Genetic determinants of general cognitive function and their association to circulating metabolites: a cross-omics study

Dina Vojinovic, Caroline Hayward, Jennifer A. Smith, Wei Zhao, Jan Bressler, Stella Trompet, Chloé Sarnowski, Murali Sargurupremraj, Jingyun Yang, Paul R.H.J. Timmers, Narelle K. Hansell, Ari Ahola-Olli, Eva Krapohl, Joshua C. Bis, Daniel E. Gustavson, Teemu Palviainen, Yasaman Saba, Anbu Thalamuthu, Sudheer Giddaluru, Leonie Weinhold, Najaf Amin, Nicola Armstrong, Lawrence F. Bielak, Anne C. Böhmer, Patricia A. Boyle, Henry Brodaty, Harry Campbell, David W. Clark, Baptiste Couvy-Duchesne, Philip L De Jager, Jeremy A. Elman, Thomas Espeseth, Jessica D. Faul, Annette Fitzpatrick, Scott D. Gordon, Thomas Hankemeier, Edith Hofer, M. Arfan Ikram, Peter K. Joshi, Rima Kaddurah-Daouk, Jaakko Kaprio, Sharon LR Kardia, Katherine A. Kentistou, Luca Kleineidam, Nicole Kochan, John Kwok, Markus Leber, Teresa Lee, Terho Lehtimäki, Anu Loukola, Anders Lundquist, Leo-Pekka Lyytikäinen, Karen Mather, Grant W. Montgomery, Simone Reppermund, Richard J. Rose, Suvi Rovio, Perminder Sachdev, Matthias Schmid, Helena Schmidt, Andre G. Uitterlinden, Eero Vuoksimaa, Michael Wagner, Holger Wagner, David R. Weir, Margaret J. Wright, Miao Yu, Lars Nyberg, Alfredo Ramirez, Stephanie Le Hellard, David Ames, Peter Schofield, Reinhold Schmidt, Danielle Dick, David Porteous, William S. Kremen, Bruce M. Psaty, Olli Raitakari, Nicholas G. Martin, James F. Wilson, David A. Bennett, Stephanie Debette, J. Wouter Jukema, Thomas H Mosley, Jr, Sudha Seshadri, Cornelia M. van Dujin

This chapter is in preparation.

The supplemental information for this paper is available at https://drive.google.com/drive/folders/1yQGEPmTmY4kcnrsx8EiosNPL1w9rUx57?usp=sharing

ABSTRACT

General cognitive function is a heritable predictor of health outcomes. Here, we performed a genome-wide association study of general cognitive function in 245,117 participants of European (EA) and African American (AA) ancestry from Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium and UK Biobank. We reported 32 novel genetic loci in individuals of EA which have previously been associated with various disorders including psychiatric illnesses such as schizophrenia, autistic disorder, bipolar disorder, depression, mood and anxiety disorders. The risk score based on our findings and previously identified genetic risk loci underlying general cognitive function in EA was associated with general cognitive function in individuals of AA ancestry (N = 2,117). Genes associated with general cognitive function could be linked to circulating metabolites including tyrosine, creatinine, 22:6 docosahexaenoic acid (DHA), glycoprotein acetyl, acetate, and citrate. Using Mendelian randomization, we examined whether these metabolites were cause or rather a consequence of biological pathways underlying cognitive function. Genes determining glycoprotein acetyls and tyrosine also determine general cognitive function, suggesting that these metabolites are in the causal pathway, whereas DHA is rather a consequence of the physiological process determining cognitive function. These results provide new insights into general cognitive function.

INTRODUCTION

General cognitive function is an important predictor of health outcomes, including mortality and morbidity varying from dementia to depression and other psychiatric diseases.[1-4] Differences in cognitive function are determined by various factors including lifestyle and genetic factors.[5] Morbidities such as cardiometabolic diseases and cancer also contribute to cognitive performance and cognitive decline in later life.[6] For long, the relationship between metabolic factors and cognitive function was poorly understood. We recently identified circulating metabolites to be associated with the general cognitive function in healthy individuals.[7] The metabolic profile included the subfractions of high-density lipoprotein particles, fatty acids, amino acids, and acute phase reaction markers.[7,8] We successfully associated these metabolites to environmental factors such as lifestyle and diet.[7] A question to answer is whether these metabolites are in the causal pathway and a target of cognitive function or rather a consequence of physiological processes underlying general cognitive function and associated lifestyle and pathology.

Our human genome is another major driver of the circulating metabolites and general cognitive function.[9] General cognitive function has heritability of 50% and over 140 genomic regions have been identified in the genome-wide association studies (GWASs) performed to date.[10-13] Despite the overwhelming progress, the polygenic profile score capturing the joint effects of those variants explained only up to 4.3% of trait variance.[11] Yet, common genetic variants underlying general cognitive function were associated with various neurological and psychiatric disorders when checking for genome-wide genetic overlap using LD score regression.[14] Based on this method, there is also evidence for a shared genetic origin with body mass index, waist to hip ratio, high-density lipoprotein cholesterol, and cardiovascular diseases which are key drivers of the human metabolism.[11] Up until now, we have not linked these genetic determinants to the metabolites in the circulation, which may bring to surface new insights in metabolic pathways that play a key role in general cognitive function. An omission of previous studies is that only participants of European ancestry were included. A question to answer is whether the findings are generalizable to other ethnic groups. Finally, the studies of general cognitive function conducted to date mainly focused on the imputations generated as part of the 1000 Genomes Project. Recently, the Haplotype Reference Consortium (HRC) made available a large haplotype reference panel which increased imputation accuracy.[15,16]

Here, we performed a GWAS of general cognitive function in 243,000 participants of European ancestry and 2,117 participants of African-American from Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium and UK biobank

ancestry using HRC imputation panel.[11] We examined the association of genes implicated in general cognitive function and circulating metabolites and evaluated their causal relationship using Mendelian randomization approach.

METHODS

Study population

Our study population encompassed 243,000 participants of European ancestry (EA) from 32 studies from CHARGE consortium and UK biobank and 2,117 participants of African-American (AA) ancestry from 3 studies that were part of the CHARGE consortium.[11,17] Participating studies are described in detail in the **Supplementary Note**. General characteristics of study populations are provided in **Supplementary Table 1**. Local ethical committees or the institutional review boards approved each of the studies and written informed consent was obtained from all participants.

Phenotype assessment

The general cognitive function was constructed from a number of cognitive tasks for each of the CHARGE cohorts.[18] Each participating study performed principal component analysis using at least three cognitive tests that assess different cognitive domains. Only one score was used from each of the cognitive tests. The general cognitive function was the first unrotated principal component. The phenotype was constructed in such a way that higher score indicated higher cognitive function. Participants with dementia and stroke were excluded. Information on cognitive tests used to create general cognitive function score in each participating study and correlation between the general cognitive function and each cognitive test is provided in **Supplementary Table 2.** General cognitive function explained between 34.7% and 59.3% of the total test variance.

The cognitive test from UK Biobank was a verbal and numerical reasoning score assessed by 13 multiple choice questions which has a high genetic correlation with general cognitive function.[11,19] A detailed information regarding the samples of UK Biobank participants with verbal-numerical reasoning scores is provided elsewhere.[11] In the current analysis, four samples of UK Biobank participants were used.

Genotyping and imputation

Description of genotyping platforms, calling method and quality control procedures in each of the CHARGE cohorts is provided in **Supplementary Table 3**. The study participants were genotyped using commercially available genotyping arrays. Each study used free imputation servers (Michigan or Sanger) to perform genotype imputation using

Haplotype Reference Consortium (HRC) reference panel.[16] Description of genotyping platforms and quality control in UK Biobank is provided in Bycroft *et al.* (http://www.biorxiv.org/content/early/2017/07/20/166298).

Genome-wide association analysis

Each of the CHARGE cohorts performed genome-wide association analysis of general cognitive function while adjusting for age, gender, principal components if needed, familial relationship if appropriate and study center if needed. Details on analysis methods for each cohort are provided in **Supplementary Table 3**. The quality control (QC) was performed using EasyQC.[20] Genetic variants with low imputation quality ($r^2 < 0.5$) or minor allele count below 5 were removed. The genome-wide association summary results of verbal and numerical reasoning score in UK Biobank were obtained from http://www.ccace.ed.ac.uk/node/335.[11] The summary statistic results of CHARGE participating studies and UK biobank were combined using sample size weighted meta-analysis in METAL.[21] The meta-analysis was performed separately for each ethnic group. For each genome-wide association analysis, LD score regression method was used to estimate intercept which can distinguish between the inflation due to a polygenic signal and the inflation due to population stratification or cryptic relatedness.[22]

Conditional association analysis

Approximate conditional genome-wide analysis was performed using Genome-wide Complex Trait Analysis (GCTA), version 1.26.0, in order to identify genetic variants conditionally independent on genetic signals previously reported in the largest GWAS of general cognitive function to date.[11,23] Genetic variants with high collinearity (0.9) were ignored. Complete linkage equilibrium was assumed for genetic variants located more than 10Mb away from each other. The linkage disequilibrium pattern (LD) between the genetic variants was calculated based on data of 11,496 individuals from the Rotterdam Study imputed with HRC reference panel.

Characterization of genomic loci

Genomic loci were characterized using Functional Mapping and Annotation of genetic associations (FUMA).[24] First, independent genetic variants were defined as genome-wide significant variants that are not in linkage disequilibrium with each other ($r^2 < 0.6$).[24] Independent significant variants with $r^2 \geq 0.1$ were assigned to the same genomic risk locus and were merged into a single locus if they were 250 bp or closer.[24] Each genomic risk locus was represented by the top lead genetic variant defined as an independent significant variant ($r^2 < 0.1$).

Functional annotation

The genome-wide significant variants were annotated using the Combined Annotation Dependent Depletion (CADD), HaploReg v4.1, and RegulomeDB tools.[25,26] Furthermore, GTEx data was used to determine whether these genetic variants have an effect on expression.[27]

Correlating genetic determinants of cognition and circulating metabolites

Association of individual genetic variants underlying general cognitive function was explored in a GWAS of circulating metabolites including ~25,000 individuals.[11,28] The metabolites were measured by nuclear magnetic resonance (^{1}H-NMR) on Nightingale Health platform. To model correlation between metabolites and linkage disequilibrium between the genetic variants, we first calculated the number of independent tests using the method of Li and Ji.[29] The Bonferroni corrected *p*-value was calculated based on the number of independent tests and set at 0.05/(32 independent metabolites × 342 independent genetic variants) = 4.57×10^{-6}.

Next, more global test was used to link general cognitive function and metabolites using the LD score regression method. Genetic correlation was estimated between general cognitive function and metabolites measured by ^{1}H-NMR on Nightingale Health platform.[30] The analyses were performed using a web interface, LD-hub.[31] The significance threshold was determined based on a number of traits tested and was set at 0.05/111 = 4.5×10^{-4}.

Mendelian randomization

To evaluate whether the association of the metabolites to general cognitive function is a cause or consequence of the physiological processes underlying general cognitive function bidirectional Mendelian randomization was performed for each metabolite associated with genetic variants underlying general cognitive function. The associations were estimated based on the present GWAS and that of circulating metabolites including ~25,000 individuals.[28] The effect of genetic risk score was constructed using the summary statistic level data and method implemented in gtx package.[32] Genetic risk scores based on more than 5 genetic variants that explain more than 1% of variance in exposure were taken forward.

RESULTS

Genome-wide association study of general cognitive function in individuals of EA

A detailed description of the genome-wide association analysis in individuals of EA is given in the supplementary material. The quantile-quantile plot suggested inflation (λ = 1.62, mean χ^2 = 1.9) (**Supplementary Figure 1, Supplementary Table 4**). However, LD score regression revealed intercept of 1.049 (SE = 0.012) and a ratio of 0.0568 suggesting that inflation is mainly due to polygenicity and only 5.68% of the inflation is due to other causes. General cognitive function was associated with 9,521 genetic variants distributed across all autosomal chromosomes at genome-wide significance level (*p*-value < 5×10^{-8}), of which 358 independent genetic variants mapped to 139 genomic loci (**Supplementary Table 5-7**). After conditioning on genetic signals previously reported in the largest GWAS of general cognitive function to date,[11] 311 genetic variants surpassed genome-wide significance threshold, including 33 novel independent genetic variants mapped to 32 genomic loci (**Figure 1, Supplementary Table 8-10**). The list of pleiotropic associations for these variants and tagged variants is provided in **Supplementary Table 11**, whereas the genes to which these independent variants were mapped to and disease they have been implicated in are listed in **Supplementary Table 12**.

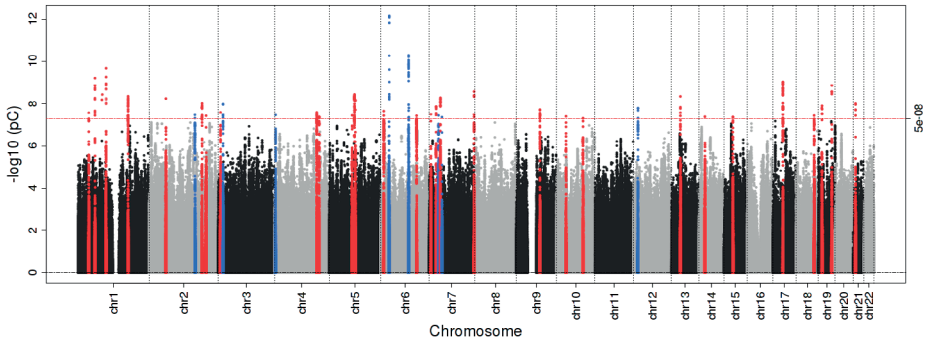


Figure 1. The results of genome-wide association meta-analysis including participants of EA in CHARGE cohorts and UK biobank after conditioning on genetic variants identified in the largest GWAS of general cognitive function to date. The novel loci defined as > 1Mb from previously reported genome-wide variants are depicted in red, whereas the known loci are depicted in blue.

Genome-wide association study of general cognitive function in individuals of AA

There was no variant that surpassed the genome-wide significant threshold in the sample of AA. Among 9,521 genome-wide significant variants associated with general cognitive function in EA cohorts, 66.7% genetic variants had the same direction of effect size in AA individuals and 10.6% variants showed at least nominal evidence of significance (*p*-value < 0.05) (**Supplementary Table 5** and **8** for all loci and independent loci,

respectively). When combining the 148 loci described previously and the 32 discovered in EA in our study into a genetic risk score, the EA risk score was significantly associated with the general cognitive function in AA (p-value = 1.88×10^{-7}).

Correlating genetic determinants of cognition and circulating metabolites

To explore association of the genes involved in general cognitive function and metabolites, we first examined the association of individual genetic variants underlying general cognitive function in the GWAS of circulating metabolites (**Supplementary Table 13**). Two associations surpassed the threshold for multiple testing (**Figure 2**). A genetic variant in *PKD1L3* was associated with lower level of tyrosine (p-value = 5.7×10^{-7}), whereas a variant near *ITIH1* was associated with lower levels of creatinine (p-value = 2.3×10^{-6}). Next, we performed a global genetic test to link general cognitive function using LD score regression. Nominally significant genome-wide genetic overlap was observed between general cognitive function and circulating metabolites including acetate ($\rho_{genetic}$ = 0.21, p-value = 3.9×10^{-3}), citrate ($\rho_{genetic}$ = 0.18, p-value = 8.3×10^{-3}), glycoprotein acetyls ($\rho_{genetic}$ = -0.12, p-value = 2.8×10^{-2}), and 22:6 docosahexaenoic acid (DHA) ($\rho_{genetic}$ = 0.12, p-value = 4.8×10^{-2}) (**Figure 3**).

Mendelian randomization

To evaluate whether the association of the metabolites that were associated with general cognitive function in the single variant or global evaluation are a cause or consequence of the physiological processes underlying general cognitive function, we performed a Mendelian randomization experiment. When testing the hypothesis that the genes determining cognition are also implicated in circulating metabolites, we found evidence for such mechanism for DHA (p-value = 1.3×10^{-5}) when adjusting for multiple testing (**Figure 4**). When testing the hypothesis that the genes determining circulating metabolites also determine general cognitive function, we found evidence for such mechanism for tyrosine (p-value = 5.8×10^{-5}) and glycoprotein acetyls (p-value = 8.99×10^{-3}) (**Figure 4**).

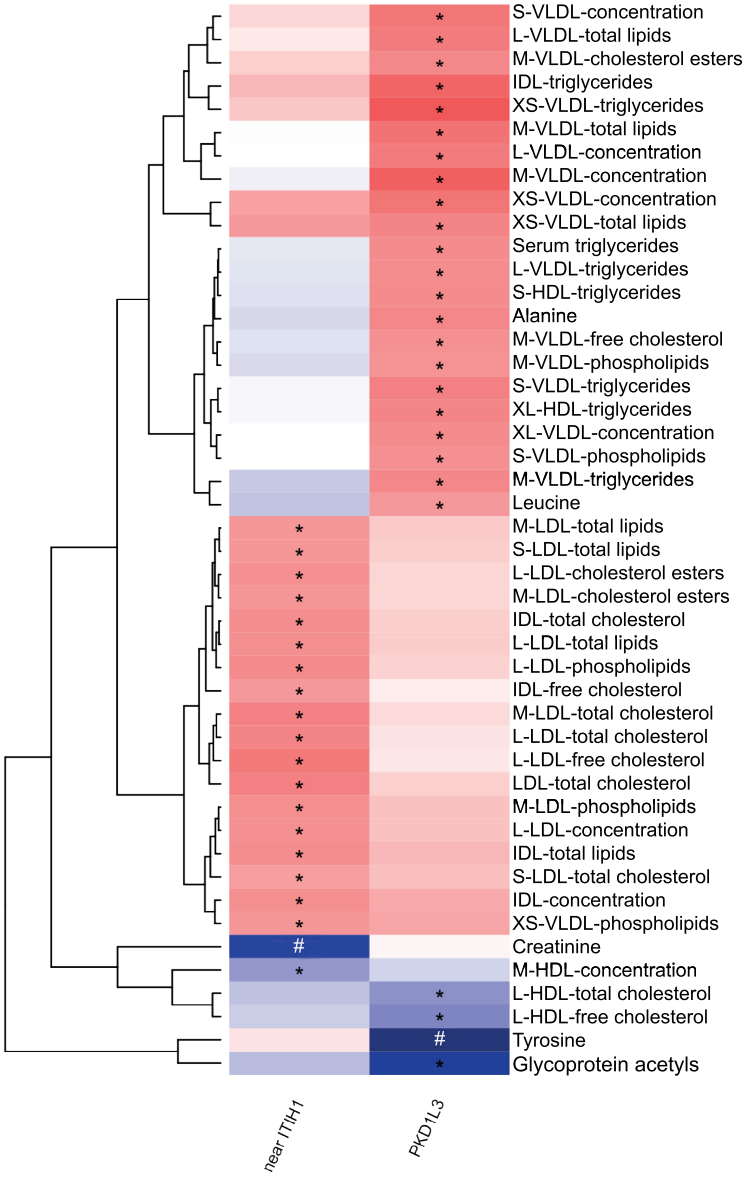


Figure 2. Genetic determinants of general cognitive function and metabolites. Red color denotes positive association and blue color depicts inverse association. Associations that surpassed threshold for multiple testing (p-value < 4.6×10^{-6}) are indicated by hash symbol, whereas nominal associations (p-value < 0.05) are labeled with star.

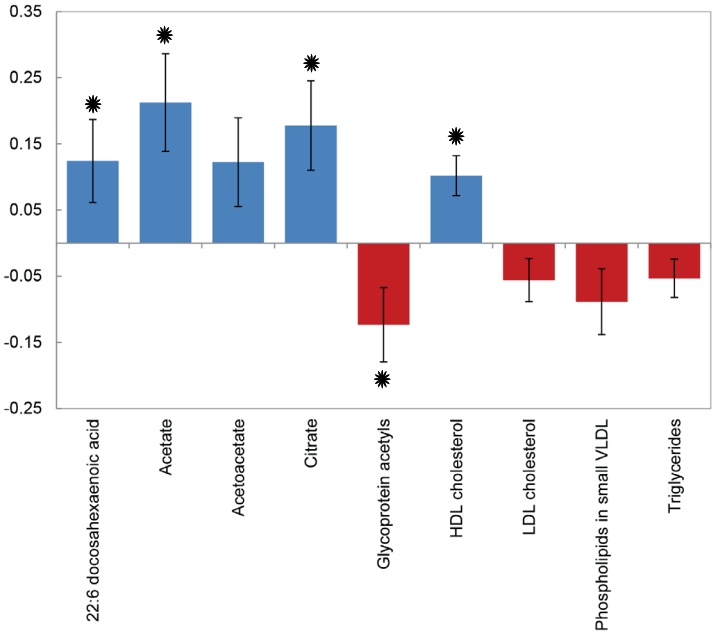


Figure 3. The top results of genetic correlation between the general cognitive function (summary statistic generated in our project) and metabolites for which summary statistic data was available on LD-hub (*p*-value < 0.2). The nominally significant associations (*p*-value < 0.05) are denoted with a star.

DISCUSSION

In a GWAS of general cognitive ability in 243,000 individuals of EA we detected 32 novel findings, bringing the total number of independent loci implicated in general cognitive function up to 180. The risk score based on 180 loci was significantly associated with general cognitive function in AA. Two genes implicated in general cognitive function could directly be linked to circulating levels of tyrosine and creatine, whereas more global genome-wide genetic overlap was found for DHA, glycoprotein acetyl, acetate, and citrate. Mendelian randomization suggests that genes determining glycoprotein acetyl and tyrosine also determine general cognitive function while DHA is rather a consequence of the physiological process determining cognitive function.

Among the 32 novel loci, the variants with the highest probability of having regulatory function based on RegulomeDB score (1f, 1d) were mapped to 17q12 (**Supplementary Table 9**). Deletion of this region has been associated with a syndrome in which about half of the people have delayed development, intellectual disability, or psychiatric disorders such as autism spectrum disorder, schizophrenia, anxiety, and bipolar disorder.[33] The lead variant at 17q12 was mapped to *DHRS11* gene that metabolizes steroid hor-

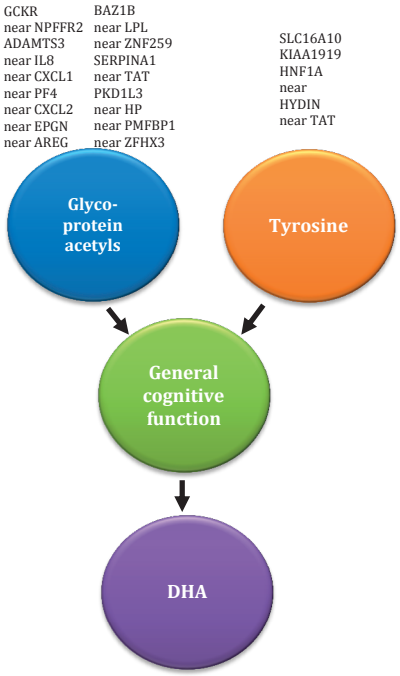


Figure 4. Suggested paths for general cognitive function and metabolites. The genes used in the genetic risk score are located above the metabolite name.

mones, prostaglandins, retinoids, lipids, and xenobiotics (**Supplementary Table 9**).[34] Additionally, other novel locus 6p21.33 had also RegulomeDB score of 1f and is located near *IER3* gene previously related to inflammatory diseases and hypertension.[35,36] Other independent significant variants at novel loci were mapped to genes previously implicated in bipolar disorder, autistic disorder, or near the genes previously implicated in depression, schizophrenia, mood and anxiety disorders (**Supplementary Table 12**).[37-41] Our findings also overlap with those of Savage *et al*. which targeted intelligence rather than general cognitive findings.[42] When performing LD score regression of the two traits the genetic correlation was 0.98. Yet there are differences observed in lead variants in the same regions which imply heterogeneity: statistical or genetic.

This study is unique in two aspects. First, it included a study of genomic variants implicated in general cognitive function in AA. Although this study was too small to yield genome-wide significant findings, the study showed that the variants implicated in EA are also highly significantly determining general cognitive function in AA. Second, our study has successfully examined the role of the genes implicated in general cognitive function in metabolic changes in the circulation. Our findings showed association with circulating metabolites including tyrosine and creatinine. A missense genetic variant in *PKD1L3* at 16q22.2 was associated with general cognitive function (p-value = 1.13×10^{-11}).

This variant, predicted to be damaging (PolyPhen = 0.99), showed association with tyrosine. Tyrosine is an amino acid that plays an important role in synthesis of dopamine, a key neurotransmitter in the brain, and is known to modulate cognitive functions in healthy population by modulating dopamine function.[43,44] Several diseases that involve dopamine dysfunction, such as Parkinson's disease, schizophrenia, and attention deficit hyperactivity, also show alteration of cognitive function.[45] Of note is that there is also a genetic correlation of general cognitive function and schizophrenia and attention deficit hyperactivity disorder, suggesting common pathogenesis. This finding not only highlights the potential mechanism through which the established genetic variants of general cognitive function may act but also opens opportunities to prevent cognitive decline by targeting tyrosine, which is according to our Mendelian randomization experiments most likely in the causal pathway. Also, the finding that the established genetic variant of general cognitive function (rs3755799, 1f RegulomeDB score) modulates levels of creatinine is interesting from a preventive perspective. Creatinine is an organic molecule and a breakdown product of muscle creatine phosphate and it is widely used as a measure of renal function. Higher levels of creatinine and renal impairment were previously associated with lower cognitive performance and increased risk of dementia, asking for further follow-up research in view of potential preventive interventions.[46,47]

In the LD regression, we identified at nominal significance genetic correlation between general cognitive function and several circulating metabolites including DHA, acetate, citrate, and glycoprotein acetyls. These were also included in a formal Mendelian randomization to evaluate whether the metabolites are more likely in the causal pathway or rather a consequence of the various biological processed determining general cognitive function. Unfortunately, the Mendelian randomization experiments for citrate and acetate were not possible because reliable data on genetics was missing. Mendelian randomization suggests that glycoprotein acetyl is causally related to general cognitive function while the association with DHA is rather a consequence of the physiological process determining cognitive function. Circulating levels of DHA, a long-chain omega-3 polyunsaturated fatty acid, have been associated with cognitive function and risk of Alzheimer's disease and dementia.[7] These are strongly associated with fish consumption.[7] However, circulating levels of DHA are also genetically determined and subject of enzymatic processes.[28] Our Mendelian randomization experiment suggests that endogenous processes related to cognitive function are related to levels of DHA rather than that DHA is a driver of general cognitive function. Glycoprotein acetyls appear to be more likely in the causal pathway. Circulating glycoprotein acetyl levels, a marker of acute phase reaction, have been implicated in chronic inflammatory disease and cancer.[48] The levels of this protein were also associated with future risk of all-cause mortality.[49] The fact

that at middle age, these proteins also associate to lower cognitive ability asks for more research on what influences circulating levels of glycoprotein acetyls.[7]

The strengths of our study are large sample size, population-based design, use of large imputation reference panel, and integration of genetic and metabolomics data. However, our study also has limitations. The association results of participating cohorts were combined using a sample-size weighted meta-analysis due to phenotypic heterogeneity limiting discussion on effect sizes. Despite concordance in the direction of effects observed between EA and AA participants, limited sample size of AA sample yielded low statistical power and influenced our ability to explore genetic determinants of general cognitive function in other ethnic groups. Future research studies should focus on non-Europeans. When exploring metabolic pathophysiology underlying genetic variants associated with the general cognitive function, we focused on circulating metabolites measured by NMR technology using Nightingale Health platform.[7] This platform detects various metabolites including amino acids, ketone bodies, fatty acids, and a large proportion of metabolites are lipoproteins and lipid subclasses which provides an excellent opportunity to study cognition.[50] However, the studied metabolites represent only small proportion of circulating metabolites, therefore, future studies focusing on a wider spectrum of metabolites are needed.[51] By improving the power of GWASs of metabolites, novel associations may be revealed.

We have reported association of general cognitive function with 32 novel genetic loci in EA sample and showed that there is genetic overlap of the loci determining general cognitive function in EA and AA. We have also found association of established and novel genetic determinants of general cognitive function and circulating metabolites, providing a starting point for new preventive studies.

REFERENCES

1. Deary, I.J., Weiss, A. & Batty, G.D. Intelligence and personality as predictors of illness and death: How researchers in differential psychology and chronic disease epidemiology are collaborating to understand and address health inequalities. *Psychological Science in the Public Interest***.11**, pp (2010).
2. Batty, G.D., Deary, I.J. & Gottfredson, L.S. Premorbid (early life) IQ and later mortality risk: systematic review. *Ann Epidemiol* **17**, 278-88 (2007).
3. Wraw, C., Deary, I.J., Gale, C.R. & Der, G. Intelligence in youth and health at age 50. *Intelligence. Vol.53 2015, pp. 23-32. Nov-Dec Intelligence* (2015).
4. Hagenaars, S.P., Gale, C.R., Deary, I.J. & Harris, S.E. Cognitive ability and physical health: a Mendelian randomization study. *Sci Rep* **7**, 2651 (2017).
5. Yaffe, K. *et al.* Predictors of maintaining cognitive function in older adults: the Health ABC study. *Neurology* **72**, 2029-35 (2009).
6. Calvin, C.M. *et al.* Intelligence in youth and all-cause-mortality: systematic review with meta-analysis. *Int J Epidemiol* **40**, 626-44 (2011).
7. van der Lee, S.J. *et al.* Circulating metabolites and general cognitive ability and dementia: Evidence from 11 cohort studies. *Alzheimers Dement* (2018).
8. Toledo, J.B. *et al.* Metabolic network failures in Alzheimer's disease: A biochemical road map. *Alzheimers Dement* **13**, 965-984 (2017).
9. Draisma, H.H.M. *et al.* Genome-wide association study identifies novel genetic variants contributing to variation in blood metabolite levels. *Nat Commun* **6**, 7208 (2015).
10. Davies, G. *et al.* Genetic contributions to variation in general cognitive function: a meta-analysis of genome-wide association studies in the CHARGE consortium (N=53 949). *Molecular Psychiatry* **20**, 183-192 (2015).
11. Davies, G. *et al.* Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. *Nat Commun* **9**, 2098 (2018).
12. Haworth, C.M. *et al.* The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Mol Psychiatry* **15**, 1112-20 (2010).
13. Plomin, R. & Deary, I.J. Genetics and intelligence differences: five special findings. *Mol Psychiatry* **20**, 98-108 (2015).
14. Davies, G. *et al.* Ninety-nine independent genetic loci influencing general cognitive function include genes associated with brain health and structure (N = 280,360). *biorxiv* (2017).
15. Iglesias, A.I. *et al.* Haplotype reference consortium panel: Practical implications of imputations with large reference panels. *Hum Mutat* **38**, 1025-1032 (2017).
16. McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet* **48**, 1279-83 (2016).
17. Psaty, B.M. *et al.* Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium: Design of prospective meta-analyses of genome-wide association studies from 5 cohorts. *Circ Cardiovasc Genet* **2**, 73-80 (2009).
18. Davies, G. *et al.* Genetic contributions to variation in general cognitive function: a meta-analysis of genome-wide association studies in the CHARGE consortium (N=53949). *Mol Psychiatry* **20**, 183-92 (2015).
19. Hill, W.D. *et al.* Molecular genetic aetiology of general cognitive function is enriched in evolutionarily conserved regions. *Translational Psychiatry* **6**(2016).

20. Winkler, T.W. *et al.* Quality control and conduct of genome-wide association meta-analyses. *Nat Protoc* **9**, 1192-212 (2014).
21. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-1 (2010).
22. Bulik-Sullivan, B.K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* **47**, 291-5 (2015).
23. Yang, J., Lee, S.H., Goddard, M.E. & Visscher, P.M. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* **88**, 76-82 (2011).
24. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* **8**, 1826 (2017).
25. Boyle, A.P. *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Research* **22**, 1790-1797 (2012).
26. Kircher, M. *et al.* A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* **46**, 310 (2014).
27. Lonsdale, J. *et al.* The Genotype-Tissue Expression (GTEx) project. *Nature Genetics* **45**, 580-585 (2013).
28. Kettunen, J. *et al.* Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA. *Nat Commun* **7**, 11122 (2016).
29. Li, J. & Ji, L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity* **95**, 221-227 (2005).
30. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* **47**, 1236-41 (2015).
31. Zheng, J. *et al.* LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* **33**, 272-279 (2017).
32. Liu, J. *et al.* A Mendelian Randomization Study of Metabolite Profiles, Fasting Glucose, and Type 2 Diabetes. *Diabetes* **66**, 2915-2926 (2017).
33. Mitchel, M.W. *et al.* 17q12 Recurrent Deletion Syndrome. (1993).
34. Persson, B. *et al.* The SDR (short-chain dehydrogenase/reductase and related enzymes) nomenclature initiative. *Chem Biol Interact* **178**, 94-8 (2009).
35. Arlt, A. & Schafer, H. Role of the immediate early response 3 (IER3) gene in cellular stress response, inflammation and tumorigenesis. *Eur J Cell Biol* **90**, 545-52 (2011).
36. Shahid, M. *et al.* Impaired 3',5'-cyclic adenosine monophosphate-mediated signaling in immediate early responsive gene X-1-deficient vascular smooth muscle cells. *Hypertension* **56**, 705-12 (2010).
37. Ikeda, M. *et al.* A genome-wide association study identifies two novel susceptibility loci and trans population polygenicity associated with bipolar disorder. *Mol Psychiatry* **23**, 639-647 (2018).
38. Li, X. *et al.* Common variants in the CDH7 gene are associated with major depressive disorder in the Han Chinese population. *Behav Genet* **44**, 97-101 (2014).
39. Hawi, Z. *et al.* The role of cadherin genes in five major psychiatric disorders: A literature update. *Am J Med Genet B Neuropsychiatr Genet* **177**, 168-180 (2018).
40. Sacchetti, E. *et al.* The GRM7 gene, early response to risperidone, and schizophrenia: a genome-wide association study and a confirmatory pharmacogenetic analysis. *Pharmacogenomics J* **17**, 146-154 (2017).
41. Wu, G. *et al.* Central functions of neuropeptide Y in mood and anxiety disorders. *Expert Opin Ther Targets* **15**, 1317-31 (2011).

42. Savage, J.E. *et al.* Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat Genet* **50**, 912-919 (2018).
43. Jongkees, B.J., Hommel, B. & Colzato, L.S. People are different: tyrosine's modulating effect on cognitive control in healthy humans may depend on individual differences related to dopamine function. *Frontiers in Psychology* **5** (2014).
44. Jongkees, B.J., Hommel, B., Kuhn, S. & Colzato, L.S. Effect of tyrosine supplementation on clinical and healthy populations under stress or cognitive demands--A review. *J Psychiatr Res* **70**, 50-7 (2015).
45. Nieoullon, A. Dopamine and the regulation of cognition and attention. *Prog Neurobiol* **67**, 53-83 (2002).
46. Elias, M.F. *et al.* Chronic kidney disease, creatinine and cognitive functioning. *Nephrology Dialysis Transplantation* **24**, 2446-2452 (2009).
47. Seliger, S.L. *et al.* Moderate renal impairment and risk of dementia among older adults: The cardiovascular health cognition study. *Journal of the American Society of Nephrology* **15**, 1904-1911 (2004).
48. Connelly, M.A., Gruppen, E.G., Otvos, J.D. & Dullaart, R.P.F. Inflammatory glycoproteins in cardiometabolic disorders, autoimmune diseases and cancer. *Clin Chim Acta* **459**, 177-186 (2016).
49. Lawler, P.R. *et al.* Circulating N-Linked Glycoprotein Acetyls and Longitudinal Mortality Risk. *Circ Res* **118**, 1106-15 (2016).
50. Hottman, D.A., Chernick, D., Cheng, S., Wang, Z. & Li, L. HDL and cognition in neurodegenerative disorders. *Neurobiol Dis* **72 Pt A**, 22-36 (2014).
51. Wishart, D.S. *et al.* HMDB 3.0--The Human Metabolome Database in 2013. *Nucleic Acids Res* **41**, D801-7 (2013).

SUPPLEMENTARY MATERIAL

Supplementary Note

Supplementary Table 1. Descriptive statistics of study population.

Supplementary Table 2. Information on phenotype assessment.

Supplementary Table 3. Information on genotyping platforms and quality control.

Supplementary Table 4. Genomic inflation factor for each study.

Supplementary Table 5. The variants associated with general cognitive factor at genome-wide significance level.

Supplementary Table 6. Annotation of genome-wide significant variants in CHARGE (EA) + UKB meta-analysis.

Supplementary Table 7. Independent significant SNPs and their distribution across the genomic loci.

Supplementary Table 8. Independent significant SNPs after conditioning on previously reported genetic signal.

Supplementary Table 9. Annotation of genome-wide significant variants after conditioning on previously reported signals.

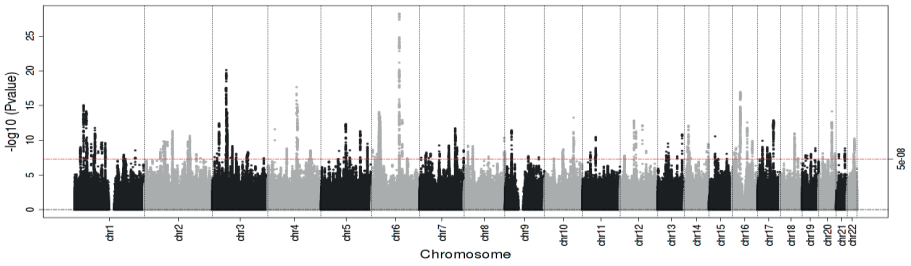
Supplementary Table 10. Independent significant SNPs and their distribution across the genomic loci after conditioning on previously reported genetic signal.

Supplementary Table 11. Overview of pleiotropic associations for novel independent significant genetic variants after conditional analysis and tagged SNPs. Only associations that passed genome-wide significance threshold were listed.

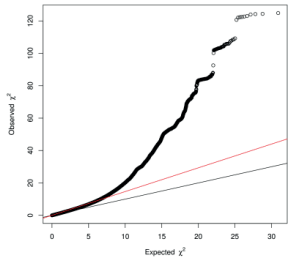
Supplementary Table 12. Genes to which independent genetic variants are mapped to and diseases they have been implicated in according to the DisGeNET database. Only association supported by at least two curated databases (Score > 0.2) are shown.

Supplementary Table 13. Independent genetic variants associated with general cognitive factor in the largest study of general cognitive ability to date and current study that were extracted from GWAS of circulating metabolites.

A)



B)



Supplementary Figure 1. (A) The results of genome-wide association meta-analysis including participants of EA in CHARGE cohorts and UK biobank. The x-axis represents chromosomes and y-axis –log10 *p*-values. Variants are represented by dots. The genome-wide significance threshold (*p*-value < 5×10^{-8}) is depicted by the red dashed line. (B) Quantile-quantile plot of the genetic variants associated with general cognitive function.