

# SCIENTIFIC REPORTS



OPEN

## Metabolomic Profiling of Serum Retinol in the Alpha-Tocopherol, Beta-Carotene Cancer Prevention (ATBC) Study

Jiaqi Huang<sup>1</sup>, Orestis A. Panagiotou<sup>1</sup> , Gabriella M. Anic<sup>1</sup>, Alison M. Mondul<sup>2</sup>, Linda M. Liao<sup>1</sup>, Andriy Derkach<sup>1</sup>, Rachael Stolzenberg-Solomon<sup>1</sup>, Stephanie J. Weinstein<sup>1</sup> & Demetrius Albanes<sup>1</sup>

The role of retinol in the prevention of multifactorial chronic diseases remains uncertain, and there is sparse evidence regarding biological actions and pathways implicated in its effects on various outcomes. The aim is to investigate whether serum retinol in an un-supplemented state is associated with low molecular weight circulating metabolites. We performed a metabolomic analysis of 1,282 male smoker participants based on pre-supplementation fasting serum in the Alpha-Tocopherol, Beta-Carotene Cancer Prevention (ATBC) Study. We examined the association between 947 metabolites measured by ultra-high performance LC-MS/GC-MS and retinol concentration (from HPLC) using linear regression that estimated the difference in metabolite concentrations per unit difference in retinol concentration as standardized  $\beta$ -coefficients and standard errors (SE). We identified 63 metabolites associated with serum retinol below the Bonferroni-corrected P-value ( $p < 5.3 \times 10^{-5}$ ). The strongest signals were for N-acetyltryptophan ( $\beta = 0.27$ ; SE = 0.032;  $p = 9.8 \times 10^{-17}$ ), myo-inositol ( $\beta = 0.23$ ; SE = 0.032;  $p = 9.8 \times 10^{-13}$ ), and 1-palmitoylglycerophosphoethanolamine ( $\beta = 0.22$ ; SE = 0.032;  $p = 3.2 \times 10^{-12}$ ). Several chemical class pathways were strongly associated with retinol, including amino acids ( $p = 1.6 \times 10^{-10}$ ), lipids ( $p = 3.3 \times 10^{-7}$ ), and cofactor/vitamin metabolites ( $3.3 \times 10^{-7}$ ). The strongest sub-pathway association was for inositol metabolism ( $p = 2.0 \times 10^{-14}$ ). Serum retinol concentration is associated with circulating metabolites in various metabolic pathways, particularly lipids, amino acids, and cofactors/vitamins. These interrelationships may have relevance to the biological actions of retinol, including its role in carcinogenesis.

The importance of vitamin A compounds for homeostasis and normal physiology is well-established<sup>1</sup>, and its key biological functions include critical roles in embryonic development and growth, cell differentiation, tissue remodeling, reproduction, integrity of the immune system, vision, maintenance of skin and membranes, and hematopoiesis<sup>2-6</sup>. Vitamin A deficiency can lead to xerophthalmia, blindness, infections, and even death, especially in low-income settings<sup>7</sup>. Retinol and vitamin A compounds including retinoic acid are also integral to lipid metabolism, insulin signaling, and energy balance<sup>4-6</sup>, and they exert pleiotropic effects by regulating or co-regulating the expression of over 500 genetic response elements after binding with its nuclear receptors retinoic acid receptor (RAR), retinoid X receptor (RXR), and peroxisome proliferator-activated receptor  $\beta/\delta$  (PPAR  $\beta/\delta$ ), as well as heterodimerizing with the vitamin D and thyroid hormone receptors<sup>8</sup>. These biological actions are thought to be responsible for their experimental anti-carcinogenic effects<sup>9-12</sup>, although increased cell proliferation and decreased cell differentiation have also been observed<sup>13</sup>.

The role of vitamin A in the prevention of common chronic diseases is less clear, however<sup>14, 15</sup>. Early pre-clinical and population-based observational studies<sup>10</sup> suggested protective effects of retinol on cancer<sup>11, 16</sup> and cardiovascular disease<sup>17-19</sup>. By contrast, subsequent evidence from randomized trials and meta-analyses<sup>20</sup>

<sup>1</sup>Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA. <sup>2</sup>Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, MI, USA. Jiaqi Huang and Orestis A. Panagiotou contributed equally to this work. Correspondence and requests for materials should be addressed to D.A. (email: [DAA@NIH.GOV](mailto:DAA@NIH.GOV))

Characteristic <sup>1</sup>	
Age at randomization (years), median (IQR)	58 (54–63)
Height (cm)	173.5 (6.1)
Weight (kg)	79.0 (12.9)
BMI (kg/m <sup>2</sup> )	26.2 (3.8)
<b>Smoking, median (IQR)</b>	
cigarettes/day	20 (14–25)
duration (years)	38 (32–43)
Physical activity (3 or more times/week, %)	21
<b>Dietary intake, median (IQR)</b>	
total energy, (kcal/d)	2640 (2170–3128)
fruit (g/d)	109 (60–175)
vegetables (g/d)	99 (63–149)
red meat (g/d)	66 (50–89)
coffee (g/d)	560 (440–840)
alcohol, ethanol (g/d)	10 (2–24)
vitamin A (µg/d)	1627 (1128–2248)
<b>Intervention group, N (%)</b>	
placebo	327 (26)
beta-carotene only	342 (27)
alpha-tocopherol only	299 (23)
alpha-tocopherol and beta-carotene	314 (24)
Serum retinol (µg/L), median (IQR)	579 (511–659)
Serum total cholesterol (mmol/L), median (IQR)	6.22 (5.45–6.94)

**Table 1.** Pre-randomization characteristics of 1,282 Finnish male smokers in the ATBC Study<sup>1</sup>Data for continuous variables are shown as mean (standard deviation) unless otherwise specified. IQR: interquartile range.

have not supported many of the observational findings<sup>21</sup>, instead showing increased risk for some outcomes such as cancer of the prostate<sup>22,23</sup> and lung<sup>24</sup>, cardiovascular disease<sup>25</sup>, and even overall mortality<sup>24</sup> for individuals with high circulating retinol concentration or following supplementation with vitamin A or  $\beta$ -carotene<sup>14,23,24,26</sup>. For example, a recent pooled analysis of 15 cohort studies that included more than 11,000 prostate cancer cases found 13% higher prostate cancer risk in the highest versus lowest category of serum retinol<sup>23</sup>. Such data have led the U.S. Preventive Services Task Force to question the public health benefits of supplementation with vitamin A in the absence of deficiency<sup>14,27</sup>.

Despite evidence from molecular and laboratory studies that metabolic derivatives of retinol could promote carcinogenesis<sup>28,29</sup>, the relevant biologic pathways are not understood. Elucidating the biological mechanisms underlying these vitamin A associations would have implications for any future prevention trials, selection of their target populations, and general population vitamin A supplementation<sup>30</sup>.

Here, we hypothesized that serum retinol in an un-supplemented state may be associated with other small, low-molecular metabolites in circulation, and conducted an agnostic metabolomic analysis to identify biologically relevant metabolites related to vitamin A status.

## Results

Characteristics at study entry for the 1,282 participants included in this analysis are shown in Table 1. The median serum concentration of retinol was 579 µg/L. The median number of cigarettes smoked per day was 20 (interquartile range, IQR, 14–25), and 21% of participants exercised 3 times per week or more. Serum retinol <0.7 µmol/L (or 200 µg/L) is an accepted definition of vitamin A deficiency in most age groups<sup>31,32</sup>. In the present subset of 1,208 men, only one had a borderline-deficient serum retinol value (i.e., 192 µg/L).

In multivariable linear regression models, 263 metabolites were associated with retinol concentrations at  $p < 0.05$  (Supplemental Table 1), of which 63 (Table 2) were significant after Bonferroni correction ( $p < 5.3 \times 10^{-5}$ ). All of them were positively associated with a standard deviation (SD) increase in retinol, except for the peptide metabolite adsgedfxaeggvr (fibrinopeptide A) that showed an inverse association ( $\beta = -0.16$ ). The strongest association with serum retinol was seen for N-acetyltryptophan ( $\beta = 0.27$  per SD increase in retinol concentration;  $p = 9.9 \times 10^{-17}$ ), followed by the lipids *myo*-inositol and 1-palmitoylglycerophosphoethanolamine, the amino acid 4-acetamidobutanoate, and the purine N6-carbamoylthreonyladenosine ( $\beta = 0.23$ ,  $p = 9.8 \times 10^{-13}$ ;  $\beta = 0.22$ ,  $p = 3.2 \times 10^{-12}$ ;  $\beta = 0.22$ ,  $p = 8.9 \times 10^{-12}$ ;  $\beta = 0.21$ ,  $p = 2.9 \times 10^{-11}$ , respectively). The findings were not altered by adjustment for serum creatinine. Also, results were the same after model adjustment for storage time (from blood collection to metabolomics measurement) (data not shown).

Of the metabolites that remained statistically significant after Bonferroni correction, the majority were lipids ( $n = 30$  of 264 lipids) and amino acids ( $n = 20$  of 157 amino acids). The remaining were nucleotides ( $n = 4$  of 31), xenobiotics ( $n = 4$  of 124), carbohydrates ( $n = 2$  of 26), energy metabolites ( $n = 1$  of 9), and peptides ( $n = 1$  of 91).

Metabolite <sup>1</sup>	Estimate <sup>2</sup>	Standard error	p value	Coefficient of variation	Chemical Sub-Class
<b>Amino acid and amino acid derivatives</b>					
N-Acetyltryptophan	0.272	0.0321	$9.89 \times 10^{-17}$	0.15	Tryptophan metabolism
4-Acetamidobutanoate	0.216	0.0314	$8.87 \times 10^{-12}$	0.18	Guanidino and acetamido metabolism; polyamine metabolism
N-Acetyl-3-methylhistidine	0.217	0.0331	$9.84 \times 10^{-11}$	0.49	Histidine metabolism
Homocitrulline	0.201	0.0320	$4.87 \times 10^{-10}$	0.14	Urea cycle; arginine and proline metabolism
N-Acetylvaline	0.192	0.0314	$1.32 \times 10^{-09}$	0.15	Valine, leucine and isoleucine metabolism
N-Acetyl-1-methylhistidine	0.200	0.0339	$4.65 \times 10^{-09}$	0.10	Histidine metabolism
1-Methylhistidine	0.187	0.0322	$8.41 \times 10^{-09}$	0.13	Histidine metabolism
N6-Acetyllysine	0.179	0.0323	$4.23 \times 10^{-08}$	0.08	Lysine metabolism
Creatinine	0.172	0.0322	$1.11 \times 10^{-07}$	0.13	Creatine metabolism
N-Acetylthreonine	0.162	0.0319	$4.51 \times 10^{-07}$	0.11	Glycine, serine and threonine metabolism
Beta-hydroxyisovaleryl carnitine	0.181	0.0363	$8.24 \times 10^{-07}$	0.14	Valine, leucine and isoleucine metabolism
Indolelactate	0.159	0.0321	$8.99 \times 10^{-07}$	0.09	Tryptophan metabolism
N-Delta-acetylorithine	0.157	0.0321	$1.14 \times 10^{-06}$	0.24	Urea cycle; arginine and proline metabolism
N2,N5-Diacetylorithine	0.204	0.0425	$2.06 \times 10^{-06}$	0.13	Urea cycle; arginine and proline metabolism
Urea	0.151	0.0323	$3.47 \times 10^{-06}$	0.12	Urea cycle; arginine and proline metabolism
3-Methoxytyramine sulfate	0.197	0.0432	$6.58 \times 10^{-06}$	0.23	Phenylalanine & tyrosine metabolism
Guanidinosuccinate	0.182	0.0414	$1.32 \times 10^{-05}$	0.38	Guanidino and acetamido metabolism
3-Methylhistidine	0.140	0.0327	$1.97 \times 10^{-05}$	0.39	Histidine metabolism
Isovaleryl carnitine	0.132	0.0310	$2.35 \times 10^{-05}$	0.12	Valine, leucine and isoleucine metabolism
Tiglyl carnitine	0.138	0.0326	$2.56 \times 10^{-05}$	0.10	Valine, leucine and isoleucine metabolism
<b>Amino acid and amino acid derivatives; xenobiotics</b>					
2-Hydroxyisobutyrate	0.205	0.0357	$1.22 \times 10^{-08}$	0.10	Valine, leucine and isoleucine metabolism; chemical
<b>Lipids</b>					
Myo-inositol	0.228	0.0316	$9.82 \times 10^{-13}$	0.17	Inositol metabolism
1-Palmitoylglycerophosphoethanolamine	0.223	0.0316	$3.16 \times 10^{-12}$	0.35	Lysolipid
4-Androsten-3beta,17beta-diol monosulfate	0.254	0.0397	$3.39 \times 10^{-10}$	0.07	Sterol/steroid
1-Palmitoleoylglycerophosphocholine	0.195	0.0312	$5.74 \times 10^{-10}$	0.57	Lysolipid
Propionyl carnitine	0.198	0.0316	$6.20 \times 10^{-10}$	0.23	Fatty acid metabolism (also BCAA metabolism)
2-Palmitoleoylglycerophosphocholine	0.199	0.0323	$1.10 \times 10^{-9}$	0.78	Lysolipid
4-Androsten-3beta,17beta-diol disulfate	0.180	0.0301	$3.35 \times 10^{-9}$	0.04	Sterol/steroid
5Alpha-androstan-3beta,17beta-diol disulfate	0.189	0.0311	$1.60 \times 10^{-9}$	0.07	Sterol/steroid
Stearoyl-arachidonoyl-glycerophosphoethanolamine	0.223	0.0409	$7.74 \times 10^{-8}$	0.12	Lysolipid
Etiocolanolone glucuronide	0.210	0.0420	$7.91 \times 10^{-7}$	0.12	Sterol/steroid
1-Linolenoylglycerophosphocholine (18:3n3)	0.175	0.0354	$9.65 \times 10^{-7}$	0.41	Lysolipid
4-Androsten-3beta,17beta-diol disulfate	0.157	0.0318	$9.83 \times 10^{-7}$	0.06	Sterol/steroid
Inositol 1-phosphate (I1P)	0.148	0.0302	$1.06 \times 10^{-6}$	0.24	Inositol metabolism
Palmitoyl-oleoyl-glycerophosphocholine	0.201	0.0412	$1.31 \times 10^{-6}$	0.09	Lysolipid
Cortisol	0.157	0.0324	$1.43 \times 10^{-6}$	0.07	Sterol/steroid
1-Oleoylglycerol (1-monoolein)	0.160	0.0338	$2.51 \times 10^{-6}$	0.34	Monoacylglycerol
Acetylcarnitine	0.147	0.0315	$3.73 \times 10^{-6}$	0.28	Carnitine metabolism
2-Palmitoylglycerophosphocholine	0.148	0.0321	$4.26 \times 10^{-6}$	0.59	Lysolipid
Malonylcarnitine	0.196	0.0424	$4.57 \times 10^{-6}$	0.30	Fatty acid synthesis
Carnitine	0.140	0.0318	$1.20 \times 10^{-5}$	0.10	Carnitine metabolism
Scyllo-inositol	0.262	0.0591	$1.31 \times 10^{-5}$	0.48	Inositol metabolism
1-Oleoylglycerophosphocholine	0.137	0.0315	$1.47 \times 10^{-5}$	0.43	Lysolipid
2-Myristoylglycerol (2-monomyristin)	0.156	0.0363	$1.94 \times 10^{-5}$	0.42	Monoacylglycerol
1-Oleoylglycerophosphoethanolamine	0.138	0.0322	$2.13 \times 10^{-5}$	0.40	Lysolipid
1-Palmitoylglycerophosphocholine	0.138	0.0322	$2.17 \times 10^{-5}$	0.28	Lysolipid
2-Palmitoylglycerol (2-monopalmitin)	0.148	0.0347	$2.33 \times 10^{-5}$	0.43	Monoacylglycerol
21-Hydroxypregnenolone disulfate	0.135	0.0321	$2.94 \times 10^{-5}$	0.06	Sterol/steroid
Palmitoyl-arachidonoyl-glycerophosphocholine	0.165	0.0395	$3.42 \times 10^{-5}$	0.19	Lysolipid
1-Linolenoylglycerol	0.138	0.0332	$3.72 \times 10^{-5}$	0.20	
1-Stearoylglycerophosphoethanolamine	0.131	0.0321	$4.64 \times 10^{-5}$	0.42	Lysolipid
Continued					

Metabolite <sup>1</sup>	Estimate <sup>2</sup>	Standard error	p value	Coefficient of variation	Chemical Sub-Class
<b>Nucleotides</b>					
N6-Carbamoylthreonyladenosine	0.212	0.0315	$2.91 \times 10^{-11}$	0.09	Purine metabolism, guanine containing; purine metabolism, adenine containing
Urate	0.205	0.0308	$4.14 \times 10^{-11}$	0.07	Purine metabolism, urate metabolism; purine metabolism, (Hypo)xanthine/inosine containing
Pseudouridine	0.191	0.0318	$2.81 \times 10^{-9}$	0.1	Pyrimidine metabolism, uracil containing
N2,N2-Dimethylguanosine	0.134	0.0317	$2.76 \times 10^{-5}$	0.2	Purine metabolism, guanine containing
<b>Carbohydrates</b>					
Erythronate	0.198	0.0313	$4.01 \times 10^{-10}$	0.14	Aminosugar metabolism
Arabitol	0.188	0.0308	$1.39 \times 10^{-9}$	0.14	Pentose metabolism
<b>Xenobiotics</b>					
Erythritol	0.192	0.0313	$1.10 \times 10^{-9}$	0.14	Sugar, sugar substitute, starch; food component/plant
O-Sulfo-L-tyrosine	0.217	0.0358	$2.13 \times 10^{-9}$	0.08	Chemical
Ethyl glucuronide	0.164	0.0319	$3.00 \times 10^{-7}$	0.06	Detoxification and metabolism; chemical
3-Hydroxycotinine glucuronide	0.143	0.0302	$2.34 \times 10^{-6}$	0.08	Tobacco metabolite
<b>Peptides</b>					
Adsgedfxaeggvr	-0.157	0.0353	$9.38 \times 10^{-6}$	0.05	Fibrinogen cleavage peptide
<b>Energy</b>					
Succinylcarnitine	0.134	0.0315	$2.37 \times 10^{-5}$	0.38	Krebs cycle / TCA cycle

**Table 2.** Metabolites associated with serum retinol concentration at  $5.3 \times 10^{-5}$  level of statistical significance. <sup>1</sup>All metabolites had detectable values in > 80% of study population with the exception of N-acetyl-3-methylhistidine, N-acetyl-1-methylhistidine, beta-hydroxyisovalerylcarnitine, N2,N5-diacetylornithine, 3-methoxytyramine sulfate, guanidinosuccinate, 4-androsten-3beta,17beta-diol monosulfate, 2-palmitoleoylglycerophosphocholine, stearoyl-arachidonoyl-glycerophosphoethanolamine, etiocholanolone glucuronide, 1-linolenoylglycerophosphocholine (18:3n3), palmitoyl-oleoyl-glycerophosphocholine, malonylcarnitine, scyllo-inositol, 2-myristoylglycerol (2-monomyristin), palmitoyl-arachidonoyl-glycerophosphocholine, 1-linolenoylglycerol and O-sulfo-L-tyrosine (71%, 77%, 68%, 46%, 45%, 45%, 47%, 78%, 47%, 47%, 68%, 47%, 47%, 39%, 66%, 47%, 79% and 68%, respectively). <sup>2</sup>Estimate is the standardized beta-coefficient

The metabolite chemical classes of amino acids, lipids, and cofactors/vitamins had strong positive associations with serum retinol ( $p = 1.6 \times 10^{-10}$ ,  $3.3 \times 10^{-7}$  and  $3.3 \times 10^{-7}$ , respectively; Table 3, Supplemental Table 2). In addition, 23 out of 59 metabolite chemical sub-classes exceeded the Bonferroni threshold ( $P < 8.5 \times 10^{-4}$ ), with the top signals being for inositol, creatine, valine/leucine/isoleucine, and purine/urate pathways ( $p = 2.0 \times 10^{-14}$ ,  $1.4 \times 10^{-12}$ ,  $2.8 \times 10^{-11}$ , and  $2.4 \times 10^{-10}$ , respectively; Table 4, Supplemental Table 3).

## Discussion

This large-scale agnostic investigation of nearly 1,300 men identified 63 metabolites in several chemical classes associated with serum retinol concentrations after stringent correction for 947 retinol-metabolite comparisons. Amino acids and lipids were most highly represented among the top metabolites, with N-acetyltryptophan and *myo*-inositol (followed by the lipid 1-palmitoylglycerophosphoethanolamine) ranked first within these two chemical classes, respectively. In addition, N6-carbamoylthreonyladenosine, erythronate, erythritol, adsgedfxaeggvr, and succinylcarnitine were represented as the top metabolites within the nucleotide, carbohydrate, xenobiotic, peptide, and energy metabolite classes.

N-acetyltryptophan prevents oxidative degradation of proteins<sup>33</sup>, and the tryptophan/kynurenine pathway influences the regulation of inflammation, oxidative stress and immune activation<sup>34,35</sup>. The relevance of several other N-acetyl amino acids among the top metabolites (e.g., N-acetyl-3-methylhistidine, -valine, -1-methylhistidine, -lysine, and -threonine) is unclear, but may indicate perturbations in acetylation activity and influence on cell growth through histone-chromatin function and gene regulation<sup>36-38</sup>.

Retinol plays an important role in lipid metabolism through activation of RAR and PPAR $\beta/\delta$  signaling and gene expression enhancing lipid accumulation through control of adipocyte differentiation, fatty acid oxidation, and lipolysis<sup>39-43</sup>. *Myo*-inositol ranked first among lipid metabolites associated with retinol, which along with its phosphorylated derivatives including inositol 1-phosphate (I1P), may be related to mediation of retinoic acid signaling and cellular functions including adhesion, growth, vesicular trafficking, and cell survival<sup>44,45</sup>, and is involved in regulation of enzyme activity and hormone secretion<sup>46</sup>. Through its role in phosphatidylinositol-4-phosphate (IP<sub>4</sub>) and phosphatidylinositol-4,5-bisphosphate (IP<sub>4,5</sub>) biosynthesis, *myo*-inositol and I1P are plasma membrane mediators of protein phosphorylation and second messenger signaling<sup>47</sup>.

Several lysolipids and androgen metabolites were associated with serum retinol. The former glycerophospho-fatty acids are also important plasma membrane cell-signaling molecules<sup>48,49</sup> that influence cell differentiation, growth, proliferation, and invasion<sup>50-55</sup>. The four serum androstenediol (adiol) sulfate metabolites

Chemical Class	No. of contributing metabolites	No. of contributing metabolites after Bonferroni correction	p value
Amino acids	157	20	$1.57 \times 10^{-10}$
Lipids	264	30	$3.28 \times 10^{-7}$
Cofactors and vitamins	28	0	$3.32 \times 10^{-7}$
Energy	9	1	$1.82 \times 10^{-5}$
Peptides	91	1	$8.58 \times 10^{-5}$
Carbohydrates	26	2	$9.22 \times 10^{-5}$
Nucleotides	31	4	0.00011
Xenobiotics	124	4	0.04470

**Table 3.** Gene set analysis for chemical class of metabolites for serum retinol

positively related to serum retinol could serve as androgenic mediators and confer the vitamin A-prostate cancer association<sup>56</sup>.

The present analysis also shows retinol is correlated with histidine pathway metabolites, including N-acetyl-3-methylhistidine, N-acetyl-1-methylhistidine, 1-methylhistidine and 3-methylhistidine. Experimental data indicate that biosynthesis of histidine metabolites may be mediated through the 5-phosphoribosyl-1-pyrophosphate signaling pathway<sup>57</sup>. 1- and 3-Methylhistidine are abundant in actin, formed by post-translational methylation<sup>58,59</sup>, and are considered markers of myofibrillar protein degradation<sup>60</sup>. Excessive vitamin A intake may stimulate 3-methylhistidine excretion both *in vivo* and *in vitro*, suggesting accelerated myofibrillar protein breakdown and protein turnover<sup>61</sup>. Aside from myofibrillar protein degradation, aberrations in histidine metabolites may also be indicative of tissue repair, tissue damage, alterations in the inflammatory signals and oxidative stress<sup>62–65</sup>.

This cross-sectional analysis was conducted in order to both contribute to elucidation of the molecular basis of retinol's diverse biological activities and provide mechanistic clues related to how higher vitamin A status might adversely impact prostate (and other) cancer risk. In this regard, the present findings regarding inositol and lysolipid metabolites are of interest in that they (e.g., I1P) were associated with aggressive prostate cancer in our recent serum metabolomic analysis<sup>66</sup>. In addition, findings from our group based on size and extent of primary prostate cancer cases demonstrated a positive association with histidine metabolites including N-acetyl-3-methylhistidine, and 1- and 3-methylhistidine in T2 prostate cancers<sup>67</sup>. Some of the sex steroids related to serum retinol here were associated with non-aggressive prostate cancer the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial<sup>68</sup>.

Fibrinopeptide A is the only metabolite we found to be inversely associated with serum retinol in this study. Fibrinopeptide A is produced from fibrinogen by thrombin during blood coagulation, and it is elevated in several malignancies, with persistent coagulation activation, thrombosis and disseminated intravascular coagulation being common complications<sup>69–72</sup>. Given that some cancers have been related to vitamin A status, the inverse fibrinopeptide A-retinol association observed here should be examined in other studies.

To our knowledge, this is the first metabolomic profiling analysis of vitamin A status. Strengths of the analysis include its relatively large sample size, assaying of serum collected after an overnight fast, and high laboratory quality control. Some limitations should also be acknowledged, including that our study was cross-sectional, which limits our ability to make causal inferences about the association between metabolites and retinol levels. In addition, the Alpha-Tocopherol, Beta-Carotene Cancer Prevention (ATBC) cohort constitutes a relatively homogeneous population of Finnish male smokers of European ancestry, which may limit the generalizability of our findings to other populations. Residual confounding by unidentified metabolic or other factors such as chronic health conditions is possible, although we extensively adjusted our models for several potential confounders including diabetes, body mass index (BMI), serum cholesterol, smoking intensity, and alcohol consumption, and the findings were unchanged after adjusting for serum creatinine, an important indicator of kidney function.

Our study identified a large number of metabolites in various metabolic pathways associated with serum retinol concentration, some of which directly relate to prostate cancer risk associations and other known biological functions. Lipid and amino acid biochemical classes are heavily represented, which may provide some molecular insights into the role of retinol in human health and disease. In particular, lipid bioactivity related to inositol, lysolipids and sterol/steroid compounds, may potentially provide biological clues relevant to the positive association between circulating retinol and the development of prostate cancer. Further studies in diverse populations are warranted to re-examine these findings.

## Materials and methods

**Study population.** The ATBC Study has been described in detail elsewhere<sup>73</sup>. In brief, it is a 2 × 2 factorial randomized controlled trial in which 29,133 male Finnish smokers aged 50–69 were assigned to receive either alpha-tocopherol (dl- $\alpha$ -tocopheryl-acetate, 50 mg/day), beta-carotene (20 mg/day), both vitamins, or placebo. Participants received the study capsules for a median of 6.1 years (range 5–8 years) until the trial ended on 30 April 1993. Information on general risk factors, smoking, and medical history was obtained through a questionnaire at study entry along with a validated food-frequency questionnaire. Serum was collected at baseline (i.e., pre-supplementation and years prior to cancer diagnoses) after overnight fast was protected from light and stored at  $-70^{\circ}\text{C}$  until assayed. Incident cancer diagnosed during 20 years of follow-up were identified from the Finnish Cancer Registry. All participants provided written informed consent. The trial was approved by institutional

Sub-Pathway	No. of contributing metabolites	p value
Inositol metabolism	4	$1.98 \times 10^{-14}$
Creatine metabolism	3	$1.41 \times 10^{-12}$
Valine, leucine and isoleucine metabolism	27	$2.75 \times 10^{-11}$
Purine metabolism, urate metabolism; purine metabolism, (Hypo)xanthine/inosine containing	3	$2.43 \times 10^{-10}$
Lysolipid	62	$1.02 \times 10^{-8}$
Urea cycle; arginine and proline metabolism	14	$1.76 \times 10^{-8}$
Monoacylglycerol	13	$2.24 \times 10^{-8}$
Histidine metabolism	12	$3.63 \times 10^{-8}$
Tryptophan metabolism	18	$1.60 \times 10^{-7}$
Sterol/steroid	46	$1.76 \times 10^{-7}$
Lysine metabolism	9	$1.99 \times 10^{-6}$
Aminosugar metabolism	2	$7.38 \times 10^{-6}$
Fatty acid metabolism (also BCAA metabolism)	5	$9.86 \times 10^{-6}$
Polyamine metabolism	3	$1.52 \times 10^{-5}$
Fibrinogen cleavage peptide	2	$3.79 \times 10^{-5}$
Pentose metabolism	10	$6.39 \times 10^{-5}$
Ascorbate and aldarate metabolism	3	0.000126
Alanine and aspartate metabolism; pyrimidine metabolism, uracil containing	3	0.000238
Glutamate metabolism	6	0.000250
Pyrimidine metabolism, uracil containing	6	0.000295
Cysteine, methionine, SAM, taurine metabolism	15	0.000558
Glutathione metabolism	2	0.000630
Essential fatty acid; polyunsaturated fatty acid (N3 and N6)	7	0.000681

**Table 4.** Gene set analysis for chemical sub-class of metabolites for serum retinol at  $8.5 \times 10^{-4}$  level of statistical significance.

review boards at the U.S. National Cancer Institute and the Finnish National Public Health Institute. All methods were performed in accordance with the relevant guidelines and regulations.

The present analysis is based on participants included in seven on-going or completed analyses (case-control and other, subsequently referred to as “metabolomic sets”) nested within the ATBC Study<sup>66,74,75</sup>. These included cases of prostate, lung, esophagus, and stomach cancer, along with their controls, as well as controls from a pancreatic cancer study. After excluding duplicate participants, 1,282 were included in the present analysis.

**Laboratory analysis.** Baseline serum retinol concentration was measured using an isocratic high-performance liquid chromatography platform<sup>76</sup>. Assays were carried out from 1986–88 in a dedicated laboratory at the National Public Health Institute in Helsinki, Finland, which was certified by a National Institute of Standards and Technology quality-control testing program. Serum samples were subsequently analyzed at Metabolon, Inc. (Durham, N.C.) using ultrahigh performance liquid chromatography (LC)/mass spectroscopy (MS) and gas chromatography (GC)/mass spectrometry (MS) as described before<sup>66,74,75</sup>. The methods of sample accessioning, sample preparation, quality control, data extraction and compound identification have also been described in detail<sup>77</sup>. Briefly, each of the samples were analyzed using three different analytical conditions: ultrahigh performance LC-MS/MS (+ electrospray ionization (ESI)), ultrahigh performance LC-MS/MS (-ESI), and GC-MS. To remove proteins, dissociate small molecules bound to protein or trapped in the precipitated protein matrix, and to recover chemically diverse metabolites, each of the samples were extracted using an aqueous methanol extraction procedure. The methanol contained four recovery standards (DL-2-fluorophenylglycine, tridecanoic acid, d6-cholesterol and 4-chlorophenylalanine) to allow confirmation of extraction efficiency. For each sample, four aliquots were obtained from the extract and dried. Two aliquots of each sample were then reconstituted in 50  $\mu$ l of 6.5 mM ammonium bicarbonate in water (pH 8) for the negative ion analysis, and the other two aliquots of each were reconstituted using 50  $\mu$ l 0.1% formic acid in water (pH ~3.5) for the positive ion analysis. The resulting extracts were divided into fractions for analysis by ultrahigh performance LC/MS/MS (positive mode), ultrahigh performance LC/MS/MS (negative mode), and GC/MS. Samples were placed briefly on a TurboVap<sup>®</sup> (Zymark) to remove the organic solvent. Each sample was then frozen and dried under vacuum. The samples were then prepared for the appropriate instrument, either ultrahigh performance LC/MS/MS or GC/MS. The raw data was further extracted, peak-identified and quality control (QC) processed using Metabolon’s hardware and software. Internal controls included extraction process (5 recovery standards), injection (up to 11 standards), and alignment standards for quality assurance/quality control procedures to control for experimental variability. Compounds were identified by comparison to library entries of purified standards or recurrent unknown entities. Metabolon maintains a library based on authenticated standards that contains the retention time/index (RI), mass to charge ratio (m/z), and chromatographic data (including MS/MS spectral data) on all molecules present in the library. Furthermore, biochemical identifications are based on three



criteria: retention index within a narrow RI window of the proposed identification, accurate mass match to the library  $\pm 0.005$  amu, and the MS/MS forward and reverse scores between the experimental data and authentic standards. The MS/MS scores are based on a comparison of the ions present in the experimental spectrum to the ions present in the library spectrum. More than 2400 commercially available purified standard compounds have been acquired and used in both the LC and GC platforms for determination of their analytical characteristics<sup>78</sup>.

Batch variability was standardized by dividing the signal strength by the batch median value for each metabolite and subject. Metabolite signal strength were log-transformed and centered according to normalization. Within each metabolomic set, metabolite values below the limit of detection were imputed to have the minimum of all non-missing values.

After excluding metabolites that had fewer than 10 non-missing values across all metabolomic sets, or single-value metabolites in individual metabolomic set, 947 identified compounds remained eligible for analysis. Based upon a standard chemical classification scheme, 920 of the metabolites were categorized in eight mutually exclusive chemical classes: amino acids and amino acid derivatives (referred to as “amino acids”), carbohydrates, cofactors and vitamins, energy metabolites, lipids, nucleotides, peptides, and xenobiotics. Blinded quality control samples (9%) from a pooled sample were included in each batch to assess the technical reliability of data, and coefficient of variation (CV) was calculated. The median and interquartile range of the CV% across the metabolites was 9% (4%–20%). These CVs are similar to those previously observed for blood samples analyzed by the same laboratory<sup>77,78</sup>, and previous studies have reported a high reliability for the metabolomics platform used in the present study<sup>79</sup>.

**Statistical analysis.** We investigated the association between serum levels of retinol and each metabolite, using linear regression adjusted for age at randomization (continuous), BMI (continuous), case status (binary, case or control status in each metabolomic set), metabolomic sets (categorical), serum cholesterol (continuous), number of cigarettes per days (continuous), and baseline alcohol consumption (continuous). We further performed additional analysis to adjust for creatinine in the model. For each model, we computed the standardized beta-coefficient which represents the change in the level of each metabolite per 1-SD increase in the levels of retinol.

For all analyses, we used log-transformed metabolite concentrations. To control for multiple comparisons ( $n = 947$  retinol-metabolite associations), we applied the Bonferroni correction<sup>80</sup> using  $p = 0.05/947 = 5.3 \times 10^{-5}$  as the adjusted significance threshold. We further performed sensitivity analyses additionally adjusted for storage time (from blood collection to metabolomics measurement).

Within each metabolomic set, we applied Gene-Set Analysis (GSA), which is a pathway analysis method, to examine whether pre-defined metabolic pathways including super- and sub-pathways were related to retinol within individual metabolomic sets<sup>81</sup>. For  $Z$  values (i.e.  $Z_1$  to  $Z_s$ ) tested from  $S$  metabolites in a pre-defined pathway, GSA examines the “maxmean” statistic max that is the average of all  $Z$  values (both positive and negative) and calculates the  $p$  values by 10,000 permutations. With the seven individual metabolomic sets, we used Fisher’s method, namely sum of logs method, to combine  $p$  value for each pre-defined pathway.

Analyses were performed in SAS 9.4, and R 3.2.3. All statistical tests and reported  $p$  values are two-sided.

## References

- Sommer, A. & Vyas, K. S. A global clinical view on vitamin A and carotenoids. *Am J Clin Nutr* **96**, 1204S–1206S, doi:10.3945/ajcn.112.034868 (2012).
- Bates, C. J. Vitamin A. *Lancet* **345**, 31–35 (1995).
- Sommer, A. & W. K. Vitamin A deficiency: health, survival, and vision. New York: Oxford University Press. (1996).
- D’Ambrosio, D. N., Clugston, R. D. & Blaner, W. S. Vitamin A metabolism: an update. *Nutrients* **3**, 63–103, doi:10.3390/nu3010063 (2011).
- Sporn, M. B., R. A., Goodman, D. S., editors. The retinoids: biology, chemistry, and medicine. New York: Raven Press. (1994).
- Dadon Bar-El, S. & Reif, R. Vitamin A and the Epigenome. *Critical Reviews in Food Science and Nutrition*, 00–00, doi:10.1080/10408398.2015.1060940 (2015).
- Underwood, B. A. Vitamin A deficiency disorders: international efforts to control a preventable “pox”. *J Nutr* **134**, 231S–236S (2004).
- Sommer, A. Vitamin A deficiency and clinical disease: an historical overview. *J Nutr* **138**, 1835–1839 (2008).
- Willis, M. S. & Wians, F. H. The role of nutrition in preventing prostate cancer: a review of the proposed mechanism of action of various dietary substances. *Clin Chim Acta* **330**, 57–83 (2003).
- Vitamin, A. retinol, carotene, and cancer prevention. *Br Med J* **281**, 957–958 (1980).
- Hunter, D. J. *et al.* A prospective study of the intake of vitamins C, E, and A and the risk of breast cancer. *N Engl J Med* **329**, 234–240, doi:10.1056/NEJM199307223290403 (1993).
- Krinsky, N. I. & Johnson, E. J. Carotenoid actions and their relation to health and disease. *Mol Aspects Med* **26**, 459–516, doi:10.1016/j.mam.2005.10.001 (2005).
- Peehl, D. M. & Feldman, D. The role of vitamin D and retinoids in controlling prostate cancer progression. *Endocr Relat Cancer* **10**, 131–140 (2003).
- Gualler, E., Stranges, S., Mulrow, C., Appel, L. J. & Miller, E. R. 3rd Enough is enough: Stop wasting money on vitamin and mineral supplements. *Ann Intern Med* **159**, 850–851, doi:10.7326/0003-4819-159-12-201312170-00011 (2013).
- Taylor, P. R. & Greenwald, P. Nutritional interventions in cancer prevention. *J Clin Oncol* **23**, 333–345, doi:10.1200/JCO.2005.06.190 (2005).
- Steinmetz, K. A. & Potter, J. D. Vegetables, fruit, and cancer. I. Epidemiology. *Cancer Causes Control* **2**, 325–357 (1991).
- Gaziano, J. M. *et al.* A prospective study of consumption of carotenoids in fruits and vegetables and decreased cardiovascular mortality in the elderly. *Ann Epidemiol* **5**, 255–260 (1995).
- Manson, J. E., Gaziano, J. M., Jonas, M. A. & Hennekens, C. H. Antioxidants and cardiovascular disease: a review. *J Am Coll Nutr* **12**, 426–432 (1993).
- Brazionis, L., Walker, K. Z., Itsiopoulos, C. & O’Dea, K. Plasma retinol: a novel marker for cardiovascular disease mortality in Australian adults. *Nutr Metab Cardiovasc Dis* **22**, 914–920, doi:10.1016/j.numecd.2011.08.009 (2012).

20. Fortmann, S. P., Burda, B. U., Senger, C. A., Lin, J. S. & Whitlock, E. P. Vitamin and mineral supplements in the primary prevention of cardiovascular disease and cancer: An updated systematic evidence review for the U.S. Preventive Services Task Force. *Ann Intern Med* **159**, 824–834, doi:10.7326/0003-4819-159-12-201312170-00729 (2013).
21. Ioannidis, J. P. Contradicted and initially stronger effects in highly cited clinical research. *JAMA* **294**, 218–228, doi:10.1001/jama.294.2.218 (2005).
22. Mondul, A. M. *et al.* Serum retinol and risk of prostate cancer. *Am J Epidemiol* **173**, 813–821, doi:10.1093/aje/kwq429 (2011).
23. Key, T. J. *et al.* Carotenoids, retinol, tocopherols, and prostate cancer risk: pooled analysis of 15 studies. *Am J Clin Nutr* **102**, 1142–1157, doi:10.3945/ajcn.115.114306 (2015).
24. Omenn, G. S. *et al.* Effects of a combination of beta carotene and vitamin A on lung cancer and cardiovascular disease. *N Engl J Med* **334**, 1150–1155, doi:10.1056/NEJM199605023341802 (1996).
25. Sesso, H. D., Buring, J. E., Norkus, E. P. & Gaziano, J. M. Plasma lycopene, other carotenoids, and retinol and the risk of cardiovascular disease in men. *Am J Clin Nutr* **81**, 990–997 (2005).
26. Albanes, D. *et al.* Alpha-Tocopherol and beta-carotene supplements and lung cancer incidence in the alpha-tocopherol, beta-carotene cancer prevention study: effects of base-line characteristics and study compliance. *J Natl Cancer Inst* **88**, 1560–1570 (1996).
27. U.S. Preventive Services Task Force. Internet: <http://www.uspreventiveservicestaskforce.org/Page/Document/RecommendationStatementFinal/vitamin-supplementation-to-prevent-cancer-and-cvd-counseling>.
28. Wright, M. E. *et al.* Effects of beta-carotene supplementation on molecular markers of lung carcinogenesis in male smokers. *Cancer Prev Res (Phila)* **3**, 745–752, doi:10.1158/1940-6207.CAPR-09-0107 (2010).
29. Liu, C., Russell, R. M. & Wang, X. D. Exposing ferrets to cigarette smoke and a pharmacological dose of beta-carotene supplementation enhance *in vitro* retinoic acid catabolism in lungs via induction of cytochrome P450 enzymes. *J Nutr* **133**, 173–179 (2003).
30. Goodman, G. E., Alberts, D. S. & Meyskens, F. L. Retinol, vitamins, and cancer prevention: 25 years of learning and relearning. *J Clin Oncol* **26**, 5495–5496, doi:10.1200/JCO.2008.19.0884 (2008).
31. Serum retinol concentrations for determining the prevalence of vitamin A deficiency in populations. Vitamin and mineral nutrition information system. *Geneva, Switzerland: World Health Organization* (2011).
32. Underwood, B. A. Hypovitaminosis A: international programmatic issues. *J Nutr* **124**, 1467S–1472S (1994).
33. Fang, L., Parti, R. & Hu, P. Characterization of N-acetyltryptophan degradation products in concentrated human serum albumin solutions and development of an automated high performance liquid chromatography-mass spectrometry method for their quantitation. *J Chromatogr A* **1218**, 7316–7324, doi:10.1016/j.chroma.2011.08.044 (2011).
34. Wang, Q., Liu, D., Song, P. & Zou, M. H. Tryptophan-kynurenine pathway is dysregulated in inflammation, and immune activation. *Front Biosci (Landmark Ed)* **20**, 1116–1143 (2015).
35. Santhanam, S., Alvarado, D. M. & Ciorba, M. A. Therapeutic targeting of inflammation and tryptophan metabolism in colon and gastrointestinal cancer. *Transl Res* **167**, 67–79, doi:10.1016/j.trsl.2015.07.003 (2016).
36. Aksnes, H., Hole, K. & Arnesen, T. Molecular, cellular, and physiological significance of N-terminal acetylation. *Int Rev Cell Mol Biol* **316**, 267–305, doi:10.1016/bs.ircmb.2015.01.001 (2015).
37. Eberharter, A. & Becker, P. B. Histone acetylation: a switch between repressive and permissive chromatin. *Second in review series on chromatin dynamics. EMBO Rep* **3**, 224–229, doi:10.1093/embo-reports/kvf053 (2002).
38. Verdone, L., Caserta, M. & Di Mauro, E. Role of histone acetylation in the control of gene expression. *Biochem Cell Biol* **83**, 344–353, doi:10.1139/o05-041 (2005).
39. Berry, D. C., Soltanian, H. & Noy, N. Repression of cellular retinoic acid-binding protein II during adipocyte differentiation. *J Biol Chem* **285**, 15324–15332, doi:10.1074/jbc.M110.110635 (2010).
40. Wang, Y. X. *et al.* Peroxisome-proliferator-activated receptor delta activates fat metabolism to prevent obesity. *Cell* **113**, 159–170 (2003).
41. Schwarz, E. J., Reginato, M. J., Shao, D., Krakow, S. L. & Lazar, M. A. Retinoic acid blocks adipogenesis by inhibiting C/EBPbeta-mediated transcription. *Mol Cell Biol* **17**, 1552–1561 (1997).
42. Neele, D. M., de Wit, E. C. & Princen, H. M. Inhibition of apolipoprotein(a) synthesis in cynomolgus monkey hepatocytes by retinoids via involvement of the retinoic acid receptor. *Biochem Pharmacol* **58**, 263–271 (1999).
43. Berry, D. C. & Noy, N. Signaling by vitamin A and retinol-binding protein in regulation of insulin responses and lipid homeostasis. *Biochim Biophys Acta* **1821**, 168–176, doi:10.1016/j.bbali.2011.07.002 (2012).
44. Toker, A. & Cantley, L. C. Signalling through the lipid products of phosphoinositide-3-OH kinase. *Nature* **387**, 673–676, doi:10.1038/42648 (1997).
45. Laserna, E. J. *et al.* Proteomic analysis of phosphorylated nuclear proteins underscores novel roles for rapid actions of retinoic acid in the regulation of mRNA splicing and translation. *Mol Endocrinol* **23**, 1799–1814, doi:10.1210/me.2009-0165 (2009).
46. Krapels, I. P. *et al.* Myo-inositol, glucose and zinc status as risk factors for non-syndromic cleft lip with or without cleft palate in offspring: a case-control study. *BJOG* **111**, 661–668, doi:10.1111/j.1471-0528.2004.00171.x (2004).
47. Tsui, M. M. & York, J. D. Roles of inositol phosphates and inositol pyrophosphates in development, cell signaling and nuclear processes. *Adv Enzyme Regul* **50**, 324–337, doi:10.1016/j.advenzreg.2009.12.002 (2010).
48. Lisacovitch, M. & Cantley, L. C. Lipid second messengers. *Cell* **77**, 329–334 (1994).
49. Moolenaar, W. H. Bioactive lysophospholipids and their G protein-coupled receptors. *Exp Cell Res* **253**, 230–238, doi:10.1006/excr.1999.4702 (1999).
50. Xu, Y., Xiao, Y. J., Baudhuin, L. M. & Schwartz, B. M. The role and clinical applications of bioactive lysolipids in ovarian cancer. *J Soc Gynecol Investig* **8**, 1–13 (2001).
51. Goetzl, E. J., Dolezalova, H., Kong, Y. & Zeng, L. Dual mechanisms for lysophospholipid induction of proliferation of human breast carcinoma cells. *Cancer Res* **59**, 4732–4737 (1999).
52. Manning, T. J., Jr., Parker, J. C. & Sontheimer, H. Role of lysophosphatidic acid and rho in glioma cell motility. *Cell Motil Cytoskeleton* **45**, 185–199, doi:10.1002/(SICI)1097-0169(200003)45:3<185::AID-CM2>3.0.CO;2-G (2000).
53. Mukai, M. *et al.* Inhibition of tumor invasion and metastasis by a novel lysophosphatidic acid (cyclic LPA). *Int J Cancer* **81**, 918–922 (1999).
54. Mills, G. B. & Moolenaar, W. H. The emerging role of lysophosphatidic acid in cancer. *Nat Rev Cancer* **3**, 582–591, doi:10.1038/nrc1143 (2003).
55. Liu, S. *et al.* Expression of autotaxin and lysophosphatidic acid receptors increases mammary tumorigenesis, invasion, and metastases. *Cancer Cell* **15**, 539–550, doi:10.1016/j.ccr.2009.03.027 (2009).
56. Mizokami, A. *et al.* The adrenal androgen androstenediol is present in prostate cancer tissue after androgen deprivation therapy and activates mutated androgen receptor. *Cancer Res* **64**, 765–771 (2004).
57. Zhang, T. *et al.* Identification of potential biomarkers for ovarian cancer by urinary metabolomic profiling. *J Proteome Res* **12**, 505–512, doi:10.1021/pr3009572 (2013).
58. Asatoor, A. M. & Armstrong, M. D. 3-methylhistidine, a component of actin. *Biochem Biophys Res Commun* **26**, 168–174 (1967).
59. Young, V. R., Alexis, S. D., Baliga, B. S., Munro, H. N. & Muecke, W. Metabolism of administered 3-methylhistidine. *Lack of muscle transfer ribonucleic acid charging and quantitative excretion as 3-methylhistidine and its N-acetyl derivative.* *J Biol Chem* **247**, 3592–3600 (1972).



60. Haverberg, L. N., Munro, H. N. & Young, V. R. Isolation and quantitation of Ntau-methylhistidine in actin and myosin of rat skeletal muscle: use of pyridine elution of protein hydrolysates on ion-exchange resins. *Biochim Biophys Acta* **371**, 226–237 (1974).
61. Hillgartner, F. B., Morin, D. & Hansen, R. J. Effect of excessive vitamin A intake on muscle protein turnover in the rat. *Biochem J* **202**, 499–508 (1982).
62. Kumar, A. *et al.* Metabolomic analysis of serum by (1) H NMR spectroscopy in amyotrophic lateral sclerosis. *Clin Chim Acta* **411**, 563–567, doi:10.1016/j.cca.2010.01.016 (2010).
63. Nicholson, J. K., Lindon, J. C. & Holmes, E. ‘Metabonomics’: understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica* **29**, 1181–1189, doi:10.1080/004982599238047 (1999).
64. Sui, W. *et al.* A proton nuclear magnetic resonance-based metabonomics study of metabolic profiling in immunoglobulin a nephropathy. *Clinics (Sao Paulo)* **67**, 363–373 (2012).
65. Watanabe, M. *et al.* Consequences of low plasma histidine in chronic kidney disease patients: associations with inflammation, oxidative stress, and mortality. *Am J Clin Nutr* **87**, 1860–1866 (2008).
66. Mondul, A. M. *et al.* Metabolomic analysis of prostate cancer risk in a prospective cohort: The alpha-tocopherol, beta-carotene cancer prevention (ATBC) study. *Int J Cancer* **137**, 2124–2132, doi:10.1002/ijc.29576 (2015).
67. Huang, J. *et al.* Prospective Serum Metabolomic Profile of Prostate Cancer by Size and Extent of Primary Tumor. *Oncotarget* **8**, 45190–45199, doi:10.18632/oncotarget.16775 (2017).
68. Huang, J. *et al.* Serum metabolomic profiling of prostate cancer risk in the prostate, lung, colorectal, and ovarian cancer screening trial. *Br J Cancer*. doi:10.1038/bjc.2016.305 (2016).
69. Bergen, H. R. 3rd *et al.* Discovery of ovarian cancer biomarkers in serum using NanoLC electrospray ionization TOF and FT-ICR mass spectrometry. *Dis Markers* **19**, 239–249 (2003).
70. Orvisky, E. *et al.* Enrichment of low molecular weight fraction of serum for MS analysis of peptides associated with hepatocellular carcinoma. *Proteomics* **6**, 2895–2902, doi:10.1002/pmic.200500443 (2006).
71. Abbasciano, V. *et al.* Usefulness of coagulation markers in staging of gastric cancer. *Cancer Detect Prev* **19**, 331–336 (1995).
72. Villanueva, J. *et al.* Serum peptidome patterns that distinguish metastatic thyroid carcinoma from cancer-free controls are unbiased by gender and age. *Mol Cell Proteomics* **5**, 1840–1852, doi:10.1074/mcp.M600229-MCP200 (2006).
73. The ATBC Cancer Prevention Study Group. The alpha-tocopherol, beta-carotene lung cancer prevention study: design, methods, participant characteristics, and compliance. *Ann Epidemiol* **4**, 1–10 (1994).
74. Mondul, A. M. *et al.* Metabolomic profile of response to supplementation with beta-carotene in the Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study. *Am J Clin Nutr* **98**, 488–493, doi:10.3945/ajcn.113.062778 (2013).
75. Mondul, A. M. *et al.* 1-stearoylglycerol is associated with risk of prostate cancer: results from serum metabolomic profiling. *Metabolomics* **10**, 1036–1041, doi:10.1007/s11306-014-0643-0 (2014).
76. Milne, D. B. & Botnen, J. Retinol, alpha-tocopherol, lycopene, and alpha- and beta-carotene simultaneously determined in plasma by isocratic liquid chromatography. *Clin Chem* **32**, 874–876 (1986).
77. Evans, A. M., DeHaven, C. D., Barrett, T., Mitchell, M. & Milgram, E. Integrated, nontargeted ultrahigh performance liquid chromatography/electrospray ionization tandem mass spectrometry platform for the identification and relative quantification of the small-molecule complement of biological systems. *Anal Chem* **81**, 6656–6667, doi:10.1021/ac901536h (2009).
78. DeHaven, C. D., Evans, A. M., Dai, H. & Lawton, K. A. Organization of GC/MS and LC/MS metabolomics data into chemical libraries. *J Cheminform* **2**, 9, doi:10.1186/1758-2946-2-9 (2010).
79. Sampson, J. N. *et al.* Metabolomics in epidemiology: sources of variability in metabolite measurements and implications. *Cancer Epidemiol Biomarkers Prev* **22**, 631–640, doi:10.1158/1055-9965.EPI-12-1109 (2013).
80. Bland, J. M. & Altman, D. G. Multiple significance tests: the Bonferroni method. *BMJ* **310**, 170 (1995).
81. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* **102**, 15545–15550, doi:10.1073/pnas.0506580102 (2005).

## Acknowledgements

The ATBC Study is supported by the Intramural Research Program of the U.S. National Cancer Institute, National Institutes of Health, and by U.S. Public Health Service contract HHSN261201500005C from the National Cancer Institute, Department of Health and Human Services.

## Author Contributions

The authors’ responsibilities were as follow—J.H., O.A.P., G.M.A., A.M.M., L.M.L., A.D., R.S.S., S.J.W., D.A.: design the study; L.M.L., R.S.S., S.J.W., D.A.: provide essential materials; J.H., O.A.P.: perform the statistical analysis; O.A.P., A.M.M., A.D., S.J.W., D.A.: advised on statistical analysis and interpretation of the findings; J.H., O.A.P.: draft the manuscript; J.H., O.A.P., G.M.A., A.M.M., L.M.L., A.D., R.S.S., S.J.W., D.A.: revised the manuscript; D.A.: had primary responsibility for the final content; and all authors: read and approved the final manuscript.

## Additional Information

**Supplementary information** accompanies this paper at doi:10.1038/s41598-017-09698-w

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher’s note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**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 license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017