bucciarelli, Carotenoids and asymptomatic carotid atherosclerosis, *J. Biol. Regul. Homeost. Agents* 24 (2010) 447–452.
- [16] M. Chatterjee, M. Janarthan, Biological activity of carotenoids: its implications in cancer risk and prevention, *Curr. Pharmaceut. Biotechnol.* 13 (2012) 180–190.
- [17] H. Zhou, X. Zhao, E.J. Johnson, A. Lim, E. Sun, J. Yu, Y. Zhang, X. Liu, T. Snellingen, F. Shang, N. Liu, Serum carotenoids and risk of age-related macular degeneration in a Chinese population sample, *Invest. Ophthalmol. Vis. Sci.* 52 (2011) 4338–4344.
- [18] N.I. Krinsky, J.I. Landrum, R.A. Bone, Biological mechanisms of the protective role of lutein and zeaxanthin in the eye, *Annu. Rev. Nutr.* 23 (2003) 171–201.
- [19] W.K. Al-Delaimy, N. Slimani, P. Ferrari, T. Key, E. Spencer, I. Johansson, G. Johansson, I. Mattisson, E. Wirfalt, S. Sieri, et al., Plasma carotenoids as biomarkers of intake of fruits and vegetables: ecological level correlations in the European Prospective Investigation into Cancer and Nutrition (EPIC), *Eur. J. Clin. Nutr.* 59 (2005) 1397–1408.
- [20] S.A. McNaughton, G.C. Marks, P. Gaffney, G. Williams, A. Green, Validation of a food-frequency questionnaire assessment of carotenoid and vitamin E intake using weighed food records and plasma biomarkers: the method of triads model, *Eur. J. Clin. Nutr.* 59 (2005) 211–218.
- [21] A.M. Hodge, J.A. Simpson, M. Fridman, K. Rowley, D.R. English, G.G. Giles, Q. Su, K. O'Dea, Evaluation of an FFQ for assessment of antioxidant intake using plasma biomarkers in an ethnically diverse population, *Publ. Health Nutr.* 12 (2009) 2438–2447.
- [22] D. Hess, H.E. Keller, B. Oberlin, R. Bonfanti, W. Schu'ep, Simultaneous determination of retinol, tocopherols, carotenes and lycopene in plasma by means of high-performance liquid-chromatography on reversed phase, *Int. J. Vitam. Nutr. Res.* 61 (1991) 232–238.
- [23] R.N. Cole, I. Ruczinski, K. Schulze, P. Christian, S. Herbrich, L. Wu, L.R. DeVine, R.N. O'Meally, S. Shrestha, T.N. Boronina, et al., The plasma proteome identifies expected and novel proteins correlated with micronutrient status in undernourished Nepalese children, *J. Nutr.* 143 (2013) 1540–1548.
- [24] K.P. West Jr., R.N. Cole, S. Shrestha, K.J. Schulze, S.E. Lee, J. Betz, B.A. Nonyane, L.S. Wu, J.D. Yager, J.D. Groopman, P. Christian, A plasma alpha-tocopherome can be identified from proteins associated with vitamin E status in school-aged children of Nepal, *J. Nutr.* 12 (2015) 2646–2656.
- [25] K.J. Schulze, R.N. Cole, R. Chaerkady, L.S. Wu, B.A. Nonyane, S.E. Lee, J.D. Yager,

- J.D. Groopman, P. Christian, K.P. West Jr., Plasma selenium protein P isoform 1 (SEPP1): a predictor of selenium status in Nepalese children detected by plasma proteomics, *Int. J. Vitam. Nutr. Res.* (2016 May 10), <http://dx.doi.org/10.1024/0300-9831/a000256> Epub ahead of print.
- [26] S.E. Lee, K.J. Schulze, L.S. Wu, J.D. Yager, P. Christian, K.P. West Jr., Biological systems of vitamin K: a plasma nutripoteomics study of subclinical vitamin K deficiency in 500 Nepalese children, *Omic* 4 (2016) 214–223.
- [27] S.E. Lee, K.P. West Jr., R.N. Cole, K.J. Schulze, P. Christian, L.S. Wu, J.D. Yager, J.D. Groopman, I. Ruczinski, Plasma proteome biomarkers of inflammation in school aged children in Nepal, *PLoS One* (2015) 10, <http://dx.doi.org/10.1371/journal.pone.0144279> e0144279.
- [28] S.E. Lee, K.P. West Jr., R.N. Cole, K.J. Schulze, L.S. Wu, J.D. Yager, J.D. Groopman, P. Christian, General intelligence is associated with subclinical inflammation in Nepalese children: a population-based plasma proteomics study, *Brain Behav. Immun.* 56 (2016) 253–263.
- [29] S.E. Lee, C.P. Stewart, K.J. Schulze, R.N. Cole, L.S. Wu, J.D. Yager, J.D. Groopman, S.K. Khatri, R.K. Adhikari, P. Christian, K.P. West Jr., The plasma proteome is associated with anthropometric status of undernourished Nepalese school-aged children, *J. Nutr.* 147 (2017) 304–313.
- [30] P. Christian, S.K. Khatri, J. Katz, E.K. Pradhan, S.C. LeClerq, S.R. Shrestha, R.K. Adhikari, A. Sommer, K.P. West Jr., Effects of alternative maternal micro-nutrient supplements on low birth weight in rural Nepal: double blind randomized community trial, *BMJ* 326 (2003) 571–577.
- [31] K.J. Schulze, P. Christian, L.S.-F. Wu, M. Arguello, H. Cui, A. Nanayakkara-Bind, C. Stewart, S.K. Khatri, S. LeClerq, K.P. West Jr., Micronutrient deficiencies are common in 6- to 8-year-old children of rural Nepal, with prevalence estimates modestly affected by inflammation, *J. Nutr.* 144 (2014) 979–987.
- [32] WHO, Growth Reference Data for 5–19 Years, WHO, Geneva, 2007 [cited 2012 Oct 11]. Available from: <http://www.who.int/growthref/en/>.
- [33] S. Yamini, K.P. West Jr., L. Wu, M.L. Dreyfuss, D.X. Yang, S.K. Khatri, Circulating levels of retinol, tocopherol and carotenoid in Nepali pregnant and postpartum women following long-term beta-carotene and vitamin A supplementation, *Eur. J. Clin. Nutr.* 55 (2001) 252–259.
- [34] S.M. Herbrich, R.N. Cole, K.P. West Jr., K. Schulze, J.D. Yager, J.D. Groopman, P. Christian, L. Wu, R.N. O Meally, D.H. May, et al., Statistical inference from multiple iTRAQ experiments without using common reference standards, *J. Proteome Res.* 12 (2013) 594–604.
- [35] J.D. Storey, A direct approach to false discovery rates, *J. Roy. Stat. Soc. B* 64 (2002) 479–498.
- [36] G.K. Robinson, That BLUP is a good thing: the estimation of random effects, *Stat. Sci.* 6 (1991) 15–32.
- [37] National Center for Biotechnology Information. Genetic sequence data bank [Internet], Bethesda (MD): National Library of Medicine (US). c2016 [cited 2016 Oct 23]. Available from: <ftp://ftp.ncbi.nih.gov/genbank/gbrel.txt>.
- [38] The Comprehensive R Archive Network [Internet], Vienna (Austria): Institute for Statistics and Mathematics, Vienna University of Economics and Business, c2004 [cited 2015 May 27]. Available from: <http://cran.r-project.org>.
- [39] M.A. Beydoun, M.R. Shroff, X. Chen, H.A. Beydoun, Y. Wang, A.B. Zonderman, Serum antioxidant status is associated with metabolic syndrome among US adults in recent national surveys, *J. Nutr.* 141 (2011) 903–913.
- [40] M.D. Shardell, D.E. Alley, G.E. Hicks, S.S. El-Kamary, R.R. Miller, R.D. Semba, L. Ferrucci, Low-serum carotenoid concentrations and carotenoid interactions predict mortality in US adults: the third national health and nutrition examination survey, *Nutr. Res.* 31 (2011) 178–189.
- [41] A.S. Shah, L. Tan, J. Lu Long, W.S. Davidson, The proteomic diversity of high density lipoproteins: our emerging understanding of its importance in lipid transport and beyond, *J. Lipid Res.* 54 (2013) 2575–2585.
- [42] M. Mangaraj, R. Nanda, S. Panda, A.-I. Apolipoprotein, A molecule of diverse function, *Indian J. Clin. Biochem.* 31 (2016) 253–259.
- [43] K. Enesa, H.P. Moll, L. Luong, C. Ferran, P.C. Evans, A20 suppresses vascular inflammation by recruiting proinflammatory signaling molecules to intracellular aggregates, *Faseb. J.* 29 (2015) 1869–1878.
- [44] L. Verstrepen, K. Verhelst, G. van Loo, I. Carpentier, S.C. Ley, R. Beyaert, Expression, biological activities and mechanisms of action of A20 (TNFAIP3), *Biochem. Pharmacol.* 80 (2010) 2009–2020.
- [45] I. Gurevich, B.J. Aneskievich, Liganded RARalpha and RARgamma interact with but are repressed by TNIP1, *Biochem. Biophys. Res. Commun.* 389 (2009) 409–414.
- [46] K.R. Mayilyan, Complement genetics, deficiencies, and disease associations, *Protein Cell* 3 (2012) 487–496.
- [47] M.J. Nielsen, H.J. Moller, S.K. Moestrup, Hemoglobin and heme scavenger receptors, *Antioxidants Redox Signal.* 12 (2010) 261–273.
- [48] M.J. Nielsen, S.V. Petersen, C. Jacobsen, C. Oxvig, D. Rees, H.J. Moller, S.K. Moestrup, Haptoglobin-related protein is a high-affinity hemoglobin-binding plasma protein, *Blood* 108 (2006) 2846–2849.
- [49] T. Hocheppied, F.G. Berger, H. Baumann, C. Libert,  $\alpha$ 1-Acid glycoprotein: an acute phase protein with inflammatory and immunomodulating properties, *Cytokine Growth Factor Rev.* 14 (2003) 25–34.
- [50] N.A. Kalsheker, Alpha-1-antichymotrypsin, *Int. J. Biochem. Cell Biol.* 28 (1996) 961–964.
- [51] S. Janciauskiene, Conformational properties of serine proteinase inhibitors (serpins) confer multiple pathophysiological roles, *Biochim. Biophys. Acta* 1535 (2001) 221–235.
- [52] M. Piñeiro, M. Andrés, M. Iturralde, et al., ITIH4 (inter-alpha-trypsin inhibitor heavy chain 4) is a new acute-phase protein isolated from cattle during experimental infection, *Infect. Immun.* 72 (2004) 3777–3782.
- [53] R.S. Parker, Absorption, metabolism, and transport of carotenoids, *FASEB J.* 10 (1996) 542–551.
- [54] S. Goulinet, M.J. Chapman, Plasma LDL and HDL subspecies are heterogeneous in particle content of tocopherols and oxygenated and hydrocarbon carotenoids: relevance to oxidative resistance and atherogenesis, *Arterioscler. Thromb. Vasc. Biol.* 17 (1997) 786–796.
- [55] T. Vaisar, Proteomics investigations of HDL: challenges and promise, *Curr. Vasc. Pharmacol.* 10 (2012) 410–421.
- [56] R.C. Chisté, M. Freitas, A.Z. Mercadante, E. Fernandes, Carotenoids inhibit lipid peroxidation and hemoglobin oxidation, but not the depletion of glutathione induced by ROS in human erythrocytes, *Life Sci.* 99 (2014) 52–60.
- [57] R.F. Burk, K.E. Hill, Selenoprotein P-expression, functions, and roles in mammals, *Biochim. Biophys. Acta* 1790 (2009) 1441–1447.
- [58] J.T. Rotruck, A.L. Pope, H.E. Ganther, A.B. Swanson, D.G. Hafeman, W.G. Hoekstra, Selenium: biochemical role as a component of glutathione peroxidase, *Science* 179 (1973) 588–590.
- [59] S. Deakin, I. Leviev, M. Gomaschi, L. Calabresi, G. Franceschini, R.W. James, Enzymatically active paraoxonase-1 is located at the external membrane of producing cells and released by a high affinity, saturable, desorption mechanism, *J. Biol. Chem.* 277 (2002) 4301–4308.
- [60] R.A. Bone, J.T. Landrum, S.L. Tarsis, Preliminary identification of the human macular pigment, *Vis. Res.* 25 (1985) 1531–1535.
- [61] F.M. Barker, D.M. Snodderly, E.J. Johnson, W. Schalch, W. Koepcke, J. Gerss, M. Neuringer, Nutritional manipulation of primate retinas, V: effects of lutein, zeaxanthin, and n-3 fatty acids on retinal sensitivity to blue-light-induced damage, *Invest. Ophthalmol. Vis. Sci.* 52 (2011) 3934–3942.
- [62] N.I. Krinsky, E.J. Johnson, Carotenoid actions and their relation to health and disease, *Mol. Aspect. Med.* 6 (2005) 459–516.
- [63] M. Sasaki, Y. Ozawa, T. Kurihara, et al., Neuroprotective effect of an antioxidant, lutein, during retinal inflammation, *Invest. Ophthalmol. Vis. Sci.* 50 (2009) 1433–1439.
- [64] M.L. Samson, S. Morrison, N. Masala, B.D. Sullivan, D.A. Sullivan, H. Sheardown, T.A. Schmidt, Characterization of full-length recombinant human proteoglycan 4 as an ocular surface boundary lubricant, *Exp. Eye Res.* 127 (2014) 14–19.
- [65] T.A. Schmidt, D.A. Sullivan, E. Knop, S.M. Richards, N. Knop, S.H. Liu, A. Sahin, R.R. Darabad, S. Morison, W.R. Kam, et al., Transcription, translation, and function of proteoglycan 4, a boundary lubricant, at the ocular surface, *Jama Ophthalmol* 131 (2013) 766–776.