

## IMMEDIATE COMMUNICATION

Genetic contributions to variation in general cognitive function: a meta-analysis of genome-wide association studies in the CHARGE consortium ( $N = 53\,949$ )

G Davies<sup>1,2</sup>, N Armstrong<sup>3</sup>, JC Bis<sup>4</sup>, J Bressler<sup>5</sup>, V Chouraki<sup>6,7</sup>, S Giddaluru<sup>8,9</sup>, E Hofer<sup>10,11</sup>, CA Ibrahim-Verbaas<sup>12,13</sup>, M Kirin<sup>14</sup>, J Lahti<sup>15,16</sup>, SJ van der Lee<sup>13</sup>, S Le Hellard<sup>8,9</sup>, T Liu<sup>17,18</sup>, RE Marioni<sup>1,19,20</sup>, C Oldmeadow<sup>21</sup>, I Postmus<sup>22,23</sup>, AV Smith<sup>24,25</sup>, JA Smith<sup>26</sup>, A Thalamuthu<sup>27</sup>, R Thomson<sup>28</sup>, V Vitart<sup>29</sup>, J Wang<sup>30,31</sup>, L Yu<sup>32</sup>, L Zgaga<sup>33,34</sup>, W Zhao<sup>26</sup>, R Boxall<sup>19</sup>, SE Harris<sup>1,19</sup>, WD Hill<sup>2</sup>, DC Liewald<sup>1</sup>, M Luciano<sup>1,2</sup>, H Adams<sup>35,36</sup>, D Ames<sup>37,38</sup>, N Amin<sup>13,36</sup>, P Amouyel<sup>6</sup>, AA Assareh<sup>27</sup>, R Au<sup>7,30</sup>, JT Becker<sup>39,40,41</sup>, A Beiser<sup>7,30</sup>, C Berr<sup>42,43</sup>, L Bertram<sup>18,44</sup>, E Boerwinkle<sup>5,45,46</sup>, BM Buckley<sup>47</sup>, H Campbell<sup>14</sup>, J Corley<sup>2</sup>, PL De Jager<sup>48,49,50</sup>, C Dufouil<sup>51,52</sup>, JG Eriksson<sup>16,53,54,55</sup>, T Espeseth<sup>56,57</sup>, JD Faul<sup>58</sup>, I Ford<sup>59</sup>, Generation Scotland<sup>60</sup>, RF Gottesman<sup>61</sup>, ME Griswold<sup>62</sup>, V Gudnason<sup>24,25</sup>, TB Harris<sup>63</sup>, G Heiss<sup>64</sup>, A Hofman<sup>35,36</sup>, EG Holliday<sup>21</sup>, J Huffman<sup>29</sup>, SLR Kardia<sup>26</sup>, N Kochan<sup>27,65</sup>, DS Knopman<sup>66</sup>, JB Kwok<sup>67,68</sup>, J-C Lambert<sup>6</sup>, T Lee<sup>27,65</sup>, G Li<sup>4</sup>, S-C Li<sup>17,69</sup>, M Loitfelder<sup>10</sup>, OL Lopez<sup>39</sup>, AJ Lundervold<sup>70,71,72</sup>, A Lundqvist<sup>73</sup>, KA Mather<sup>27</sup>, SS Mirza<sup>35,36</sup>, L Nyberg<sup>73,74,75</sup>, BA Oostra<sup>13</sup>, A Palotie<sup>76,77,78</sup>, G Papenberg<sup>17,79</sup>, A Pattie<sup>2</sup>, K Petrovic<sup>10</sup>, O Polasek<sup>80</sup>, BM Psaty<sup>4,81,82,83</sup>, P Redmond<sup>2</sup>, S Reppermund<sup>27</sup>, JI Rotte<sup>84,85</sup>, H Schmidt<sup>10,86</sup>, M Schuur<sup>12,13</sup>, PW Schofield<sup>87</sup>, RJ Scott<sup>21</sup>, VM Steen<sup>8,9</sup>, DJ Stott<sup>88</sup>, JC van Swieten<sup>12</sup>, KD Taylor<sup>84,89</sup>, J Trollor<sup>27,90</sup>, S Trompet<sup>22,91</sup>, AG Uitterlinden<sup>35,36,92</sup>, G Weinstein<sup>7,30</sup>, E Widen<sup>77</sup>, BG Windham<sup>93</sup>, JW Jukema<sup>91,94,95</sup>, AF Wright<sup>29</sup>, MJ Wright<sup>96</sup>, Q Yang<sup>30,31</sup>, H Amieva<sup>52</sup>, JR Attia<sup>21</sup>, DA Bennett<sup>32</sup>, H Brodaty<sup>27,97</sup>, AJM de Craen<sup>22,23</sup>, C Hayward<sup>29</sup>, MA Ikram<sup>12,35,36,98</sup>, U Lindenberg<sup>17</sup>, L-G Nilsson<sup>99</sup>, DJ Porteous<sup>1,19,60</sup>, K Räikkönen<sup>15</sup>, I Reinvang<sup>57</sup>, I Rudan<sup>14</sup>, PS Sachdev<sup>27,65</sup>, R Schmidt<sup>10</sup>, PR Schofield<sup>100,101</sup>, V Srikanth<sup>28,102</sup>, JM Starr<sup>1,103</sup>, ST Turner<sup>104</sup>, DR Weir<sup>58</sup>, JF Wilson<sup>14</sup>, C van Duijn<sup>13,36</sup>, L Launer<sup>63</sup>, AL Fitzpatrick<sup>81,105</sup>, S Seshadri<sup>7,30</sup>, TH Mosley Jr<sup>93</sup> and IJ Deary<sup>1,2</sup>

<sup>1</sup>Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK; <sup>2</sup>Department of Psychology, University of Edinburgh, Edinburgh, UK; <sup>3</sup>School of Mathematics and Statistics, University of Sydney, Sydney, NSW, Australia; <sup>4</sup>Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA, USA; <sup>5</sup>Human Genetics Center, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX, USA; <sup>6</sup>Inserm-UMR744, Institut Pasteur de Lille, Unité d'Epidémiologie et de Santé Publique, Lille, France; <sup>7</sup>Department of Neurology, Boston University School of Medicine, Boston, MA, USA; <sup>8</sup>K.G. Jebsen Centre for Psychosis Research and the Norwegian Centre for Mental Disorders Research (NORMENT), Department of Clinical Science, University of Bergen, Bergen, Norway; <sup>9</sup>Dr Einar Martens Research Group for Biological Psychiatry, Center for Medical Genetics and Molecular Medicine, Haukeland University Hospital, Bergen, Norway; <sup>10</sup>Department of Neurology, Medical University of Graz, Graz, Austria; <sup>11</sup>Institute for Medical Informatics, Statistics and Documentation, Medical University of Graz, Graz, Austria; <sup>12</sup>Department of Neurology, Erasmus University Medical Center, Rotterdam, The Netherlands; <sup>13</sup>Genetic Epidemiology Unit, Department of Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands; <sup>14</sup>Centre for Population Health Sciences, University of Edinburgh, Edinburgh, UK; <sup>15</sup>Institute of Behavioural Sciences, University of Helsinki, Helsinki, Finland; <sup>16</sup>Folkhälsan Research Centre, Helsinki, Finland; <sup>17</sup>Max Planck Institute for Human Development, Berlin, Germany; <sup>18</sup>Max Planck Institute for Molecular Genetics, Berlin, Germany; <sup>19</sup>Medical Genetics Section, University of Edinburgh Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, Western General Hospital, Edinburgh, UK; <sup>20</sup>Queensland Brain Institute, The University of Queensland, Brisbane, QLD, Australia; <sup>21</sup>Hunter Medical Research Institute and Faculty of Health, University of Newcastle, Newcastle, NSW, Australia; <sup>22</sup>Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands; <sup>23</sup>Netherlands Consortium for Healthy Ageing, Leiden, The Netherlands; <sup>24</sup>Icelandic Heart Association, Kopavogur, Iceland; <sup>25</sup>University of Iceland, Reykjavik, Iceland; <sup>26</sup>Department of Epidemiology, University of Michigan, Ann Arbor, MI, USA; <sup>27</sup>Centre for Healthy Brain Ageing, School of Psychiatry, University of New South Wales, Sydney, NSW, Australia; <sup>28</sup>Menzies Research Institute, Hobart, Tasmania; <sup>29</sup>MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK; <sup>30</sup>Framingham Heart Study, Framingham, MA, USA; <sup>31</sup>Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA; <sup>32</sup>Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, IL, USA; <sup>33</sup>Department of Public Health and Primary Care, Trinity College Dublin, Dublin, Ireland; <sup>34</sup>Andrija Stampar School of Public Health, Medical School, University of Zagreb, Zagreb, Croatia; <sup>35</sup>Department of Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands; <sup>36</sup>Netherlands Consortium for Healthy Ageing, Leiden, The Netherlands; <sup>37</sup>National Ageing Research Institute, Royal Melbourne Hospital, Melbourne, VIC, Australia; <sup>38</sup>Academic Unit for Psychiatry of Old Age, St George's Hospital, University of Melbourne, Kew, Australia; <sup>39</sup>Department of Neurology, University of Pittsburgh, Pittsburgh, PA, USA; <sup>40</sup>Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA, USA; <sup>41</sup>Department of Psychology, University of Pittsburgh, Pittsburgh, PA, USA; <sup>42</sup>Inserm, U106, Montpellier, France; <sup>43</sup>Université Montpellier I, Montpellier, France; <sup>44</sup>Faculty of Medicine, School of Public Health, Imperial College, London, UK; <sup>45</sup>Brown Foundation Institute of Molecular Medicine for the Prevention of Human Diseases, University of Texas Health Science Center at Houston, Houston, TX, USA; <sup>46</sup>Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA; <sup>47</sup>Department of Pharmacology and Therapeutics, University College Cork, Cork, Ireland; <sup>48</sup>Program in Translational NeuroPsychiatric Genomics, Department of Neurology, Brigham and Women's Hospital, Boston, MA, USA; <sup>49</sup>Harvard Medical School, Boston, MA, USA; <sup>50</sup>Program in Medical and Population Genetics, Broad Institute, Cambridge, MA, USA; <sup>51</sup>Inserm U708, Neuroepidemiology, Paris, France; <sup>52</sup>Inserm U897, Université Bordeaux Segalen, Bordeaux, France; <sup>53</sup>National Institute for Health and Welfare, Helsinki, Finland; <sup>54</sup>Department of General Practice and Primary Health Care, University of Helsinki, Helsinki, Finland; <sup>55</sup>Unit of General Practice, Helsinki University Central Hospital, Helsinki, Finland; <sup>56</sup>K.G. Jebsen Centre for Psychosis Research, Norwegian Centre For Mental Disorders Research (NORMENT), Division of Mental Health and Addiction, Oslo University Hospital and Institute of Clinical Medicine, University of Oslo, Oslo, Norway; <sup>57</sup>Department of Psychology, University of Oslo, Oslo, Norway; <sup>58</sup>Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI, USA; <sup>59</sup>Robertson Center for Biostatistics, Glasgow, UK; <sup>60</sup>Generation Scotland, University of Edinburgh Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, Western General Hospital, Edinburgh, UK; <sup>61</sup>Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD, USA; <sup>62</sup>Center of Biostatistics and Bioinformatics, University of Mississippi Medical Center, Jackson, MS, USA; <sup>63</sup>Intramural Research Program National Institutes on Aging, National Institutes of Health, Bethesda, MD, USA; <sup>64</sup>Department of Epidemiology, University of North Carolina Gillings School of Global Public Health, Chapel Hill, NC, USA; <sup>65</sup>Neuropsychiatric Institute, Prince of Wales Hospital, Sydney, NSW, Australia; <sup>66</sup>Department of Neurology, Mayo Clinic, Rochester, MN, USA; <sup>67</sup>Neuroscience Research Australia, Randwick, NSW, Australia; <sup>68</sup>School of Medical Sciences, University of New South Wales, Sydney, NSW, Australia; <sup>69</sup>Technische Universität Dresden, Dresden, Germany; <sup>70</sup>Department of Biological and Medical Psychology, University of Bergen, Bergen, Norway; <sup>71</sup>Kavli Research Centre for Aging and Dementia, Haraldsplass Deaconess Hospital, Bergen, Norway; <sup>72</sup>K.G. Jebsen Centre for Research on Neuropsychiatric Disorders, University of Bergen, Bergen, Norway; <sup>73</sup>Umeå Center for Functional Brain Imaging (UFBI), Umeå University, Umeå, Sweden; <sup>74</sup>Department of Radiation Sciences, Umeå University, Umeå, Sweden; <sup>75</sup>Department of Integrative Medical Biology, Umeå University, Umeå, Sweden; <sup>76</sup>Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Cambridge, UK; <sup>77</sup>Institute for Molecular Medicine Finland (FIMM), University of Helsinki,

General cognitive function is substantially heritable across the human life course from adolescence to old age. We investigated the genetic contribution to variation in this important, health- and well-being-related trait in middle-aged and older adults. We conducted a meta-analysis of genome-wide association studies of 31 cohorts ( $N = 53\,949$ ) in which the participants had undertaken multiple, diverse cognitive tests. A general cognitive function phenotype was tested for, and created in each cohort by principal component analysis. We report 13 genome-wide significant single-nucleotide polymorphism (SNP) associations in three genomic regions, 6q16.1, 14q12 and 19q13.32 (best SNP and closest gene, respectively: rs10457441,  $P = 3.93 \times 10^{-9}$ , *MIR2113*; rs17522122,  $P = 2.55 \times 10^{-8}$ , *AKAP6*; rs10119,  $P = 5.67 \times 10^{-9}$ , *APOE/TOMM40*). We report one gene-based significant association with the *HMGN1* gene located on chromosome 21 ( $P = 1 \times 10^{-6}$ ). These genes have previously been associated with neuropsychiatric phenotypes. Meta-analysis results are consistent with a polygenic model of inheritance. To estimate SNP-based heritability, the genome-wide complex trait analysis procedure was applied to two large cohorts, the Atherosclerosis Risk in Communities Study ( $N = 6617$ ) and the Health and Retirement Study ( $N = 5976$ ). The proportion of phenotypic variation accounted for by all genotyped common SNPs was 29% (s.e. = 5%) and 28% (s.e. = 7%), respectively. Using polygenic prediction analysis, ~1.2% of the variance in general cognitive function was predicted in the Generation Scotland cohort ( $N = 5487$ ;  $P = 1.5 \times 10^{-17}$ ). In hypothesis-driven tests, there was significant association between general cognitive function and four genes previously associated with Alzheimer's disease: *TOMM40*, *APOE*, *ABCG1* and *MEF2C*.

*Molecular Psychiatry* (2015) **20**, 183–192; doi:10.1038/mp.2014.188; published online 3 February 2015

## INTRODUCTION

The systems with which humans face the challenges of the external and internal environments tend to show deterioration in their mean level of efficiency as people age. For example, the immune system,<sup>1</sup> cardiovascular<sup>2</sup> and respiratory systems,<sup>3</sup> renal function<sup>4</sup> and stress responses<sup>5</sup> are, on average, not as efficient in old age as they were in young adulthood. When met with a challenge to a system, healthy older people tend to have a reduced reserve capacity compared with younger adults. The nervous system also declines, with sensory function<sup>6,7</sup> and motor strength and co-ordination<sup>8</sup> not at their highest mean level in older age.

The brain's cognitive functions show a heterogeneity of age-related changes. Some capabilities such as vocabulary, some number skills and general knowledge withstand the aging trend and tend (in the absence of neurological disease) to remain intact into older age, although they decline eventually.<sup>9–11</sup> These cognitive functions, which rely on the access to stored information, are called crystallized abilities.<sup>12</sup>

By contrast, fluid cognitive abilities, which rely on on-the-spot information processing, show more age-related decline in mean levels.<sup>9–14</sup> Through fluid cognitive abilities, the brain equips us to, for example, recognize and recall previously unseen stimuli, and to make associations between previously unrelated stimuli (memory); induce general rules from sets of occurrences and to apply these rules to new situations (reasoning); perform simple,

repetitive cognitive tasks accurately and at high speed (processing speed); compute mental two- and three-dimensional transformations of objects and locations (spatial ability); and organize thinking (executive functioning). By comparison with crystallized abilities, tests of fluid cognitive abilities tend to employ unfamiliar and often abstract materials, and draw less on stored knowledge, education and broader enculturation. Each domain of fluid cognitive ability is important and is a major object of research in its own right.<sup>15,16</sup> However, when taken together the cognitive domains show an interesting regularity. If we apply a number of cognitive tests that assess diverse cognitive functions to a large and varied sample, the correlations are universally positive; people who tend to do well on one domain of fluid cognitive ability tend to do well in all of the others, although the associations are far from perfect.<sup>17</sup> This means that some of the observed inter-individual variation in any single cognitive domain can be attributed to at least four sources: how good they are generally at all fluid cognitive tests, how good they are at that fluid cognitive domain, how good they are at specific tests within any one domain, and error of measurement.<sup>18</sup>

In the present study, we focus on the variance that crosses a number of different fluid cognitive functions, that is, on general fluid cognitive ability. We examine the genetic contributions to this ability in middle and older age. There are good reasons for doing so. General cognitive ability remains a strong source of cognitive variation in older age, probably accounting for about as

Helsinki, Finland; <sup>78</sup>Department of Medical Genetics, University of Helsinki and University Central Hospital, Helsinki, Finland; <sup>79</sup>Karolinska Institutet, Aging Research Center, Stockholm University, Stockholm, Sweden; <sup>80</sup>Faculty of Medicine, Department of Public Health, University of Split, Split, Croatia; <sup>81</sup>Department of Epidemiology, University of Washington, Seattle, WA, USA; <sup>82</sup>Department of Health Services, University of Washington, Seattle, WA, USA; <sup>83</sup>Group Health Research Unit, Group Health Cooperative, Seattle, WA, USA; <sup>84</sup>Institute for Translational Genomics and Population Sciences Los Angeles BioMedical Research Institute, Harbor-UCLA Medical Center, Los Angeles, CA, USA; <sup>85</sup>Division of Genetic Outcomes, Department of Pediatrics, Harbor-UCLA Medical Center, Los Angeles, CA, USA; <sup>86</sup>Centre for Molecular Medicine, Institute of Molecular Biology and Biochemistry, Medical University of Graz, Graz, Austria; <sup>87</sup>School of Medicine and Public Health, University of Newcastle, Newcastle, NSW, Australia; <sup>88</sup>Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, UK; <sup>89</sup>Department of Pediatrics, Harbor-UCLA Medical Center, Los Angeles, CA, USA; <sup>90</sup>Department of Developmental Disability Neuropsychiatry, School of Psychiatry, University of New South Wales, Sydney, NSW, Australia; <sup>91</sup>Department of Cardiology, Leiden University Medical Center, Leiden, The Netherlands; <sup>92</sup>Department of Internal Medicine, Erasmus University Medical Center, Rotterdam, The Netherlands; <sup>93</sup>Division of Geriatrics, Department of Medicine, University of Mississippi Medical Center, Jackson, MS, USA; <sup>94</sup>Durrer Center for Cardiogenetic Research, Amsterdam, The Netherlands; <sup>95</sup>Interuniversity Cardiology Institute of the Netherlands, Utrecht, The Netherlands; <sup>96</sup>Neuroimaging Genetics Group, QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia; <sup>97</sup>Dementia Collaborative Research Centre, University of New South Wales, Sydney, NSW, Australia; <sup>98</sup>Department of Radiology, Erasmus University Medical Center, Rotterdam, The Netherlands; <sup>99</sup>ARC, Karolinska Institutet, Stockholm and UFBI, Umeå University, Umeå, Sweden; <sup>100</sup>Neuroscience Research Australia, Sydney, NSW, Australia; <sup>101</sup>Faculty of Medicine, University of New South Wales, Sydney, NSW, Australia; <sup>102</sup>Stroke and Ageing Research, Medicine, Southern Clinical School, Monash University, Melbourne, VIC, Australia; <sup>103</sup>Alzheimer Scotland Dementia Research Centre, University of Edinburgh, Edinburgh, UK; <sup>104</sup>Division of Nephrology and Hypertension, Department of Internal Medicine, Mayo Clinic, Rochester, MN, USA and <sup>105</sup>Department of Global Health, University of Washington, Seattle, WA, USA. Correspondence: Dr IJ Deary, Centre for Cognitive Ageing and Cognitive Epidemiology, Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, Scotland, UK.  
E-mail: i.deary@ed.ac.uk

Received 17 October 2014; revised 11 November 2014; accepted 24 November 2014; published online 3 February 2015

much cognitive variation as in young adulthood.<sup>19</sup> General cognitive function accounts for a substantial proportion of the age-related variance<sup>9,20</sup> and genetic variance<sup>21,22</sup> of people's cognitive test differences in middle and older age. Across adulthood, and especially in older age, lower fluid cognitive ability and greater decline in it across the life course are associated with earlier mortality.<sup>23–25</sup> It can also be argued that if mean levels of fluid cognitive tests show decline with age, then people's differences in cognitive functions become more important as they grow older, because the levels of cognitive functioning that remain in older age are nearer to the critical levels that are needed for everyday functions.<sup>26</sup>

To date, behavior genetic research—using twin, adoption and family designs—shows that general cognitive ability is substantially heritable across the life course, from late childhood to old age.<sup>22</sup> The heritability of general cognitive functioning in old age might decrease slightly from its levels in young and middle adulthood. Candidate gene studies have found that variation in *APOE* genotype is the only reliable individual genetic associate of cognitive function in older age, but that might apply especially to cognitive change rather than cognitive level in old age.<sup>27–29</sup> Using the genome-wide complex trait analysis procedure (GCTA), genome-wide association studies (GWAS) found that ~51% (the s.e. was large, at 11%) of the variation in general fluid cognitive function in late middle age and older age could be accounted for by genetic variation that is tagged by single-nucleotide polymorphisms (SNPs) on the Illumina610-QuadV1 chip.<sup>30</sup> That study was conducted in a total discovery sample of 3511 individuals, with replication in 670 independent individuals. It found no genome-wide significant single SNP associations. From other GWAS studies of complex traits, we now know that this sample size is likely to be too small, by an order of magnitude, to detect genome-wide significant SNPs.<sup>31</sup>

In summary, general fluid cognitive functioning is a key aspect of health in older age. It is important to identify the determinants of its individual differences, both environmental and genetic. To date, studies have been too small to detect the expected small genetic effects. Here we conduct a meta-analysis of GWAS studies of general fluid cognitive functioning in middle and older age from the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium, with a total of 53 949 individuals.

## MATERIALS AND METHODS

### Participants

This report includes individuals from 31 population-based cohorts participating in the Cohorts for Heart and Aging Research in Genomic Epidemiology consortium (Supplementary Table S1 and Supplementary Information 1 Section 1). All participants were of European ancestry and aged 45 years or older. Exclusion criteria included prevalent dementia and clinical stroke (including self-reported stroke). The total sample size was 53 949 individuals ( $N_{\text{men}} = 23\,030$ ,  $N_{\text{women}} = 30\,919$ ).

### General fluid cognitive function phenotype

For each of the cohorts, a general fluid cognitive function component phenotype was constructed from a number of cognitive tasks, each testing a different cognitive domain. In order to construct this measure, each cohort was required to have tasks that tested at least three different cognitive domains. Principal component analysis was applied to the cognitive task scores to derive a measure of general cognitive function, which was the score on the first unrotated principal component. Correlations between each test score and the general cognitive function score were calculated for each cohort, to confirm that all cohorts' general cognitive function scores had been scored in the required direction, that is, with a higher test score indicating higher cognitive function. Further details of the cognitive tasks undertaken and of the phenotype construction in each cohort are provided in Supplementary Information 1 Section 2. In summary, there was a clear single component accounting for between

33.7% and 62.3% (mean = 49.6%) of the total cognitive test variance in all cohorts.

Cohorts used different batteries of cognitive tests, which means that there will be phenotypic heterogeneity. However, it has been shown that the individual differences found on the general cognitive component derived from different cognitive test batteries are very similar.<sup>32,33</sup> Here we give an example of the similarity of the scores obtained when using two different sets of tests to derive the general cognitive ability component. This can be illustrated in the Lothian Birth Cohort 1936, because it has such a large battery of cognitive tests.<sup>34</sup> Two general fluid-type cognitive function component phenotypes could be derived, each using a different battery of cognitive tests. Of course, only one of these was used in the GWAS study. The first comprised six non-verbal tests from the Wechsler Adult Intelligence Scale-III UK; these were Block Design, Digit Symbol, Symbol Search, Letter-Number Sequencing, Backward Digit Span and Matrix Reasoning. The second contained the Moray House Test, Logical Memory, Spatial Span, Four Choice Reaction Time and Verbal Fluency. These two general cognitive function phenotypes calculated from two non-overlapping batteries of cognitive tests in the Lothian Birth Cohort 1936 had a correlation of  $r = 0.79$  ( $P < 0.001$ ). Thus, we use this as a methodological illustration to show that there is substantial overlap between the general fluid cognitive ability components from different sets of tests.

### Genotyping and quality control

Each cohort applied quality control (QC) measures based on SNP and sample-based missingness, minor allele frequency, Hardy-Weinberg equilibrium, relatedness and evidence of non-Caucasian descent. Cohort-specific thresholds for these QC measures along with details of genotyping, imputation methods and reference panels are detailed in Supplementary Table S2.

### Statistical analyses

*Genotype-phenotype association analyses.* Genotype-phenotype association analyses were performed, using an additive model, on imputed SNP dosage scores within each cohort (Supplementary Figures S1 and S2). Adjustments for age, sex and population stratification, if required, were included in the model. Cohort-specific covariates—for example, site or familial relationships—were also fitted as required. The cohort-specific association results were subjected to QC procedures before meta-analysis. SNPs were excluded based on imputation quality (IMPUTE info  $< 0.4$ , MACH  $r^2 < 0.3$ , BAMBAM  $r^2 < 0.3$ ) and minor allele frequency ( $< 0.5\%$ ). Only SNPs that passed these QC criteria in  $> 50\%$  of individuals were included in the meta-analysis (2 478 500 SNPs). A meta-analysis of the 31 cohorts was performed using the METAL package with an inverse variance weighted model implemented and single genomic control applied (<http://www.sph.umich.edu/csg/abecasis/Metal>). SNP-based results were also compared to published results for educational attainment<sup>35</sup> (Social Science Genetic Association Consortium) and general cognitive function in childhood<sup>36</sup> (Childhood Intelligence Consortium).

Gene-based tests of association were performed using the VEGAS software.<sup>37</sup> In addition, we examined the gene-based findings for association with published candidate genes previously associated with Alzheimer's disease (AD) or neuropathological features of AD and related dementias,<sup>38–44</sup> educational attainment<sup>35</sup> and general cognitive function in childhood.<sup>36</sup>

*Prediction analyses.* In order to perform prediction analyses in Generation Scotland (GS), a meta-analysis was performed, which excluded this cohort. A multi-SNP prediction model was created using the profile scoring method implemented in PLINK.<sup>45</sup> This uses the effect sizes estimated in the meta-analysis. The GS best-guess imputed data were used for this analysis, an imputation quality threshold of  $r^2 > 0.9$  was applied, and the remaining SNPs were pruned to remove those in linkage disequilibrium (based on  $r^2 > 0.25$  within a 200-SNP sliding window). The estimated effect sizes from the meta-analysis for each of these SNPs were then used to calculate a prediction score for each GS individual. A series of prediction scores was created based on the inclusion of SNPs with a range of association  $P$ -values: all SNPs and SNPs with  $P < 0.5$ ,  $P < 0.1$ ,  $P < 0.05$  or  $P < 0.01$ . Linear or logistic regressions of the prediction score and cognitive phenotypes, and some health outcomes previously associated with cognitive function<sup>46</sup> were performed. We calculated the proportion of phenotypic variance that was predicted by adding the prediction score to a 'null' model that

adjusted for age, sex and population stratification (four principal components). The cognitive phenotypes from GS that were included in the prediction analysis were general cognitive function, general fluid cognitive function, Wechsler Digit Symbol Substitution Task,<sup>47</sup> Wechsler Logical Memory Test,<sup>48</sup> Verbal Fluency<sup>49</sup> and the Mill Hill Vocabulary Scale (junior+senior synonyms).<sup>50</sup> The health outcomes included in the prediction analysis were self-reported cardiovascular disease, type 2 diabetes and hypertension. Polygenic prediction analyses were also performed in GS for cognitive phenotypes using published GWAS results for AD<sup>39</sup> (International Genomics of Alzheimer's Project) and educational attainment.<sup>35</sup>

**Estimation of SNP-based heritability using GCTA analysis.** The GCTA program<sup>51</sup> was used to estimate the proportion of variance explained by all common SNPs for general cognitive function in the Atherosclerosis Risk in Communities Study (ARIC) and Health and Retirement Study (HRS) cohorts. These cohorts were selected to be used for this analysis, because they are two of the largest cohorts in the study. The total numbers of individuals included in these analyses were 6617 for the ARIC cohort and 5976 for the HRS cohort. These totals differ from the single SNP analyses, because close relatives were excluded from this analysis. One individual was excluded from any pair of individuals that had an estimated coefficient of relatedness of  $> 0.025$  to ensure that effects due to shared environment were not included. The same covariates were included in the GCTA analyses as for the SNP-based association analyses.

**Pathway and functional genomic analyses.** The following pathway and functional genomic methods were performed: INRICH<sup>52</sup> and Ingenuity Pathway Analysis (Ingenuity Systems, www.ingenuity.com), and reference was made to The Genotype-Tissue Expression Portal (<http://www.gtexportal.org>), Regulome DB database<sup>53</sup> and the Human Brain Transcriptome Project ([hbatlas.org](http://hbatlas.org)).<sup>54</sup> Full details are given in Supplementary Methods.

## RESULTS

The SNP-based meta-analysis identified 13 SNPs associated with general cognitive function at a genome-wide significance level ( $P < 5 \times 10^{-8}$ ) (Figure 1, Figure 2 and Supplementary Figure S3). These SNPs were located in three genomic regions, 6q16.1, 14q12 and 19q13.32. The top SNP in each region and genes contained in the region were, 6q16.1, rs10457441 ( $P = 3.93 \times 10^{-9}$ ; *MIR2113*), 14q12, rs17522122 ( $P = 2.55 \times 10^{-8}$ ; *AKAP6/NPAS3*) and 19q13.32, rs10119 ( $P = 5.67 \times 10^{-9}$ ; *TOMM40/APOE*) (Figure 2). The effect size of rs10119 was significantly correlated with mean age of the cohort ( $r = -0.424$ ,  $P = 0.022$ ; Figure 3 and Supplementary Figure S4). There was no significant correlation with cohort age for the other two SNPs (rs10457441 and rs17522122) (Supplementary Figures S5 and S6). All 361 SNPs from the meta-analysis with a  $P$ -value less than a suggestive significance threshold of  $P < 1 \times 10^{-5}$  are detailed in Supplementary Table S3.

Gene-based tests of association yielded one genome-wide significant result ( $P < 2.8 \times 10^{-6}$ ), for the gene *HMG1* ( $P = 1 \times 10^{-6}$ ) located on chromosome 21. All 184 gene-based association results with a  $P$ -value below the suggestive threshold ( $P < 1 \times 10^{-3}$ ) are detailed in Supplementary Table S4.

Supplementary Table S5 shows the gene-based association results for 29 genes previously reported to be associated with AD or neuropathologic features of AD and related dementias.<sup>38–44</sup> Four of these genes, *TOMM40*, *APOE*, *MEF2C* and *ABCG1*, were associated with general fluid cognitive function at  $P < 0.01$ . Owing to linkage disequilibrium, the association  $P$ -values for *APOE* and *TOMM40* are not independent. *APOE* and *TOMM40* are located in the region of chromosome 19 identified by the SNP-based meta-analysis.

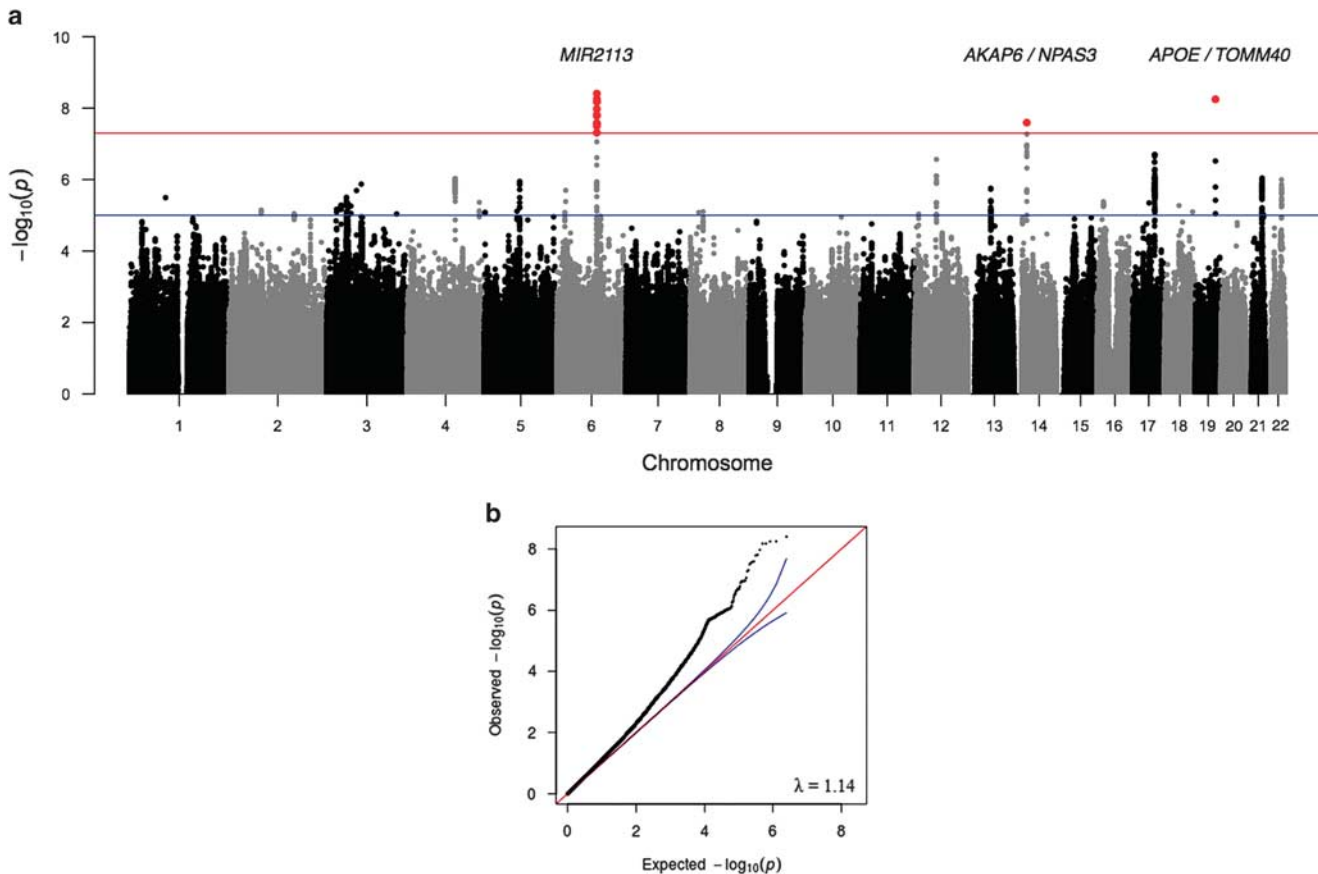
Supplementary Tables S6A and S7A show the top SNP-based results from published GWAS of educational attainment<sup>35</sup> and general cognitive function in childhood.<sup>36</sup> It should be noted that these are not independent studies due to sample overlap in some cohorts between the current study and both of these previously published studies (overlaps are: educational attainment

$N \sim 30\,000$ ; general cognitive function in childhood  $N \sim 1500$ ). Of the 361 suggestively significant SNPs from the current meta-analysis (Supplementary Table S3), 188 and 192 SNPs demonstrated nominal significance ( $P < 0.05$ ) with the educational attainment phenotypes of years of education and college completion, respectively (Supplementary Table S7B). Sixteen SNPs achieved  $P < 1 \times 10^{-6}$  in the educational attainment analyses<sup>35</sup>; six of these achieved nominal significance in the current meta-analysis (Supplementary Table S7A). Of the top 100 SNPs in the general cognitive function in childhood GWAS,<sup>36</sup> 11 reached a nominal level of significance in the current study (Supplementary Table S6A). Comparisons of gene-based results are shown in Supplementary Tables S6B, S7C and S7D. For the educational attainment phenotypes, 17 and 14 of the top 25 genes associated with years of education and college completion,<sup>35</sup> respectively, were nominally significant in the current gene-based results. For childhood general cognitive function,<sup>36</sup> eight of the top 20 gene-based findings achieve nominal significance in the current study.

Results from the meta-analysis are consistent with a polygenic model of inheritance. This is demonstrated by the clear early deviation from the null hypothesis observed in the QQ plot ( $\lambda = 1.14$ ) (Figure 1). To investigate further whether the general cognitive phenotype is under polygenic control as is indicated by the association meta-analyses, we applied the analysis within the GCTA set of methods<sup>51</sup> that estimates the proportion of phenotypic variance explained by all SNPs in two of the largest single cohorts: ARIC and HRS. For the ARIC and HRS cohorts, respectively, the GCTA method returned SNP heritability estimates of 0.29 (s.e. = 0.05;  $N = 6617$ ;  $P = 2.34 \times 10^{-9}$ ) and 0.28 (s.e. = 0.07;  $N = 5976$ ;  $P = 2.00 \times 10^{-5}$ ).

The results from the polygenic prediction analyses are shown in Supplementary Table S8. The maximum proportion of phenotypic variance explained in GS using the prediction set derived from the meta-analysis excluding GS was 1.27% ( $P = 1.5 \times 10^{-17}$ ) for general cognitive function when using the  $P < 0.50$  SNP set ( $N = 47\,322$ ). The proportion of variance explained in other cognitive domains ranged from near-zero values to about 1% (Supplementary Table S8). The polygenic score did not significantly predict cardiovascular disease, hypertension or type 2 diabetes in GS (all  $P > 0.01$ ). Supplementary Table S9 shows the results from the polygenic prediction using the published results for educational attainment<sup>35</sup> (years of education and completion of a college degree) and AD.<sup>39</sup> For educational attainment, the maximum proportion of phenotypic variance explained in GS for the general cognitive phenotype was 0.54% ( $P = 2.78 \times 10^{-8}$ ) when using the  $P < 0.50$  SNP set ( $N = 40\,239$ ). For AD, significant predictions were observed for only the more inclusive SNP sets ( $P < 0.50$  and  $P < 1$ ) and the maximum proportion of phenotypic variance explained was 0.19%. Owing to the interdependency of the cognitive function measures, health outcomes and polygenic scores, no correction for multiple testing was applied to the  $P$ -values presented.

Supplementary Table S10 lists Gene Ontology gene sets that showed nominal enrichment before correction for multiple testing in the INRICH analysis. None of the Gene Ontology gene sets remained significant after correction for multiple testing. The highest ranked 70-node network in the Ingenuity Pathway Analysis included 58 molecules with *CDK2* as a central hub (Supplementary Figure S7). The second-ranked 70-node network included 38 molecules and had three multi-connected hubs: *RHOA*, *NUPR1* and *SRF*. The highest ranked 140-node network includes 103 molecules and combines the top two 70-node networks (Supplementary Figure S8). The top function categories associated with this network are cell cycle, cell death and survival, and gene expression. The top canonical pathways were inositol pyrophosphates biosynthesis, tRNA charging, Ga12/13 signaling, IL-15 production and role of NFAT in regulation of the immune response.



**Figure 1.** Manhattan (a) and Q-Q (b) plots of  $P$ -values of the association between single-nucleotide polymorphisms (SNPs) and general cognitive function in the meta-analysis. The threshold for genome-wide significance ( $P < 5 \times 10^{-8}$ ) is indicated by the red line and the threshold for suggestive significance ( $P < 1 \times 10^{-5}$ ) is indicated by the blue line.

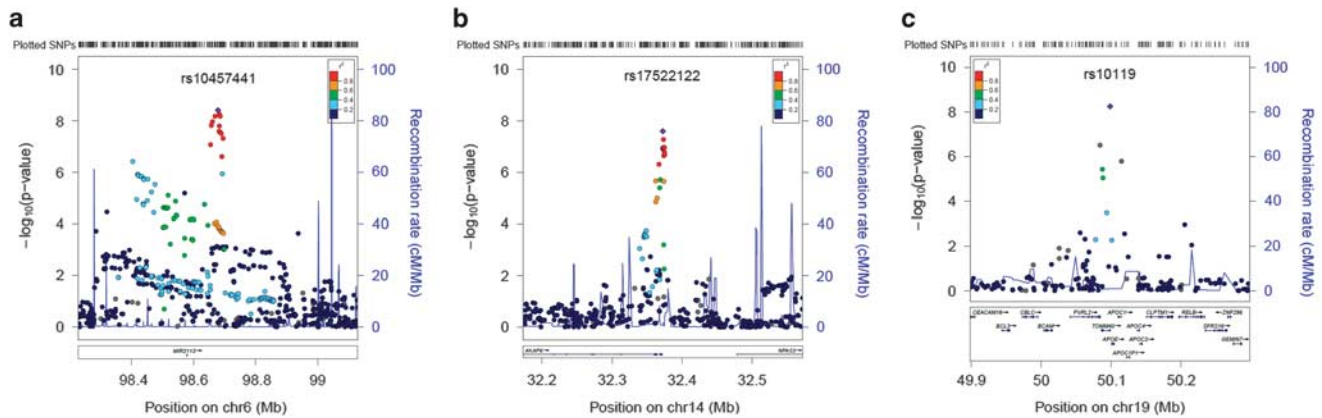
Using the Genotype-Tissue Expression Portal database (<http://www.gtexportal.org>), no *cis*-expression quantitative trait loci associations were identified for the 13 genome-wide significant SNPs. Supplementary Figure S9 shows differential expression of the top two genes from the VEGAS analysis, *HMGN1* and *BRWD1*, in six brain regions over the human lifecourse (<http://hbatlas.org/pages/hbtd>). Neither gene demonstrated differential expression across the brain regions shown in middle and older age. For this study, data mining of regulatory elements was restricted to normal brain relevant cell lines/tissues. Supplementary Table S11 demonstrates evidence of regulatory elements associated with 7 of the 13 genome-wide significant SNPs (<http://www.regulomedb.org/>).

## DISCUSSION

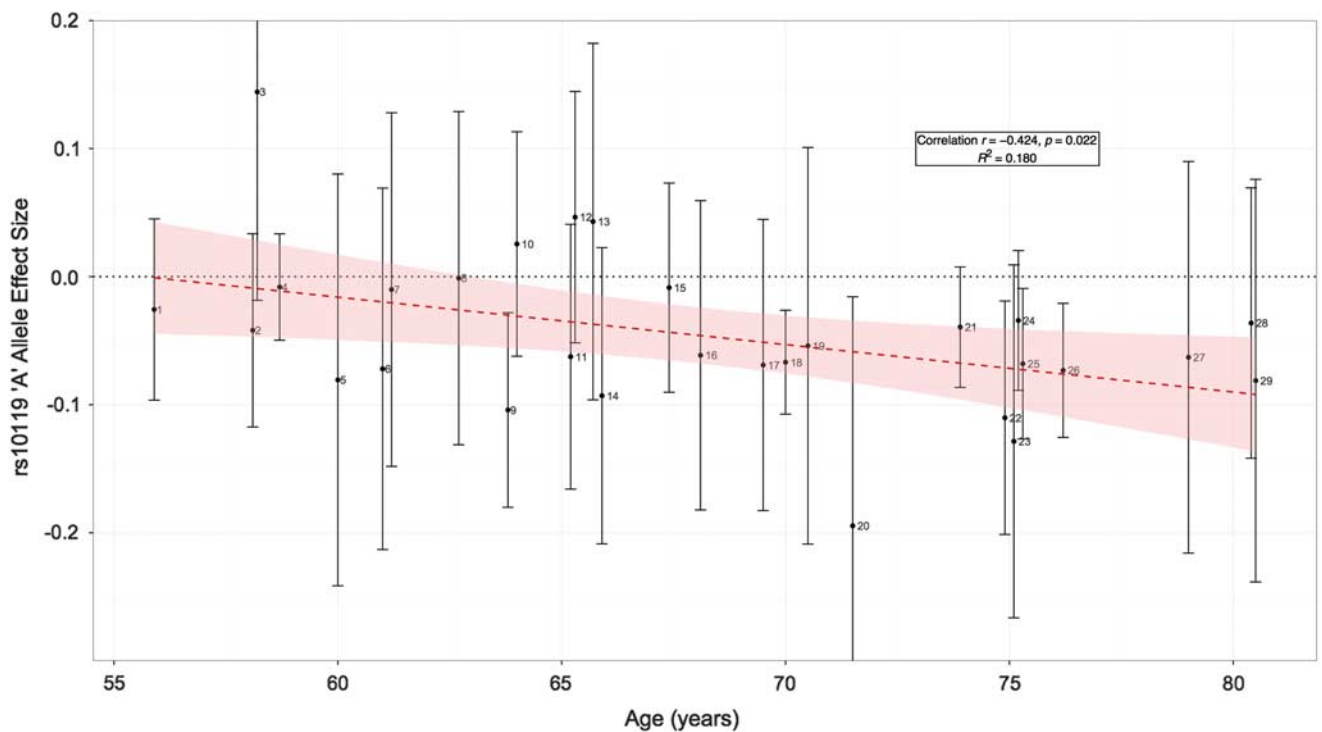
In this genome-wide association study of general cognitive function in middle and older age, with a total  $N$  of 53 949 participants, we report 13 genome-wide significant SNP-based associations in the three genomic regions 6q16.1, 14q12 and 19q13.32. There was one gene-based significant association with the *HMGN1* gene located on chromosome 21. We observed association of general cognitive function with four genes previously associated with AD or neuropathological features of AD and related dementias (*TOMM40*, *APOE*, *MEF2C* and *ABCG1*). The results from the meta-analysis are consistent with a polygenic model of inheritance, indicating that many variants of small effect contribute to the additive genetic influences on general cognitive function. Using GCTA, we present consistent estimates of the

lower bound of the narrow sense heritability of general fluid cognitive function, 0.29 and 0.28, from two of the largest cohorts (ARIC and HRS). We are able to predict, using only common SNP data to create a polygenic score,  $\sim 1.2\%$  of the variance in general cognitive function in an independent sample (GS). Pathway and network analyses did not produce significant findings. None of the 13 SNPs achieving genome-wide significance were coding variants. However, functional annotation provided evidence of regulatory elements for seven SNPs, suggesting that they might have a functional non-protein coding effect.

The 19q13.32 region, which includes the *APOE*  $\epsilon 2/3/4$  haplotype and was associated with general cognitive function in this study, has previously been associated with cognitive phenotypes in old age,<sup>55–58</sup> AD<sup>42,59–61</sup> and non-pathological cognitive aging.<sup>28,29</sup> Here we find that the *APOE/TOMM40* region is also associated with general cognitive function in middle and older age. The only published GWAS of general cognitive function in older age, to date, did not report any significant *APOE/TOMM40* findings.<sup>30</sup> From the data presented here, it is not possible to identify a single SNP or gene within this region that is driving the association, as it is a gene-dense region that is known to exhibit a strong pattern of linkage disequilibrium. Davies *et al.*<sup>28</sup> performed a fine-mapping analysis of this region, which indicated that the observed association with non-pathological cognitive aging was being driven by *APOE*-based variation.<sup>28</sup> A functional analysis of the *APOE* locus, including *TOMM40*, found that multiple *APOE* locus *cis*-regulatory elements influence both *APOE* and *TOMM40* promoter activity.<sup>62</sup> Functional annotation of the top SNP (rs101119) in the present study demonstrated evidence of



**Figure 2.** Regional association plots of genomic regions, which demonstrated genome-wide significance ( $P < 5 \times 10^{-8}$ ) in the meta-analysis, for chromosomes 6 (a), 14 (b) and 19 (c). The circles represent each genotyped single-nucleotide polymorphism (SNP), with the color indicating pairwise linkage disequilibrium (LD) in relation to the top hit (calculated from 1000 Genomes Nov 2010 EUR);  $-\log_{10} P$ -values are also indicated (y axis). The purple diamond represents the top SNP in each region. The solid blue line indicates the recombination rate.



**Figure 3.** Plot of effect size against mean age of cohort for rs10119 (top SNP *APOE/TOMM40* region). Each numbered point represents a cohort (1, RSIII; 2, ERF; 3, SPLIT; 4, GS; 5, KORCULA; 6, NCNG; 7, GENOA; 8, ORCADES; 9, RSI; 10, FHS; 11, ASP; 12, BASEI; 13, BETULA; 14, HCS; 15, RSII; 16, HBCS; 17, LBC1936; 18, HRS; 19, OATS; 20, TASCOG; 21, 3C; 22, PROSPER-Netherlands; 23, ROS; 24, PROSPER-Scotland; 25, PROSPER-Ireland; 26, AGES; 27, LBC1921; 28, CHS; 29, MAP). Two cohorts (ARIC and Sydney MAS) did not have data available for rs10119. Dashed regression line and shaded 95% confidence interval are shown. For full details of abbreviations, see Supplementary Information 2: Cohort Abbreviations.

regulatory elements, indicating regions of active transcription and epigenetic modification. All of these factors suggest that there is a complex transcriptional regulatory structure modulating regional gene expression at the *APOE/TOMM40* locus. Drawing on the large number of cohorts in the present paper, we found a significant correlation between mean cohort age and the effect size of rs10119 (located in the *APOE/TOMM40* region) on general cognitive function in older age. The effect was near to zero at younger mean ages and larger at older ages. This might help toward resolving some of the uncertainty about age moderation

of *APOE*  $\epsilon 4$  status associations with cognitive function that are based on single studies.<sup>63–65</sup>

The 14q12 region identified by the meta-analysis contains both the *AKAP6* and *NPAS3* genes. *AKAP6* (A kinase (PRKA) anchor protein 6) is highly expressed in various brain regions, and cardiac and skeletal muscle. It is specifically localized to the sarcoplasmic reticulum and nuclear membrane, and is involved in anchoring protein kinase A to the nuclear membrane or sarcoplasmic reticulum.<sup>66</sup> SNPs within this gene have tentatively been associated with several diseases/traits by GWAS: rs4296166 has been

associated with risk of AD at a suggestive level of significance,<sup>67</sup> rs2383378 was suggestively associated with anorexia nervosa,<sup>68</sup> rs2300835 was associated with fasting insulin levels in the discovery sample of a GWAS of fasting glycemic traits, but failed to replicate,<sup>69</sup> and rs1951681 and rs3784178 were suggestively associated with economic and political preferences (environmentalism).<sup>70</sup>

*NPAS3* (neuronal PAS domain protein 3) encodes a member of the basic helix-loop-helix and PAS domain-containing family of transcription factors that is expressed in the brain. The encoded protein is localized to the nucleus and may regulate genes involved in neurogenesis.<sup>71</sup> It has been associated with brain development and potentially human brain evolution.<sup>72</sup> A balanced reciprocal translocation t(9;14)(q34.2;q13) that disrupts *NPAS3* was identified in a mother and daughter with schizophrenia and schizophrenia co-morbid with mild learning disability, respectively.<sup>73,74</sup> A GWAS has since reported three coding SNPs (rs12434716, c.1654G>C,  $P=0.009$ ; rs10141940, c.2208C>T,  $P=0.01$ ; and rs10142034, c.2262C>G,  $P=0.01$ ) to be associated with schizophrenia.<sup>75</sup>

The genome-wide significant SNPs in the 6q16.1 region are located ~100 kb downstream of an uncharacterized microRNA gene *MIR2113*. This region has been previously associated with bipolar disorder (rs12202969,  $P=1.08 \times 10^{-8}$ )<sup>76</sup> and educational attainment.<sup>35</sup> Evidence of regulatory elements associated with six of the top SNPs in this region, within normal brain-related tissues and cell lines, was identified using the Regulome DB database. The regulatory elements identified include histone modifications, DNase hypersensitive sites and position weight matrix sites. This evidence suggests that the associated SNPs are in sites of active transcription and could have a regulatory role on transcription.

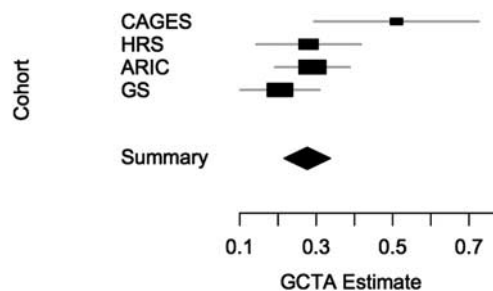
The significant finding from the gene-based analysis, *HMGN1* (high mobility group nucleosome-binding domain 1), encodes a nucleosome-binding protein that is associated with transcriptionally active chromatin.<sup>77</sup> *HMGN1* negatively regulates methyl CpG-binding protein 2 (*MeCP2*), a DNA-binding protein that is mutated in the neurodevelopmental disorder Rett syndrome<sup>78</sup> and is known to affect neurological functions including X-linked mental retardation, various autism spectrum disorders in humans and the behavior of mice. It is overexpressed in Down syndrome and it has been suggested that epigenetic changes resulting from altered *HMGN1* levels could have a role in the etiology of several neurodevelopmental disorders including autism and Down syndrome.<sup>79</sup>

Four of the 29 genes previously reported to be associated with AD<sup>38–42,44</sup> or neuropathological features of AD and related dementias<sup>43</sup> were associated with general cognitive function (at  $P < 0.01$ ). These were *TOMM40*, *APOE*, *MEF2C* and *ABCG1*. These results suggest that there is overlap between the genetic contribution to 'normal' and 'pathological' cognitive variation in older age. A polygenic prediction analysis using a large published GWAS of AD<sup>39</sup> significantly predicted only 0.19% of the phenotypic variation in general cognitive function in GS. However, Harris *et al.*<sup>80</sup> reported no significant association of polygenic score for AD with general cognitive ability when using smaller sample sizes for both the creation of the polygenic score and the prediction analysis. All known cases of clinical dementia were removed from the contributing cohorts. Of course, some or all of the effects we found could be driven by the inadvertent inclusion of individuals in a prodromal stage of dementia, and that we have picked up the genetic effects on this. This is an important issue that is impossible to eliminate entirely. Ideally, one would wish to know, prospectively, which people in the current cohorts eventually developed dementia and then re-run the analyses after omitting them. However, some people will die or be lost to contact before such an assessment could be made, thus preventing complete ascertainment. On the other hand, it is possible to envisage a study that tracks individuals and re-analyses those who have kept in contact and those who do not, say over a 10-year

period, develop dementia. This could clarify whether the effects we found here were on the normal range of age-related cognitive change. The present study includes individuals within and beyond the ninth decade. It is also important to note that the cognitive phenotype we measured is a composite of people's stable trait levels and any age-related change that has occurred. Therefore, genetic effects might be contributing to either.

The results of this study (SNP and gene-based) were compared with those of previously published large GWAS of educational attainment<sup>35</sup> and general cognitive function in childhood.<sup>36</sup> Before discussing these results further, it should be noted that there is sample overlap between both of these studies and the current study (overlaps are: educational attainment,  $N \sim 30\,000$ ; childhood general cognitive function,  $N \sim 1500$ ). Around 50% of the suggestive SNPs from the current study are nominally significant for educational attainment and the 6q16.1 region is reported to be genome-wide significant for both general cognitive function and educational attainment. The gene-based findings also demonstrated some consistency, with more than half of the top 25 genes for years of education and college completion achieving nominal significance in the current study. Of the top 100 SNPs for childhood cognitive function, 11 achieved nominal significance in the current study along with eight of the top 20 gene-based findings. These findings suggest that there is overlap between the genetic contribution to general cognitive function in late-middle and older age, and both educational attainment and childhood general cognitive function. This has also been explored in a study, which used education as a proxy phenotype for general cognitive function.<sup>81</sup> The bivariate heritability of educational attainment and general cognitive function has been previously estimated in GS using both pedigree-based and SNP-based methods (biv  $h^2=0.78$ ,  $N \sim 20\,500$  and biv  $h^2=0.59$ ,  $N \sim 6600$ , respectively).<sup>82</sup> The genetic correlation and bivariate heritability of childhood and older age general cognitive functions have also been previously estimated in a relatively small sample ( $N \sim 1900$ ), ( $r_g=0.62$ ; biv SNP  $h^2=0.21$ ).<sup>83</sup>

This study provides further evidence that general cognitive function is heritable and under polygenic control. These findings are consistent with, and add considerably to those from the Cognitive Ageing in Genetics in England and Scotland consortium.<sup>30</sup> The early deviation from the expected distribution in the QQ plot could be indicative of two outcomes; first, polygenic effects and, second, population stratification. All of the individual cohorts applied the required number of principal component analysis/multidimensional scaling components to their initial association analyses to adjust for stratification that may be present within the individual cohorts. None of the individual cohorts demonstrate both an early deviation and an inflated  $\lambda$ -value, which would indicate that stratification may be present (Supplementary Figure S2). Single genomic control was implemented in the meta-analysis. The proportion of phenotypic variance explained by all SNPs is an estimate of the lower bound of the narrow sense heritability. The estimates calculated from the ARIC and HRS cohorts suggest that 29% (s.e.=0.05) and 28% (s.e.=0.07), respectively, of the variation in general cognitive function can be attributed to common SNPs that are in linkage disequilibrium with causal variants in these cohorts. Whereas these estimates are lower than a previously-published estimate from the Cognitive Ageing in Genetics in England and Scotland consortium in a smaller sample (51%; s.e.=0.11;  $N=3511$ ),<sup>30</sup> they are slightly higher than an estimate reported from a similar sample size in GS (21%; s.e.=0.05;  $N=6648$ ).<sup>82</sup> To date, these are the only GCTA-based estimates of the general cognitive function phenotype in unrelated early-middle to older age individuals. The sample-size weighted mean of these estimates is 0.30 (Figure 4). It should be noted that these four estimates were not based on the same set of common SNPs, as only directly genotyped SNPs were included, and that the general cognitive function phenotypes in



**Figure 4.** Forest plot of four GCTA-based estimates for the single-nucleotide polymorphism (SNP)-based heritability ( $x$  axis) of general fluid cognitive function. The summary mean and s.e. were estimated using inverse-variance weighting. Abbreviations: ARIC, The Atherosclerosis Risk in Communities Study; CAGES, Cognitive Ageing Genetics in England and Scotland Consortium;<sup>30</sup> GS, Generation Scotland.<sup>78</sup> HRS, Health and Retirement Study.

each cohort were constructed using a different set of cognitive tests. Therefore, the mean of 0.30 is likely to be an underestimate.

Whereas the combined sample size in the present report is a strength, the results observed here suggest that an even larger sample size is required in order to seek replication of the findings of the present study and to identify more genome-wide significant findings. For this reason, we chose to present a single discovery meta-analysis of the largest combined sample available. If general cognitive function is similar to other complex traits, the individual effects of common SNPs will be very small. From studies of other polygenic complex traits, it has been observed that the number of discovered variants is strongly correlated with experimental sample size.<sup>31</sup> This predicted increase in detectable associations for complex traits, when sample sizes increase, has been observed in GWASs of both height—a study of 183 727 individuals reported 180 significant associations, of which >100 were novel loci compared with previous studies of fewer individuals ( $N < 40\,000$ )—and schizophrenia, in which a study of 21 246 cases and 38 072 controls reported 22 significant associations, of which 13 were novel loci compared with previous studies of fewer individuals ( $N < 18\,000$  cases).<sup>84,85</sup> This is also demonstrated in the present study when compared with the previously published Cognitive Ageing in Genetics in England and Scotland consortium study ( $N = 3511$ ), which reported no genome-wide significant SNP associations with general cognitive function in older age.<sup>30</sup>

Phenotypic heterogeneity is a limitation of this study. Each cohort used a different set of cognitive tests to create the general cognitive function phenotype. We demonstrated using the Lothian Birth Cohort 1936 cohort that two general cognitive function phenotypes calculated in the same set of individuals but using a different battery of tests for the principal component analysis are highly ( $\sim 0.8$ ) but imperfectly correlated. This is consistent with the finding that general cognitive function phenotypes derived from different test batteries are highly correlated and measurement of this phenotype is not dependent on the use of specific cognitive tasks.<sup>32,33</sup> Given, especially, that a relatively small number of tests contributed to the general cognitive component in some of the cohorts, the heterogeneity of the phenotype will be a limitation on the discovery of genetic effects.

In order to dissect the regions of association identified in the present study further, deep sequencing of these regions in a larger sample would be required followed by in-depth functional genomics studies. This type of approach may elucidate the mechanisms underlying the observed associations and identify the causal variants within these associated genomic regions.

In conclusion, we report the largest meta-analysis of GWAS studies, to date, of fluid general cognitive function in middle and older age. We also report results showing that general cognitive

function is heritable and highly polygenic, extending findings of previous studies involving general cognitive function in older individuals. We show genome-wide significant SNP-based associations within three genomic regions 6q16.1 (*MIR2113*), 14q12 (*AKAP6/NPAS3* region) and 19q13.32 (*TOMM40/APOE* region), and a genome-wide significant gene-based association with the *HMG1* gene located on chromosome 21. The 19q13.32 region has long been associated with AD and more recently was associated with non-pathological cognitive aging,<sup>28</sup> the 6q16.1, 14q12 and *HMG1* regions contain genes associated with development of the brain,<sup>78,79</sup> neurological function,<sup>71,72,79</sup> psychiatric disease<sup>73–76</sup> and educational attainment.<sup>35</sup>

## CONFLICT OF INTEREST

Bruce M Psaty serves on a DSMB for a clinical trial of a device funded by the manufacturer (Zoll LifeCor). All other authors declare no conflict of interest.

## ACKNOWLEDGMENTS

We thank Peter M Visscher for advice on statistical genetic analyses and for comments on earlier versions of the manuscript, and Stuart J Ritchie for assistance with figures. For cohort acknowledgements, see Supplementary Information 1 Section 3.

## REFERENCES

- Vallejo AN. Immunological hurdles of ageing: indispensable research of the human model. *Ageing Res Rev* 2011; **10**: 315–318.
- Laurent S. Defining vascular aging and cardiovascular risk. *J Hypertens* 2012; **30**: S3–S8.
- Vaz Fragoso A, Gill TM. Respiratory impairment and the aging lung: a novel paradigm for assessing pulmonary function. *J Gerontol Ser A Biol Sci Med Sci* 2012; **67**: 264–275.
- Bolignano D, Mattace-Raso F, Sijbrands EJG, Zoccali C. The aging kidney revisited: a systematic review. *Ageing Res Rev* 2014; **14**: 65–80.
- Ferrari E, Magri F. Role of neuroendocrine pathways in cognitive decline during aging. *Ageing Res Rev* 2008; **7**: 225–233.
- Lindenberger U, Baltes PB. Sensory functioning and intelligence in old age: a strong connection. *Psychol Aging* 1994; **9**: 339–355.
- Pathai S, Shiels PG, Lawn SD, Cook C, Gilbert C. The eye as a model of ageing in translational research: molecular, epigenetic and clinical aspects. *Ageing Res Rev* 2013; **12**: 490–508.
- Hardy R, Cooper R, Aihie Sayer A, Ben-Shlomo Y, Cooper C, Deary IJ et al. Body mass index, muscle strength, and physical performance in older adults from eight cohort studies: the HALCYON programme. *PLoS One* 2013; **8**: e56483.
- Salthouse T. Localizing age-related individual differences in a hierarchical structure. *Intelligence* 2004; **32**: 541–561.
- Schaie KW. *Developmental Influences on Adult Intelligence*. Oxford University Press: New York, USA, 2005.
- Finkel D, Reynolds CA, McArdle JJ, Pedersen NL. Cohort differences in trajectories of cognitive ageing. *J Gerontol Ser B Psychol Sci Soc Sci* 2007; **62B**: P286–P294.
- Horn JL. Theory of fluid and crystallised intelligence. In: Sternberg RJ (ed). *Encyclopedia of Human Intelligence*. Macmillan: New York, NY, USA, 1994.
- Salthouse T. When does age-related cognitive decline begin? *Neurobiol Aging* 2009; **30**: 507–514.
- Josefsson M, de Luna X, Pudas S, Nilsson L-G, Nyberg L. Genetic and lifestyle predictors of 15-year longitudinal change in episodic memory. *JAGS* 2012; **60**: 2308–2312.
- Craik FIM, Salthouse TA. *The Handbook of Aging and Cognition* 3rd Edition, Psychology Press: New York, USA, 2008.
- Hofer SM, Alwin DF. *Handbook of Cognitive Aging*. Sage: Los Angeles, USA, 2008.
- Carroll JB. *Human Cognitive Abilities: A Survey of Factor-Analytic Studies*. Cambridge University Press: Cambridge, UK, 1993.
- Deary IJ, Penke L, Johnson W. The neuroscience of human intelligence differences. *Nat Rev Neurosci* 2010; **11**: 201–211.
- Tucker-Drob E. Differentiation of cognitive abilities across the life span. *Dev Psychol* 2009; **45**: 1097–1118.
- Tucker-Drob E. Global and domain-specific changes in cognition throughout adulthood. *Dev Psychol* 2011; **47**: 331–343.
- Petrill SA, Plomin R, Berg S, Johansson B, Pedersen NL, Ahern F et al. The genetic and environmental relationship between general and specific cognitive abilities in twins age 80 and older. *Psychol Sci* 1998; **9**: 183–189.



- 22 Deary IJ, Johnson W, Houlihan L. Genetic foundations of human intelligence. *Hum Genet* 2009; **126**: 215–232.
- 23 Murray C, Pattie A, Starr JM, Deary IJ. Does cognitive ability predict mortality in the ninth decade? The Lothian Birth Cohort 1921. *Intelligence* 2012; **40**: 490–498.
- 24 Shipley BA, Der G, Taylor MD, Deary IJ. Association between mortality and cognitive change over 7 years in a large representative sample of UK residents. *Psychosom Med* 2007; **69**: 640–650.
- 25 Calvin CM, Deary IJ, Fenton C, Roberts B, Der G, Leckenby N *et al*. Intelligence in youth and all-cause mortality: systematic review with meta-analysis. *Int J Epidemiol* 2011; **40**: 626–644.
- 26 Tucker-Drob E. Neurocognitive functions and everyday functions change together in old age. *Neuropsychology* 2011; **25**: 368–377.
- 27 Payton A. The impact of genetic research on our understanding of normal cognitive ageing: 1995 to 2009. *Neuropsychol Rev* 2009; **19**: 451–477.
- 28 Davies G, Harris SE, Reynolds CA, Payton A, Knight HM, Liewald DC *et al*. A genome-wide association study implicates the *APOE* locus in non-pathological cognitive ageing. *Mol Psychiatr* 2014; **19**: 76–87.
- 29 Zhang C, Perce BL. Genetic susceptibility to accelerated cognitive decline in the US Health and Retirement Study. *Neurobiol Aging* 2013; **35**: e18.
- 30 Davies G, Tenesa A, Payton A, Yang J, Harris SE, Liewald DC *et al*. Genome-wide association studies establish that human intelligence is highly heritable and polygenic. *Mol Psychiatr* 2011; **16**: 996–1005.
- 31 Visscher PM, Brown MA, McCarthy ML, Yang J. Five years of GWAS discovery. *Am J Hum Genet* 2012; **90**: 7–24.
- 32 Johnson W, Bouchard TJ, Krueger RF, McGue M, Gottesman II. Just one g: consistent results from three test batteries. *Intelligence* 2004; **32**: 95–107.
- 33 Johnson W, te Nijenhuis J, Bouchard T Jr. Still just 1 g: Consistent results from five test batteries. *Intelligence* 2008; **36**: 81–95.
- 34 Deary IJ, Gow AJ, Taylor MD, Corley J, Brett C, Wilson V *et al*. The Lothian Birth Cohort 1936: a study to examine influences on cognitive ageing from age 11 to age 70 and beyond. *BMC Geriatr* 2007; **7**: 28.
- 35 Rietveld CA, Medland SE, Derringer J, Yang J, Esko T, Martin NW *et al*. GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science* 2013; **340**: 1467–1471.
- 36 Benyamin B, Pourcain B, Davis OS, Davies G, Hansell NK, Brion MJ *et al*. Childhood intelligence is heritable, highly polygenic and associated with *FBNP1L*. *Mol Psychiatry* 2014; **19**: 253–258.
- 37 Liu JZ, McRae AF, Nyholt DR, Medland SE, Wray NR, Brown KM *et al*. A versatile gene-based test for genome-wide association studies. *Am J Hum Genet* 2010; **87**: 139–145.
- 38 Lambert JC, Heath S, Even G, Campion D, Sleegers K, Hiltunen M *et al*. Genome-wide association study identifies variants at *CLU* and *CR1* associated with Alzheimer's disease. *Nat Genet* 2009; **41**: 1094–1099.
- 39 Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C *et al*. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* 2013; **45**: 1452–1458.
- 40 Seshadri S, Fitzpatrick AL, Ikram MA, DeStefano AL, Gudnason V, Boada M *et al*. Genome-wide analysis of genetic loci associated with Alzheimer disease. *JAMA* 2010; **303**: 1832–1840.
- 41 Hollingworth P, Harold D, Sims R, Gerrish A, Lambert JC, Carrasquillo MM *et al*. Common variants at *ABCA7*, *MS4A6A/MS4A4E*, *EPHA1*, *CD33* and *CD2AP* are associated with Alzheimer's disease. *Nat Genet* 2011; **43**: 429–435.
- 42 Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Hamshere ML *et al*. Genome-wide association study identifies variants at *CLU* and *PICALM* associated with Alzheimer's disease. *Nat Genet* 2009; **41**: 1088–1093.
- 43 Beecham GW, Hamilton K, Naj AC, Martin ER, Huentelman M, Myers AJ *et al*. Genome-wide association meta-analysis of neuropathologic features of Alzheimer's disease and related dementias. *PLoS Genet* 2014; **10**: e1004606.
- 44 Bertram L, Lange C, Mullin K, Parkinson M, Hsiao M, Hogan MF *et al*. Genome-wide association analysis reveals putative Alzheimer's disease susceptibility loci in addition to *APOE*. *Am J Hum Genet* 2008; **83**: 623–632.
- 45 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D *et al*. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–575.
- 46 Deary IJ, Weiss A, Batty GD. 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. *Psychol Sci Publ Interest* 2010; **11**: 53–79.
- 47 Wechsler D. *WAIS-III UK Administration and Scoring Manual*. Psychological Corporation: London, UK, 1998.
- 48 Wechsler D. *WMS-III UK Administration and Scoring Manual*. Psychological Corporation: London, UK, 1998.
- 49 Lezak M. *Neuropsychological Assessment*. Oxford University Press: Oxford, UK, 1995.
- 50 Raven JC, Court JH, Raven J. *Manual for Raven's Progressive Matrices and Vocabulary Scales*. Lewis HK: London, UK, 1977.
- 51 Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR *et al*. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* 2010; **42**: 565–569.
- 52 Lee PH, O'Dushlaine C, Thomas B, Purcell SM. INRICH: interval-based enrichment analysis for genome-wide association studies. *Bioinformatics* 2012; **28**: 2.
- 53 Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M *et al*. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* 2012; **22**: 1790–1797.
- 54 Kang HJ, Kawasawa YI, Cheng F, Zhu Y, Xu X, Li M *et al*. Spatio-temporal transcriptome of the human brain. *Nature* 2011; **478**: 483–489.
- 55 Wisdom NM, Callahan JL, Hawkins KA. The effects of apolipoprotein E on non-impaired cognitive functioning: a meta-analysis. *Neurobiol Aging* 2011; **32**: 63–74.
- 56 De Jager PL, Shulman JM, Chibnik LB, Keenan BT, Raj T, Wilson RS *et al*. A genome-wide scan for common variants affecting the rate of age-related cognitive decline. *Neurobiol Aging* 2011; **33**: 15.
- 57 Small BJ, Rosnick CB, Fratiglioni L, Backman L, Apolipoprotein E and cognitive performance: a meta-analysis. *Psychol Aging* 2004; **19**: 592–600.
- 58 Deary IJ, Whiteman MC, Pattie A, Starr JM, Hayward C, Wright AF *et al*. Cognitive change and the *APOE* e4 allele. *Nature* 2002; **418**: 932.
- 59 Roses AD, Lutz MW, Amrine-Madsen H, Saunders AM, Crenshaw DG, Sundseth SS *et al*. A *TOMM40* variable-length polymorphism predicts the age of late-onset Alzheimer's disease. *Pharmacogenomics J* 2009; **10**: 375–384.
- 60 Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GA *et al*. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* 1993; **261**: 921–923.
- 61 Strittmatter WJ, Saunders AM, Schmechel D, Pericak-Vance M, Enghild J, Salvesen GS *et al*. Apolipoprotein E: high avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer's disease. *Proc Natl Acad Sci USA* 1993; **90**: 1977–1981.
- 62 Bekris LM, Lutz F, Yu C. Functional analysis of *APOE* locus genetic variation implicates regional enhancers in the regulation of both *TOMM40* and *APOE*. *J Hum Genet* 2011; **57**: 18–25.
- 63 Bunce D, Bielak AAM, Anstey KJ, Cherbuin N, Batterham PJ, Easteal S. *APOE* genotype and cognitive change in young, middle-aged and older adults living in the community. *J Gerontol A Biol Sci Med Sci* 2013; **69**: 379–386.
- 64 Nilsson LG, Adolfsen R, Bäckman L, Cruts M, Nyberg L, Small BJ *et al*. The influence of *APOE* status on episodic and semantic memory: data from a population-based study. *Neuropsychology* 2006; **20**: 645–657.
- 65 Blair CK, Folsom AR, Knopman DS, Bray MS, Mosley TH, Boerwinkle E. *APOE* genotype and cognitive decline in a middle-aged cohort. *Neurology* 2005; **64**: 268–276.
- 66 Kapiloff MS, Schillace RV, Westphal AM, Scott JD. mAKAP: an A-kinase anchoring protein targeted to the nuclear membrane of differentiated myocytes. *J Cell Sci* 1999; **112**: 2725–2736.
- 67 Seshadri S, Fitzpatrick AL, Ikram MA, DeStefano AL, Gudnason V, Boada M *et al*. Genome-wide analysis of genetic loci associated with Alzheimer disease. *JAMA* 2010; **303**: 1832–1840.
- 68 Wang K, Zhang H, Bloss CS, Duvvuri V, Kaye W, Schork NJ *et al*. A genome-wide association study on common SNPs and rare CNVs in anorexia nervosa. *Mol Psychiatry* 2011; **16**: 949–959.
- 69 Manning AK, Hivert MF, Scott RA, Grimsby JL, Bouatia-Naji N, Chen H *et al*. A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nat Genet* 2012; **44**: 659–669.
- 70 Benjamin DJ, Cesarini D, van der Loos MJHM, Dawes CT, Koellinger PD, Magnusson PKE *et al*. The genetic architecture of economic and political preferences. *Proc Natl Acad Sci USA* 2012; **109**: 8026–8031.
- 71 Ledent V, Paquet O, Vervoort M. Phylogenetic analysis of the human basic helix-loop-helix proteins. *Genome Biol* 2002; **3**: RESEARCH0030.
- 72 Kamm GB, Pisciotto F, Klinger R, Franchini LF. The developmental brain gene *NPAS3* contains the largest number of accelerated regulatory sequences in the human genome. *Mol Biol Evol* 2013; **30**: 1088–1102.
- 73 Kamnasaran D, Muir WJ, Ferguson-Smith MA, Cox DW. Disruption of the neuronal *PAS3* gene in a family affected with schizophrenia. *J Med Genet* 2003; **40**: 325–332.
- 74 Pickard BS, Malloy MP, Porteous DJ, Blackwood DH, Muir WJ. Disruption of a brain transcription factor, *NPAS3*, is associated with schizophrenia and learning disability. *Am J Med Genet B Neuropsychiatr Genet* 2005; **136B**: 26–32.
- 75 Macintyre G, Alford T, Xiong L, Rouleau GA, Tibbo PG, Cox DW. Association of *NPAS3* exonic variation with schizophrenia. *Schizophr Res* 2010; **120**: 143–149.
- 76 Mühleisen TW, Leber M, Schulze TG, Strohmaier J, Degenhardt F, Treutlein J *et al*. Genome-wide association study reveals two new risk loci for bipolar disorder. *Nat Commun* 2014; **11**: 3339.

- 77 Postnikov Y, Bustin M. Regulation of chromatin structure and function by *HMGN* proteins. *Biochim Biophys Acta* 2010; **1799**: 62–68.
- 78 Amir R, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. Rett syndrome is caused by mutations in X-linked *MECP2*, encoding methyl-CpG-binding protein 2. *Nat Genet* 1999; **23**: 185–188.
- 79 Abuhatzira L, Shamir A, Schones DE, Schäffer AA, Bustin M. The chromatin-binding protein *HMGN1* regulates the expression of methyl CpG-binding protein 2 (*MECP2*) and affects the behavior of mice. *J Biol Chem* 2011; **286**: 42051–42062.
- 80 Harris SE, Davies G, Luciano M, Payton A, Fox HC, Haggarty P *et al*. Polygenic risk for Alzheimer's disease is not associated with cognitive ability or cognitive aging in non-demented older people. *J Alzheimers Dis* 2014; **39**: 565–574.
- 81 Rietveld CA, Esko T, Davies G, Pers TH, Turley P, Benyamin B *et al*. Common genetic variants associated with cognitive performance identified using the proxy-phenotype method. *Proc Natl Acad Sci USA* 2014; **111**: 13790–13794.
- 82 Marioni RE, Davies G, Hayward C, Liewald D, Kerr SM, Campbell A *et al*. Molecular genetic contributions to socioeconomic status and intelligence. *Intelligence* 2014; **44**: 26–32.
- 83 Deary IJ, Yang J, Davies G, Harris SE, Tenesa A, Liewald D *et al*. Genetic contributions to stability and change in intelligence from childhood to old age. *Nature* 2012; **482**: 212–215.
- 84 Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F *et al*. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 2010; **467**: 832–838.
- 85 Ripke S, O'Dushlaine C, Chambert K, Moran JL, Kähler AK, Akterin S *et al*. Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat Genet* 2013; **45**: 1150–1159.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

Supplementary Information accompanies the paper on the Molecular Psychiatry website (<http://www.nature.com/mp>)