Characterization of the interindividual variability of lutein and zeaxanthin concentrations in the adipose tissue of healthy males and identification of combinations of genetic variants associated with it

Mark Pretzel Zumaraga, Charles Desmarchelier, Beatrice Gleize, Marion Nowicki, Djaffar Ould-Ali, and Patrick Borel

Supporting Tables, Figures and Methods

Supplemental Table S1. Candidate genes selected.

Gene Name	Gene symbol	References
Genes that play, or are assumed to play, a role in Xanth or triglyceride metabolism in adipose tissue		
Apolipoprotein C1	*APOC1*	*1*
Apolipoprotein C2	*APOC2*	*1*
Apolipoprotein C4	*APOC4*	*1*
Apolipoprotein E	*APOE*	*1*
β-carotene 15,15'oxygenase-1	*BCO1*	*2*
β-carotene 9,10' oxygenase-2	*BCO2*	*3*
Cholesterol ester transfer protein	*CETP*	*4*
Cluster of differentiation 36	*CD36*	*5*
ELOVL fatty acid elongase 5	*ELOVL5*	*6*
GRAM domain containing 1A	*GRAMD1A*	*7,8*
GRAM domain containing 1C	*GRAMD1C*	*7,8*
Lecithin-cholesterol acyltransferase	*LCAT*	*4*
Low density lipoprotein receptor	*LDLR*	*9*
Lipase E	*LIPE*	*10*
Lipoprotein lipase	*LPL*	*11*

Monoglyceride lipase	*MGLL*	*12,13*
Polycystic kidney disease 1-like 2	*PKD1L2*	*2*
Phospholipid transfer protein	*PLTP*	*4*
Patatin-like phospholipase domain-containing 2	*PNPLA2*	*10*
Peroxisome proliferator activated receptor gamma	*PPARG*	*14*
Scavenger receptor class B member 1	*SCARB1*	*15*

Genes that have been associated with serum XANTH concentration in genome wide association studies

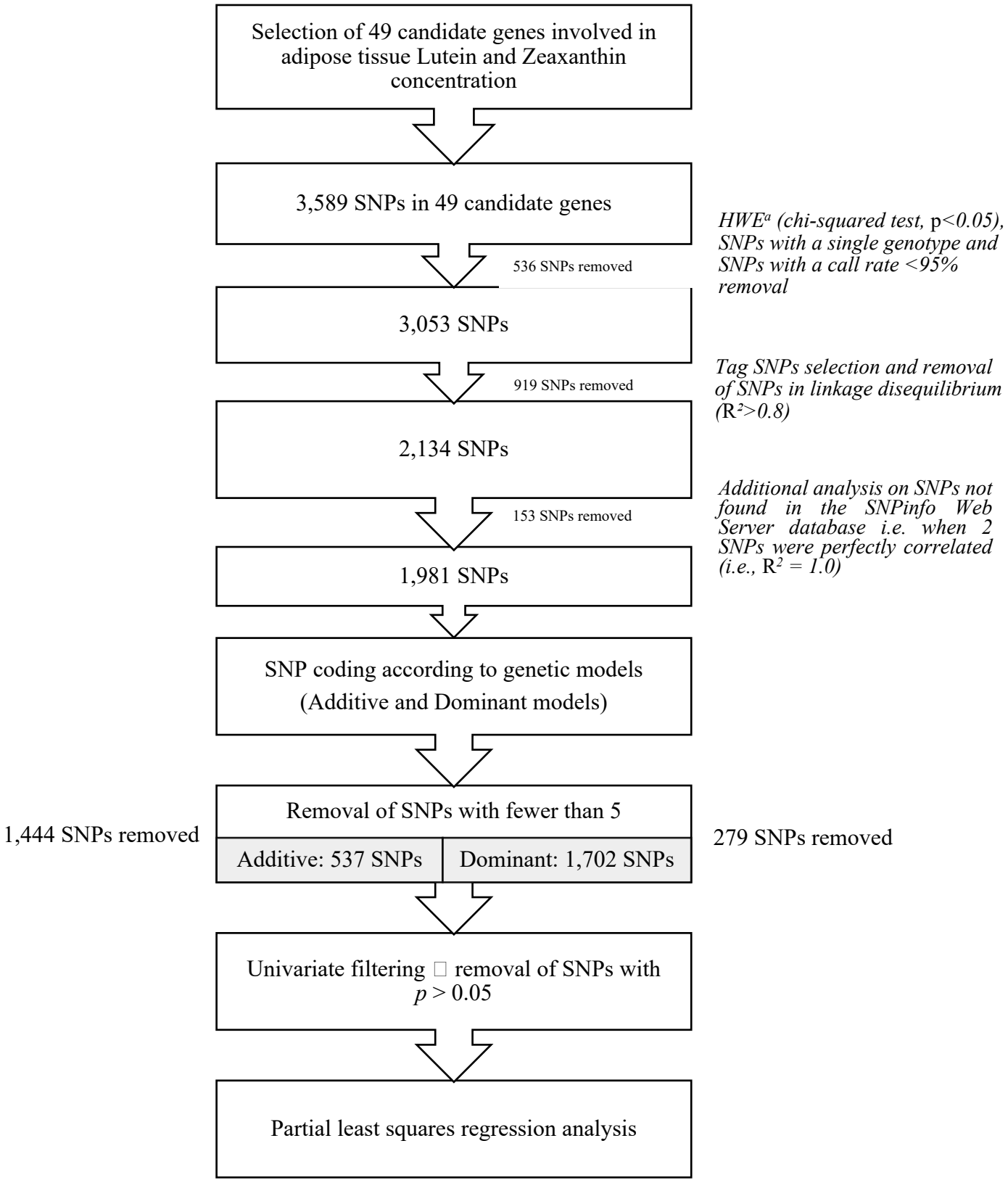
Nuclear receptor subfamily 1 group H member 3	*NR1H3*	16
StAR related lipid transfer domain containing	*STARD3*	17

Genes whose SNPs have been associated with the postprandial chylomicron carotenoid or triacylglycerol response in the same group of participants

ATP binding cassette subfamily G member 2	*ABCG2*	18,19
ATP binding cassette subfamily A member 1	*ABCA1*	18–21
ATP binding cassette subfamily G member 5	*ABCG5*	20
Apolipoprotein A1	*APOA1*	19
Apolipoprotein A3	*APOA3*	19
Apolipoprotein A4	*APOA4*	19
Apolipoprotein A5	*APOA5*	19,21
Apolipoprotein B	*APOB*	18–20
COBL-like 1	*COBLL1*	19
Chemokine (C-X-C motif) ligand 8	*CXCL8*	20
ELOVL fatty acid elongase 2	*ELOVL2*	18–20
Fatty acid desaturase 1	*FADS1*	22

Fatty acid desaturase 2	*FADS2*	[22]
Fatty acid desaturase 3	*FADS3*	[22]
Insulin induced gene 2	*INSIG2*	[18,19]
Insulin receptor substrate 1	*IRS1*	[19]
Intestine specific homebox	*ISX*	[18–20,22]
Lipase, hepatic	*LIPC*	[18–20]
Melanocortin 4 receptor	*MC4R*	[19]
Microsomal triglyceride transfer protein	*MTTP*	[18,19]
Niemann-Pick disease, type C1, gene-like 1	*NPC1L1*	[18]
Pancreatic lipase	*PNLIP*	[18]
Retinal pigment epithelium-specific protein 65kDa	*RPE65*	[19,20]
Solute carrier family 27, member 6	*SLC27A6*	[18]
Superoxide dismutase 2, mitochondrial	*SOD2*	[18,20]
Transcription factor 7 like 2	*TCF7L2*	[20]

Supplemental Figure S1. Candidate SNPs selection flowchart.



[a]Deviations from Hardy-Weinberg equilibrium (HWE) can indicate inbreeding, population stratification, and genotyping errors.

Supplemental Table S3A-B. Comparison between adipose tissue lutein (L) and zeaxanthin (Z) concentrations measured at fast and 8 h after consumption of the 3 test meals.

A. Linear mixed model

Parameters[a]	Adipose tissue L concentration		Adipose tissue Z concentration	
	F	**Sig.**	**F**	**Sig.**
Intercept	134.8	0.000	140.8	0.000
Time (Fasting vs 8 hr)	0.5	0.494	0.6	0.458
Type of Meal (Control vs Vitamin E vs Tomato Puree)	1.1	0.342	0.5	0.604
Time * Type of Meal	2.2	0.124	2.4	0.103

[a]Unstructured linear mixed model. Adipose tissue L and Z concentrations measured at fast and 8 h after consumption of the 3 test meals were analyzed with linear mixed models, using a full factorial design with meal (control, vitamin E and tomato puree) and time (fasting and 8 h post-meal) as fixed within-subject variables and participant as the random variable. Of the 5 linear mixed models tested, the unstructured model was selected based on Akaike's Information Criterion [23]

B. Paired *t*-test

Type of Meal	Paired Differences					*t*	df	Sig.
	Mean	SD	SEM	95% CI				
				Lower	Upper			
Adipose tissue L concentration								
Control Meal	-45.5	326.5	56.0	-159.4	68.4	-0.8	33	0.4
Vitamin E Meal	30.2	270.4	47.8	-24.7	-67.2	0.6	31	0.5
Tomato Puree Meal	84.5	272.1	46.7	-10.4	179.5	1.8	33	0.1
Adipose tissue Z concentration								
Control Meal	-24.4	186.4	32.0	-89.5	40.6	-0.8	33	0.5
Vitamin E Meal	9.2	124.3	22.0	-35.6	54.0	0.4	31	0.7
Tomato Puree Meal	48.4	148.1	25.4	-3.2	100.1	1.9	33	0.1

Abbr: df, degrees of freedom ; sig, significance (*p*-value) for paired *t*-test

Supplemental Table S4. Characteristics of the partial least squares regression models [a]

A. Lutein

Number of predictors	R^2	**Adjusted R^2[b]**	**R^2 after 100 permutations[c]**	**R^2 after cross-validation[c]**	**Cross-validation-ANOVA *p*-value[d]**
110 SNPs (and total cholesterol)	0.87	1.08	0.52	0.79	7.48×10^{-14}
58	0.82	1.44	0.39	0.74	3.83×10^{-12}
38	0.74	-2.46	0.28	0.68	3.10×10^{-10}
22	0.74	0.43	0.20	0.66	5.90×10^{-10}
20	0.71	0.44	0.20	0.64	2.07×10^{-9}
19	0.72	0.48	0.20	0.65	1.45×10^{-9}
16	0.69	0.50	0.18	0.62	6.38×10^{-9}
15	0.69	0.51	0.17	0.62	6.46×10^{-9}
12	0.68	0.55	0.15	0.61	9.29×10^{-9}
8	0.65	0.57	0.12	0.59	2.62×10^{-8}
7	**0.65**	**0.58**	**0.11**	**0.58**	**3.75×10^{-8}**
6	0.61	0.54	0.10	0.56	1.06×10^{-7}
5	0.58	0.52	0.09	0.54	2.34×10^{-7}
3	0.56	0.52	0.06	0.51	8.55×10^{-7}

B. Zeaxanthin

Number of predictors	R^2	**Adjusted R^{2b}**	**R^2 after 100 permutations[c]**	**R^2 after cross-validation[c]**	**Cross-validation-ANOVA *p*-value[d]**
99 SNPs (and total and LDL-cholesterol)	0.84	1.12	0.50	0.74	2.90×10^{-12}
56	0.80	1.54	0.37	0.72	2.04×10^{-11}
28	0.72	0.10	0.26	0.61	1.21×10^{-8}
23	0.72	0.37	0.24	0.62	7.51×10^{-9}
18	0.66	0.40	0.19	0.56	1.01×10^{-7}
16	0.65	0.42	0.18	0.54	2.85×10^{-7}
13	**0.65**	**0.49**	**0.17**	**0.55**	**1.52×10^{-7}**
10	0.61	0.48	0.13	0.54	2.63×10^{-7}
9	0.59	0.47	0.11	0.52	6.65×10^{-7}
7	0.50	0.39	0.09	0.43	1.65×10^{-5}
6	0.40	0.30	0.07	0.34	2.82×10^{-4}
5	0.40	0.32	0.06	0.35	2.64×10^{-4}
3	0.27	0.22	0.04	0.24	5.22×10^{-3}

[a]The selected model is highlighted in bold font. All models had 1 component.

[b]See Equation 1.

[c]See **Supplemental Figure S3** for further explanation of the procedure.

[d]See [24].

Pairwise linkage disequilibrium (LD) test: Identification of LD SNPs in the final PLS model

The PLS regression model chosen for zeaxanthin contained 13 SNPs (**Table 3B**, main manuscript). Among these, 3 were in LD according to an online calculator tool for pairwise LD (available at https://www.ensembl.org/Homo_sapiens/Tools/LD, population: European CEU, queried in 5 April 2024). Since these SNPs conveyed redundant information in the model, we kept the one which had the higher VIP value, leaving 11 SNPs in the final selected PLS regression model (Table 3B).

Supplemental Table S5. SNPs in LD in the final PLS regression model.[a]

SNP not retained in the final PLS regression model	VIP[b]	SNP retained in the final PLS regression model	VIP[b]	LD R^2[c]
rs709157	1.357	rs709158	1.364	0.87
rs1151996	1.328	rs709158	1.364	0.91

[a] Gene names can be found in Supplemental Table S1.

[b] VIP: Variable Importance in the Projection

[c]Pairwise LD R^2 results were generated from population *1000GENOMES:phase_3:CEU (accessible at* https://www.ensembl.org/Homo_sapiens/Tools/LD)

Supplemental Information: additional validations of the partial least squares (PLS) regression model.

1) Leave-k-*out* cross-validation

The leave-*k*-out validation was based on Steyerberg *et al.*[25]. Briefly, the two selected PLS regression models were accessed by randomly taking out *k* participants (*k*={1,2,3,4}) from the original dataset, thus leaving a training dataset. The *k* participants taken out were then reintroduced into this training set to assess whether the models built without these *k* participants would be able to predict accurately their adipose tissue Xanth concentrations. This test was performed up to 42 times so that each participant was taken out once (*i.e.* 42 times for *k*=1, 21 times for *k*=2, 14 times for *k*=3 and 10 times for *k*=4).

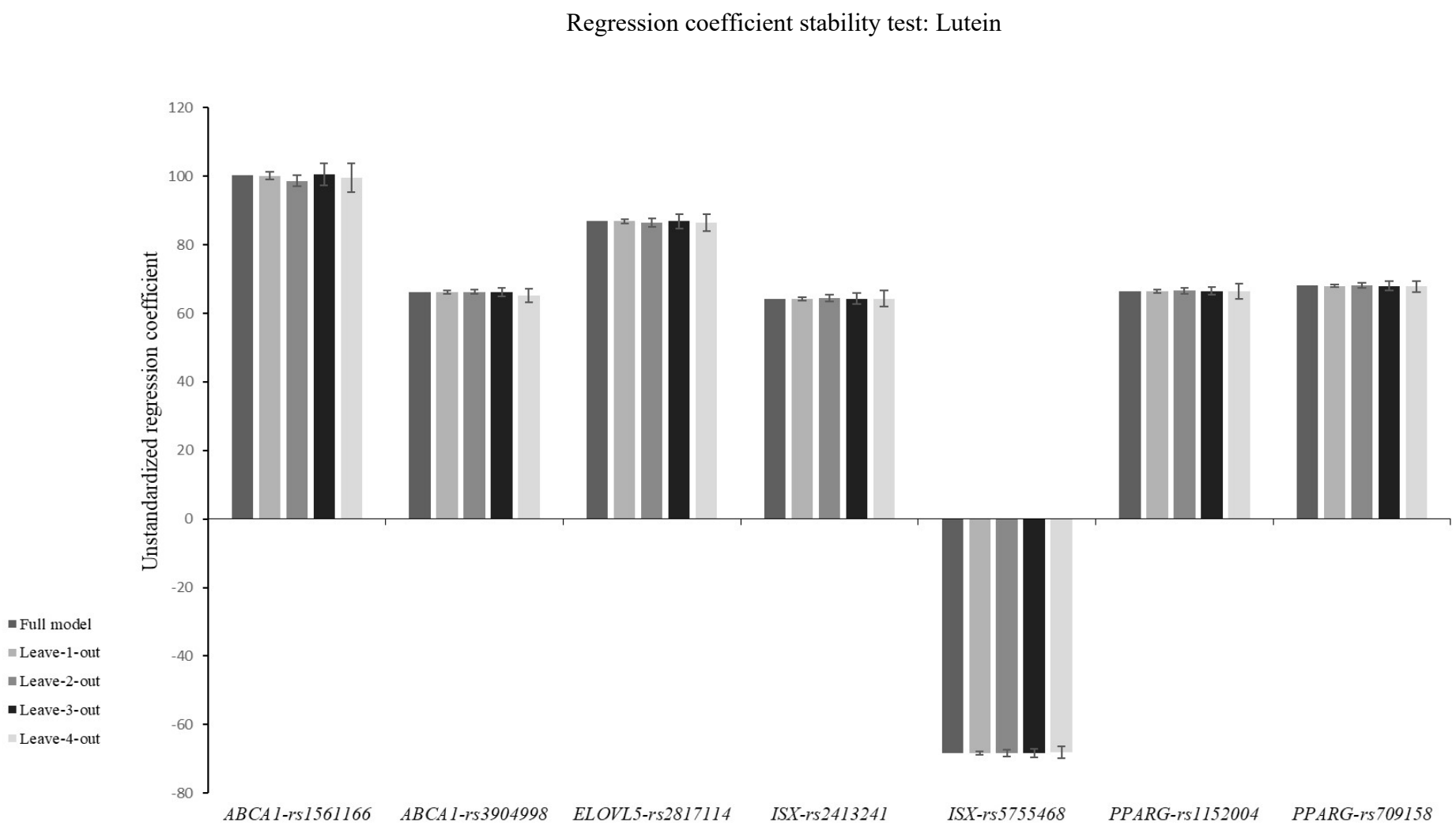
The relative prediction error (%) between the predicted and the measured adipose tissue Xanth concentrations for each *k* are presented below (**Supplemental Table S6)**. The percentage of error remained relatively stable, after leave-*k*-out cross-validation test, suggesting that the PLS regression model was relatively robust.

Supplemental Table S6. Average relative prediction error (%) following the leave-*k*-out procedure.

	Number of participants left out				
	0	1	2	3	4
Lutein (7 SNPs)	35.4	39.1	38.7	38.7	39.5
Zeaxanthin (11 SNPs)	37.1	42.4	42.5	41.9	42.7

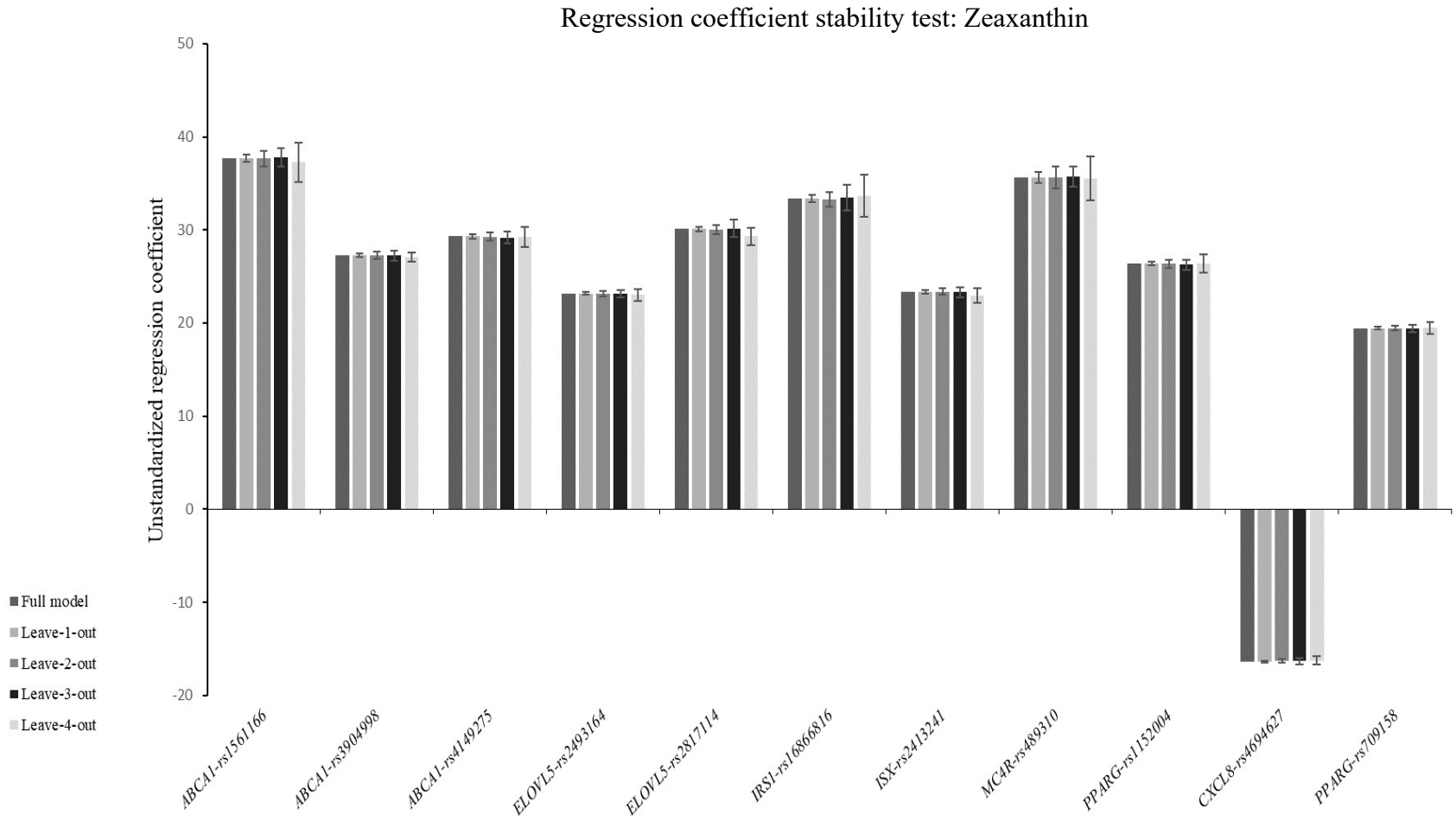
*2) Regression coefficient stability testing following the leave-*k*-out procedure*

The stability of the regression coefficients of each selected model (**Table 4**) remained unchanged (p>0.05; ANOVA) even when up to 4 participants were left out of the model (*detailed procedure is shown in Supplemental information 1*). **Supplemental Figure S2-A (lutein) and S2-B (zeaxanthin)** shows good stability of the regression coefficients with this validation.



Supplemental Figure S2-A: Regression coefficient stability validation for adipose tissue lutein concentration following the leave-*k*-out procedure. *k* participants (*k*={1,2,3,4}) were randomly removed from the original dataset, thus leaving a training subset. These participants were

then reintroduced in the training subset to assess the regression coefficients of the 7 SNPs from the selected model. This test was performed at most 42 times so that each participant was taken out once. One-way ANOVA performed for each gene showed no significant differences between the full model and the 4 training subsets generated by the procedure. Gene names are found in **Supplemental Table 1**.

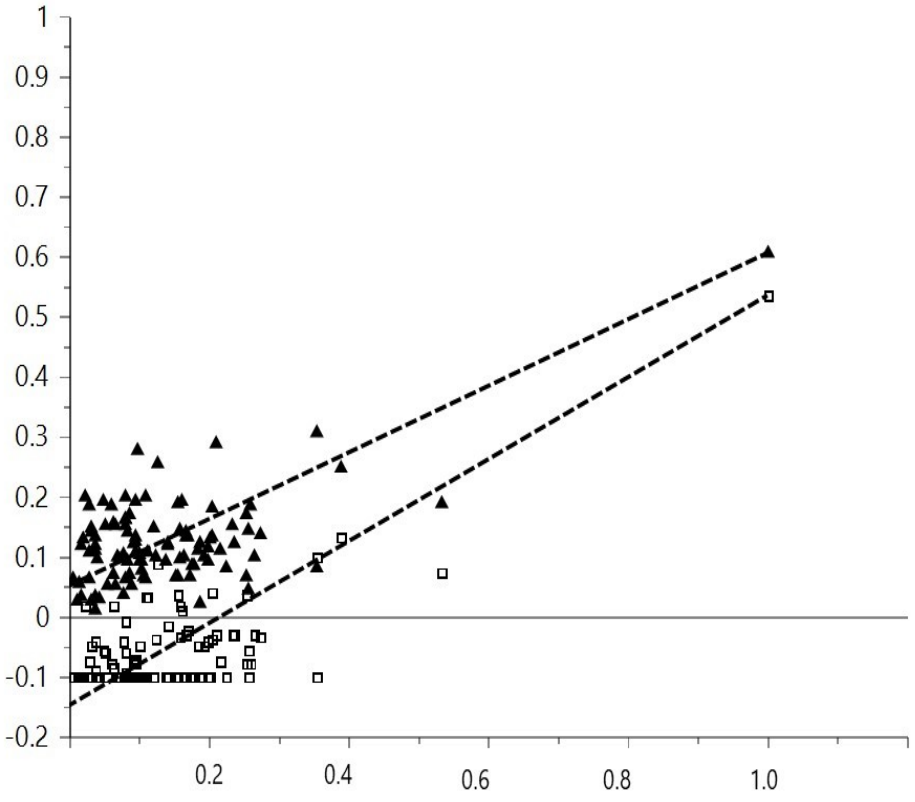


Supplemental Figure S2-B: Regression coefficient stability validation for adipose tissue zeaxanthin concentration following the leave-*k*-out procedure. *k* participants (*k*={1,2,3,4}) were randomly removed from the original dataset, thus leaving a training subset. These participants were then reintroduced in the training subset to assess the regression coefficients of the 11 SNPs from the selected model. This test was performed at

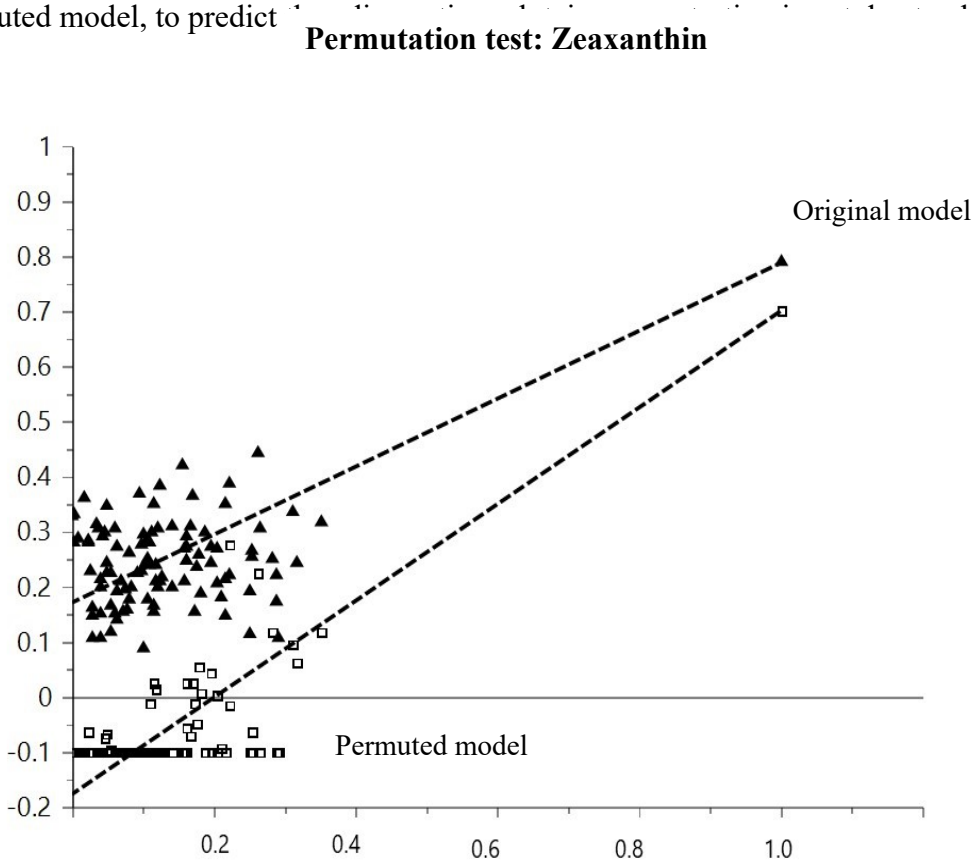
most 42 times so that each participant was taken out once. One-way ANOVA performed for each gene showed no significant differences between the full model and the 4 training subsets generated by the procedure. Gene names are found in **Supplemental Table 1**.

3) R^2 *and adjusted* R^2 *of the selected model after 100 permutations.*

This procedure 1) assessed the risk that the PLS regression model is spurious, *i.e.* the model fits the current data set well but does not predict Y well for new observations, and 2) tested for over-fitting. Briefly, the accuracy of fit (R^2 and R^2 after cross-validation) of the original model was compared with the accuracy of fit of 100 models based on data where the order of the *Y* matrices for the participants (adipose tissue L and Z concentrations) were randomly permuted, while the *X* matrices (list of selected SNPs) were kept intact. A robust model (where the fit between *X* and *Y* is high) should not be able to correctly predict the permuted *Y* variables with the intact *X* variables. **Supplemental Figure S3-A (L) and supplemental Figure S3-B (Z)** present results of these permutations using their corresponding PLS regression model.



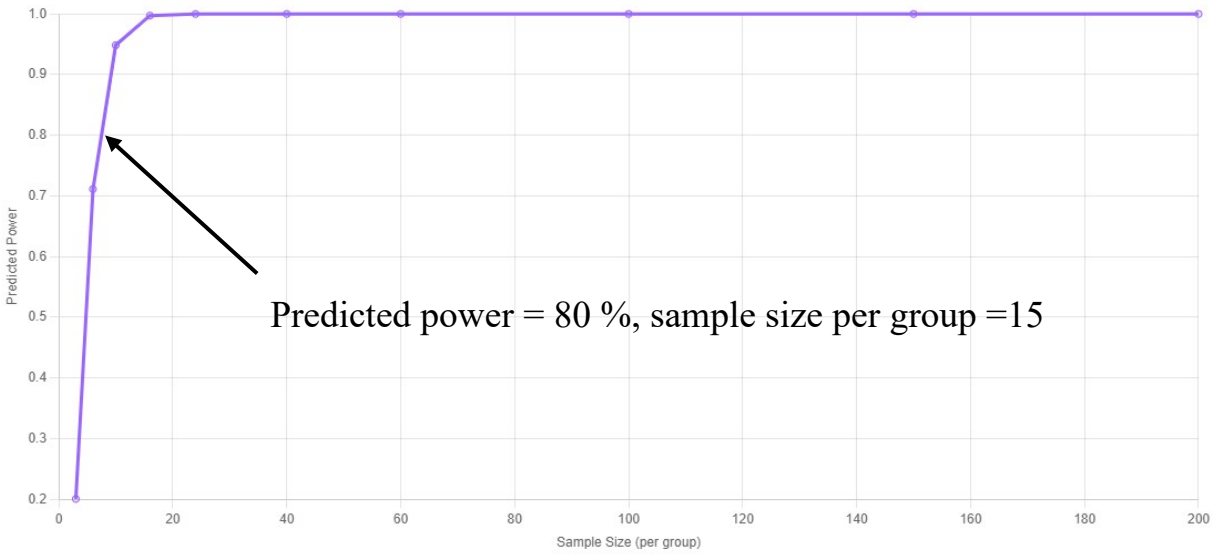
Supplemental Figure S3-A. The horizontal axis represents the correlation between the permuted *Y*'s and the original *Y*'s. The vertical axis represents the R^2 (dashed line and black triangles) and R^2 after cross-validation (dashed line and squares) values obtained in the permuted models. Values of the original model are on the far right (at correlation = 1) while values of the 100 *Y*-permuted models are further to the left. The average R^2 after 100 permutations was 0.11. This strongly supports the conclusion that the ability of the original, non-permuted model, to predict



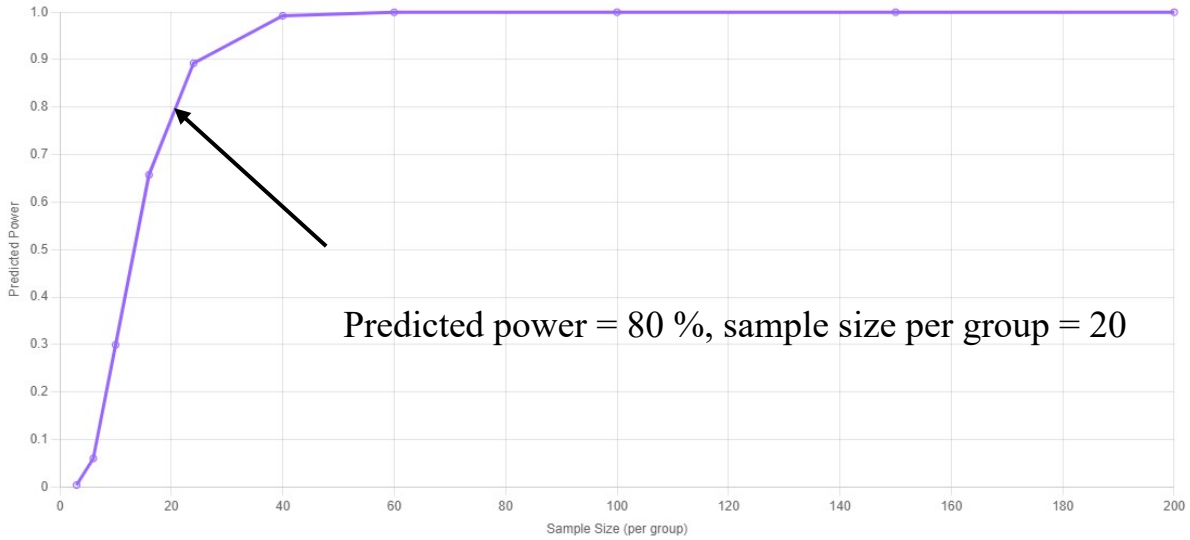
Supplemental Figure S3-B. The horizontal axis represents the correlation between the permuted *Y*'s and the original *Y*'s. The vertical axis represents the R^2 (dashed line and black triangles) and R^2 after cross-validation (dashed line and squares) values obtained in the permuted models. Values of the original model are on the far right (at correlation = 1) while values of the 100 *Y*-permuted models are further to the left. The average R^2 after 100

permutations was 0.18. This strongly supports the conclusion that the ability of the original, non-permuted model, to predict the adipose tissue zeaxanthin concentration is not due to chance.

(a) Adipose tissue lutein concentration



(b) Adipose tissue zeaxanthin concentration



Supplemental Figures S4 A and B. Retrospective multivariate power calculation was performed for 2 PLS models that incorporated the SNP variables associated with both adipose tissue lutein (7 SNPs) and zeaxanthin (11 SNPs) concentrations (refer to **Table 3**, main manuscript). With an FDR-adjusted p-value of 0.001, the predicted multivariate power of 80 % (on the vertical axis) was computed for a sample size of 15 per group for (a) adipose tissue lutein concentration and 20 per group for (b) adipose tissue zeaxanthin concentration (on the horizontal axis), affirming that the chosen sample sizes for this study (total sample size of 42 participants) were adequate. Image obtained from MetaboAnalyst 6.0 website, accessible at https://new.metaboanalyst.ca .

References:

1 P. Huebbe, J. Lange, G. Lietz and G. Rimbach, Dietary beta-carotene and lutein metabolism is modulated by the APOE genotype, *BioFactors*, 2016, **42**, 388–396.

2 M. L. Bonet, J. Ribot, S. Galmés, F. Serra and A. Palou, Carotenoids and carotenoid conversion products in adipose tissue biology and obesity: Pre-clinical and human studies, *Biochim Biophys Acta Mol Cell Biol Lipids*, 2020, **1865**, 158676.

3 J. Amengual, G. P. Lobo, M. Golczak, H. N. M. Li, T. Klimova, C. L. Hoppel, A. Wyss, K. Palczewski and J. von Lintig, A mitochondrial enzyme degrades carotenoids and protects against oxidative stress, *FASEB J*, 2011, **25**, 948–959.

4 V. Tyssandier, G. Choubert, P. Grolier and P. Borel, Carotenoids, Mostly the Xanthophylls, Exchange Between Plasma Lipoproteins, *Int J Vitam Nutr Res*, 2002, **72**, 300–308.

5 C. C. Allred, T. Krennmayr, C. Koutsari, L. Zhou, A. H. Ali and M. D. Jensen, A novel ELISA for measuring CD36 protein in human adipose tissue, *J Lipid Res*, 2011, **52**, 408–415.

6 H. Guillou, D. Zadravec, P. G. P. Martin and A. Jacobsson, The key roles of elongases and desaturases in mammalian fatty acid metabolism: Insights from transgenic mice, *Prog Lipid Res*, 2010, **49**, 186–199.

7 S. Bandara, J. Moon, S. Ramkumar and J. von Lintig, ASTER-B regulates mitochondrial carotenoid transport and homeostasis, *J Lipid Res*, 2023, **64**, 100369.

8 S. Bandara, S. Ramkumar, S. Imanishi, L. D. Thomas, O. B. Sawant, Y. Imanishi and J. von Lintig, Aster proteins mediate carotenoid transport in mammalian cells, *Proc Natl Acad Sci U S A*, 2022, **119**, e2200068119.

9 M. S. F. Lavrador, M. S. Afonso, D. E. Cintra, M. Koike, V. S. Nunes, M. Demasi, C. J. Lin, L. M. M. Beda, L. A. Gioielli, R. de P. A. Bombo, R. M. Machado, S. Catanozi, E. R. Nakandakare and A. M. Lottenberg, Interesterified Fats Induce Deleterious Effects on Adipose Tissue and Liver in LDLr-KO Mice, *Nutrients*, 2019, **11**, 466.

10 X. Zhang, C. C. Zhang, H. Yang, K. G. Soni, S. P. Wang, G. A. Mitchell and J. W. Wu, An Epistatic Interaction between Pnpla2 and Lipe Reveals New Pathways of Adipose Tissue Lipolysis, *Cells*, 2019, **8**, 395.

11 R. Zechner, J. Strauss, S. Frank, E. Wagner, W. Hofmann, D. Kratky, M. Hiden and S. Levak-Frank, The role of lipoprotein lipase in adipose tissue development and metabolism, *Int J Obes Relat Metab Disord*, 2000, **24 Suppl 4**, S53-56.

12 J. Z. Long and B. F. Cravatt, The metabolic serine hydrolases and their functions in mammalian physiology and disease, *Chem Rev*, 2011, **111**, 6022–6063.

13 R. Zechner, P. C. Kienesberger, G. Haemmerle, R. Zimmermann and A. Lass, Adipose triglyceride lipase and the lipolytic catabolism of cellular fat stores, *J Lipid Res*, 2009, **50**, 3–21.

14 A. Koppen and E. Kalkhoven, Brown vs white adipocytes: the PPARgamma coregulator story, *FEBS Lett*, 2010, **584**, 3250–3259.

15 L. Yvan-Charvet, A. Bobard, P. Bossard, F. Massiéra, X. Rousset, G. Ailhaud, M. Teboul, P. Ferré, G. Dagher and A. Quignard-Boulangé, In vivo evidence for a role of adipose tissue SR-BI in the nutritional and hormonal regulation of adiposity and cholesterol homeostasis, *Arterioscler Thromb Vasc Biol*, 2007, **27**, 1340–1345.

16 I. Domarkienė, A. Mažeikienė, G. Petrauskaitė, Z. A. Kučinskienė and V. Kučinskas, New associations of serum β-carotene, lycopene, and zeaxanthin concentrations with NR1H3, APOB, RDH12, AND CYP genes, *Food Sci Nutr*, 2022, **10**, 763–771.

17 K. J. Meyers, E. J. Johnson, P. S. Bernstein, S. K. Iyengar, C. D. Engelman, C. K. Karki, Z. Liu, R. P. Igo, B. Truitt and M. L. Klein, Genetic determinants of macular pigments

in women of the Carotenoids in Age-Related Eye Disease Study, *Invest Ophthalmol Vis Sci*, 2013, **54**, 2333–2345.

18 P. Borel, C. Desmarchelier, M. Nowicki and R. Bott, Lycopene bioavailability is associated with a combination of genetic variants, *Free Radic Biol Med*, 2015, **83**, 238–244.

19 P. Borel, C. Desmarchelier, M. Nowicki, R. Bott, S. Morange and N. Lesavre, Interindividual variability of lutein bioavailability in healthy men: characterization, genetic variants involved, and relation with fasting plasma lutein concentration, *Am J Clin Nutr*, 2014, **100**, 168–175.

20 P. Borel, C. Desmarchelier, M. Nowicki and R. Bott, A combination of single-nucleotide polymorphisms is associated with interindividual variability in dietary β-carotene bioavailability in healthy men, *J Nutr*, 2015, **145**, 1740–1747.

21 C. Desmarchelier, J.-C. Martin, R. Planells, M. Gastaldi, M. Nowicki, A. Goncalves, R. Valéro, D. Lairon and P. Borel, The postprandial chylomicron triacylglycerol response to dietary fat in healthy male adults is significantly explained by a combination of single nucleotide polymorphisms in genes involved in triacylglycerol metabolism, *J Clin Endocrinol Metab*, 2014, **99**, E484-488.

22 M. P. Zumaraga, P. Borel, R. Bott, M. Nowicki, D. Lairon and C. Desmarchelier, The Interindividual Variability of Phytofluene Bioavailability is Associated with a Combination of Single Nucleotide Polymorphisms, *Mol Nutr Food Res*, 2022, 2200580.

23 H. Bozdogan, Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions, *Psychometrika*, 1987, **52**, 345–370.

24 L. Eriksson, J. Trygg and S. Wold, CV-ANOVA for significance testing of PLS and OPLS® models, *J Chemom*, 2008, **22**, 594–600.

25 E. W. Steyerberg, F. E. Harrell, G. J. J. M. Borsboom, M. J. C. Eijkemans, Y. Vergouwe and J. D. F. Habbema, Internal validation of predictive models, *J Clin Epidemiol*, 2001, **54**, 774–781.