University of Basel

Why complex human phenotypes need complex data analytics - insights from the fields of molecular and cognitive neuroscience

A cumulative dissertation

Submitted to the Faculty of Psychology, University of Basel,
in partial fulfillment of the requirements for the degree of Doctor of Philosophy

by

Dr. med. Dipl.-Psych. Annette Milnik

from Aschaffenburg, Germany

Basel, Switzerland

January 2016

First-Supervisor: Prof. Dr. med. Andreas Papassotiropoulos

Second-Supervisor: Prof. Dr. med. Dominique J-F. de Quervain

University of Basel

Approved by the Faculty of Psychology
at the request of
Prof. Dr. med. Andreas Papassotiropoulos
Prof. Dr. med. Dominique J.-F. de Quervain

Basel, the 26th of January 2016

Prof. Dr. phil. Roselind Lieb

Abstract

Epidemiological research investigates the natural occurring variation of complex traits and the covariance between these traits in the general population. By doing so, epidemiological research is an important tool to understand influential factors on complex traits such as neuropsychiatric diseases and related phenotypes. However, epidemiological studies are challenged by interpretational difficulties and are often limited to inferential data analysis especially when based on a cross-sectional design. Different strategies exist to optimize the impact generated by such inferential data analyses. One strategy is to increase the depth of information by adding intermediate related traits, which is especially done in the field of genetics. However, complex covariance pattern typically underlie the relation between e.g. genotype, intermediate phenotype and primary phenotype of interest, which have to be resolved. In this situation, more complex analytical strategies might help to identify the most plausible model of relationship. The downside of these more comprehensive analytical models lies in the increase of model complexity, that might result in a less stable outcome. Finding a good balance between model complexity and analytical simplicity is a major challenge when performing combined analyses with several complex phenotypes. In the current thesis I presented three different works dealing with complex analytical strategies. The main goal behind all three of them was not to build up comprehensive theoretical frameworks, but to perform more preparatory analytical steps. The meta-analysis validated and extended a genetic association finding of the single nucleotide polymorphism rs17070145 with human memory performance by accumulating information of about 6 years of research performed worldwide. The heritability analysis verified that a common SNP-chip array is an appropriate dataset to perform more complex genetic analysis with human working memory performance measurements. The analysis of common epigenetic variation validates the DNA CpG methylation dataset in the context of complex analyses in mentally healthy young adults. Additionally, when comparing the three analyses, they also shed light on the varying complexity of putative intermediate phenotypes in human research. This knowledge can be used to build up comprehensive theoretical models and complex statistical analyses that combine several complex phenotypes.

Table of contents

Figure index

Acknowledgments

This thesis represents an important milestone in my professional and scientific life. I would like to thank all people that helped me reaching this destination. First of all I want to thank Andreas Papassotiropoulos and Dominique de Quervain for their guidance through the professional and scientific jungle the last 5 years. Special thanks to Andreas for helping me to prioritize if things start to grow over my head and if my professional direction has been fuzzy, and special thanks to Dominique for helping me to seek for simplicity, which appears to be the basis of a good scientific story. I also want to thank all my colleagues for their scientific spirit, which is a suitable mixture of collaborative work and scientific discourse that surrounds me every day. I would like to thank Christian Vogler and Vanja Vukojevic for the many fruitful discussion that accompanied this thesis. Angela Heck I would like to thank for her clear view that brings things down to earth if it is needed. Tobias Egli and Eva Loos I want to thank for their immediate support, if things need to be done. David Coynel I want to thank for his background support, that is just there when needed most. Last but not least I want to thank all team members of the departments of cognitive and molecular neuroscience for their daily effort to get the best out of our data, on every single level of the scientific knowledge building process; it is an honor for me to be part of this team.

Abbreviations

CpG	C-phosphate-G
DNA	Deoxyribonucleic acid
h^2	Heritability
LD	Linkage disequilibrium
miRNA	Micro ribonucleic acid
mRNA	Messenger ribonucleic acid
SNP	Single nucleotide polymorphism

1 Introduction

Members of the same species show considerable variability in their appearance and behavior. This variability can be due to differences in environmental factors, genetic factors, as well as interactions within and between both of them. By randomly selecting samples from a population, one can determine the distribution of these traits, which is often conformable with a Gaussian distribution (Fisher, 1918; Lyon, 2013). This implies that we are dealing with complex traits and that the observed variability is based on a larger number of independent influential factors (for a putative example see Figure 1). Epidemiological research investigates the natural occurring variation of complex traits and the covariance between these traits in the general population, and is mainly based on observations (Potter, 2001).

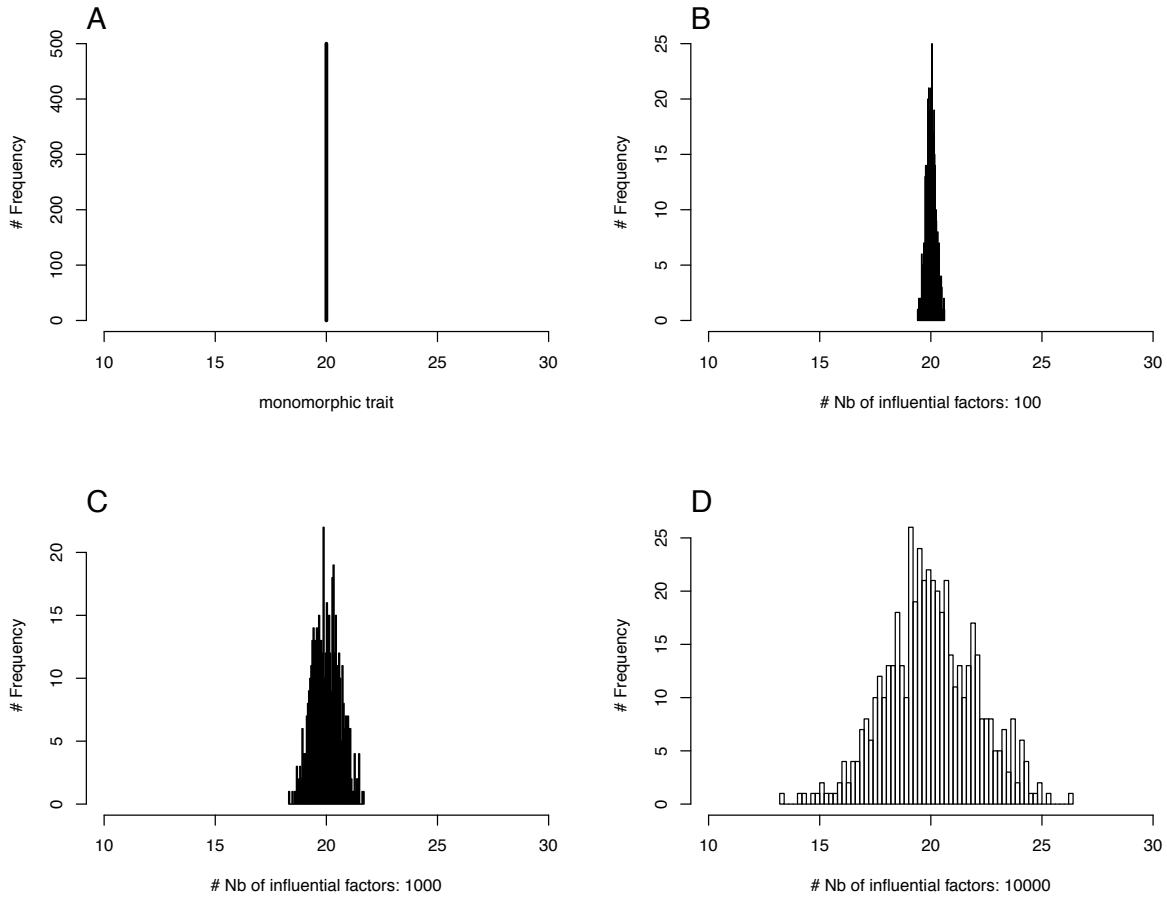


Figure 1 Starting with a monomorphic trait in a group of 500 individuals **(A)**, influential factors with an average effect-size of *r* = 0.01 were added with increasing numbers **(B-D)**. An effect-size or *r* = 0.01 is considered as small, but realistic e.g. for genetic association studies (Lango Allen et al., 2010; Ripke et al., 2014).

In contrast to that experimental research typically applies an experimental manipulation that might introduce additional variation, which does not necessarily appear under natural conditions (for a putative example see Figure 2). This is typically done in animal research by systematically changing environmental or genetic factors. Examples are e.g. studies that investigate the cognitive development of animals in an enriched in comparison to a non-enriched environment. Another example is investigating the impact of a given gene by experimentally knocking out the gene or suppressing the gene function.

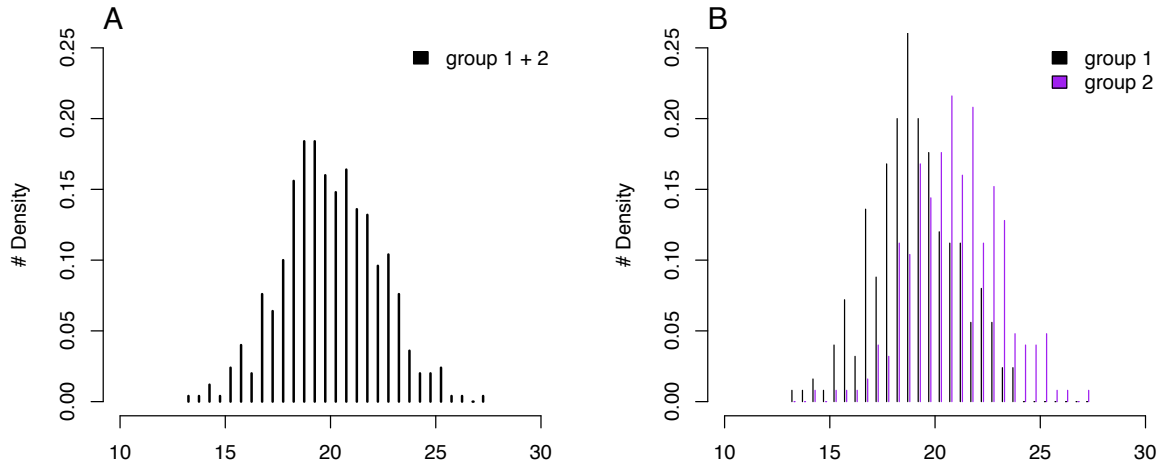


Figure 2 Based on the data from Figure 1D, an additional binary grouping factor is added. The difference between the two groups corresponds to a large effect size. In **(A)** the distribution of the total sample is depicted. In **(B)** the distribution is depicted separately for the two groups.

With respect to experimental manipulation, human research is obviously limited in its possibilities. Hence, epidemiological studies focusing on the variability and covariance observed under natural conditions are an important method to gain further insight into trait variability. Understanding these factors may allow shaping traits in a desired way in daily life or may allow preventing disease (Besen & Gan, 2014; Schacter, 2013). As a concrete example, psychiatric disorders are complex traits (Gelernter, 2015; Preston & Weinberger, 2005; Ripke et al., 2014), for which we still lack sufficient pharmacological treatment (Papassotiropoulos & de Quervain, 2015). Understanding the influential factors of complex psychiatric disorders and related traits may allow finding appropriate treatment strategies (Gottesman & Shields, 1973;

Gottesman, 2003; Papassotiropoulos & de Quervain, 2011; Papassotiropoulos et al., 2013). However, epidemiological studies are challenged by interpretational difficulties and are often limited to inferential data analysis (Leek & Peng, 2015) especially when based on a cross-sectional design. Here, the existence of covariance rarely allows to draw conclusions regarding the direction of effect or with respect to the exact nature of the covariance (Ioannidis, 2008; Potter, 2001; Smith et al., 2007; van der Sijde, Ng, & Fu, 2014), not even in the case of genetic association studies (van der Sijde et al., 2014).

Different strategies exist to optimize the impact generated by such inferential data analyses. One possibility is to apply data-mining techniques such as factor-analysis, cluster-analysis or multivariate linear modeling to identify more general and robust communalities and covariance patterns. Another approach is to add external annotation information to further weight and stratify the obtained association results. This is typically done in genetic association analyses, when it is possible to add functional information to the most-promising genetic findings. A third strategy is to increase the depth of information by adding intermediate related traits (Gottesman, 2003; Papassotiropoulos & de Quervain, 2015; van der Sijde et al., 2014). As an example, when investigating memory performance, it is possible to add information from functional brain imaging data that has been recorded while performing the task. All these strategies aim at generating testable hypotheses for future studies. The gold standard is to investigate these newly generated hypotheses by using different research designs, such as longitudinal studies and studies emphasizing on experimental designs with humans and animals (Kramer, 1998), in order to validate the importance of the initial finding.

To summarize, all the strategies mentioned above point to the fact that simple data analytical approaches might not be sufficient when investigating complex phenotypes in humans. Instead, more complex analytical strategies in combination with targeted longitudinal studies, experimental designs and animal research are the key factors to develop an in-depth understanding of the nature of complex traits (Mayr, 1997). This doctoral thesis aims at contributing to the fields of molecular and cognitive neuroscience by investigating the natural occurring variation of complex phenotypes in healthy young subjects. The phenotypes under investigation were human memory performance and human DNA (deoxyribonucleic acid) CpG (C-

phosphate-G) methylation derived from blood, which are both related to neuropsychiatric disorders. The following analytical approaches have been used: a meta-analysis, a heritability estimation based on unrelated individuals and a large-scale replication analysis. All three analyses are per se complex analytical strategies. In the scope of this thesis they were mainly done as independent validation steps before applying comprehensive modeling strategies that combine several complex phenotypes.

This thesis is based on the following three publications. The letters indicate my contributions to each publication and are listed after each reference: **A** - Designed the experiment; **B** - Performed the experiment; **C** - Analyzed the data or contributed to the analysis; **D** - Wrote the paper or contributed to paper writing.

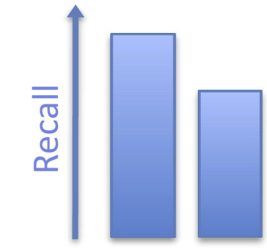
- Milnik, A., Heck, A., Vogler, C., Heinze, H.-J., de Quervain, D.J.-F., Papassotiropoulos, A., 2012. Association of KIBRA with episodic and working memory: a meta-analysis. Am. J. Med. Genet. B. Neuropsychiatr. Genet. 159B, 958–69. doi:10.1002/ajmg.b.32101 **(A-D)**
- Vogler, C., Gschwind, L., Coynel, D., Freytag, V., Milnik, A., Egli, T., Heck, A., de Quervain, D.J.-F., Papassotiropoulos, A., 2014. Substantial SNP-based heritability estimates for working memory performance. Transl. Psychiatry 4, e438. doi:10.1038/tp.2014.81 **(C-D)**
- Milnik, A., Vogler, C., Demougin, P., Egli, T., Freytag, V., Heck, A., Peter, F., Spalek, K., de Quervain, D.J.-F., Papassotiropoulos, A., Vukojevic, V., *in preparation*. Common epigenetic variation in a European population of mentally healthy young adults. **(A-D)**

2 Theoretical Background

2.1 Complex data analytics with complex phenotypes

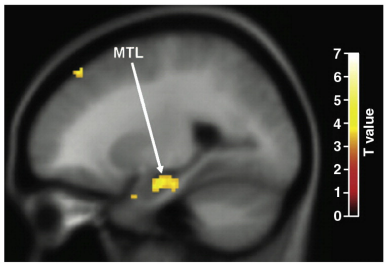
Endophenotype, intermediate phenotype and biomarker are related theoretical constructs used to describe the relationship between genetic variation and complex traits (Papassotiropoulos & de Quervain, 2015; van der Sijde et al., 2014). The terminology endophenotype was introduced by Gottesman and Shields in the seventies (Gottesman & Shields, 1973) and further adapted to the field of psychiatry (Gottesman, 2003; Papassotiropoulos & de Quervain, 2015). In the field of neuropsychiatric diseases, a suitable cognitive endophenotype has to fulfill several criteria: It has to be disturbed in diseases, it has to show a neuronal correlate that can be investigated further and it has to be heritable (Papassotiropoulos & de Quervain, 2015). Furthermore, it is thought that the complexity of an intermediate phenotype is lower than the complexity of the disease itself and that it is positioned along the causative pathway between genotype and disease (Mark & Toulopoulou, 2015; Preston & Weinberger, 2005; Rasch, Papassotiropoulos, & de Quervain, 2010; Toulopoulou et al., 2015) (see Figure 3).

Typically, complex covariance patterns underlie the relation between genotype, intermediate phenotype and primary phenotype of interest (Rasch et al., 2010; Relton & Davey Smith, 2010; Smith et al., 2007), which have to be resolved. Applying more complex analytical approaches might help to identify the most-plausible relationship (Schadt et al., 2005). However, comprehensive analytical models can also result in a less-well to understand outcome as well as in a reduced model stability, just because of the complexity of the model (Blalock, 1966; Kreft, Kreft, & de Leeuw, 1998). Finding a good balance between model complexity and analytical simplicity is a major challenge when it comes to suitable analytical strategies that combine several complex phenotypes.



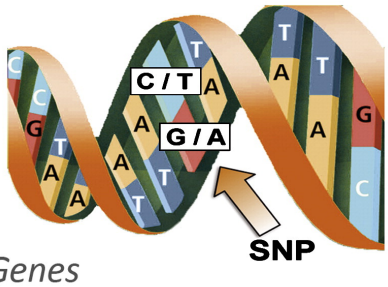
Memory performance

> 100 subjects needed to reliably detect significant genotype-dependent differences in memory performance[(b)]



Brain circuits

20-60 subjects needed to reliably detect significant genotype-dependent differences in brain activity[(a)]



Genes

Naturally occuring genetic variations, e.g., single nucleotide polymorphisms (SNPs)

Figure 3 Levels of analysis in genetic studies of episodic memory. At the level of genes, subjects are genotyped with regard to naturally occurring genetic variations in the human genome. On the level of brain circuits, neuroimaging is used to examine genotype-dependent differences in brain activity or functional coupling between brain regions (Rasch et al., 2010).[1] In this schema, the brain circuit is an intermediate phenotype with respect to memory performance, whereas memory performance is also a putative endophenotype with respect to neuropsychiatric diseases. The schema further evaluates possible sample-sizes that are necessary for genetic association studies. In principle, the less complex a phenotype is, the smaller the necessary sample-size should be.

Simplifying the research design is an appropriate strategy to circumvent overwhelmingly complex analytical situations. One possibility is to keep factors constant that are known to add variability to the dataset. An example might come from the field of neuropsychiatric diseases. Here, complex dependencies exist between the disease status, brain functioning and cognitive endophenotypes

[1] Reprinted from NeuroImage, 53(3), Fig. 3, Rasch, B. et. al., Imaging genetics of cognitive functions: Focus on episodic memory, 870–877, 2010, with permission from Elsevier.

(Toulopoulou et al., 2015). Instead of investigating the association between disease status, intermediate phenotypes and genetic background at once in a population that includes affected individuals, it is also possible to choose a less complex study population: by focusing on healthy young subjects one avoids the confounding factor of disease status. The findings derived from healthy young subjects can be used to further test them for associations with disease status in independent populations (Heck et al., 2014, 2015). With this stepwise approach, one avoids overly complex dependencies between the related phenotypes within one study population.

Another strategy is to carefully investigate and validate the input variables. This may e.g. allow reducing the complexity of the dataset by applying techniques of dimensionality reduction or by focusing on the most-promising phenotypes. Especially for comprehensive modeling strategies it is necessary to know the expected effect sizes and the biologically or environmentally driven variation in a given dataset. This allows estimating necessary sample-sizes as well as choosing appropriate analytical models. The here reported analyses were done to investigate and validate the complex phenotypes human memory performance and DNA CpG methylation before performing combined analyses and comprehensive modeling approaches.

2.2 Human memory performance

Human memory performance is a complex behavioral phenotype (Cahill, McGaugh, & Weinberger, 2001; Papassotiropoulos & de Quervain, 2011). Several classification approaches exist to group distinct memory processes in homogenous entities, based on observations from lesion-studies or based on functional classifications (Eustache & Desgranges, 2008; Henke, 2010). Memory processes can be broadly distinguished in working memory and episodic memory. Both characterize quite different cognitive functions used in everyday life. Working memory describes a system that temporarily stores transitory information for manipulation (Miller, 2013), whereas episodic memory describes the explicit memory of autobiographical events (Ferbinteanu, Kennedy, & Shapiro, 2006; Tulving, 2002). With these two definitions for working and episodic memory in mind, we used a rather simple classification to assign tasks to one of the two categories, based on the timing of stimulus presentation and the amount of distraction during the tasks: (1) the more time elapses between the encoding and the retrieval of the stimuli and (2) the more cognitive distraction takes place in between the two, the lower is the working memory component of the task. The N-Back task represents a typical and widely used working-memory task, in which subjects have to keep a stream of items (e.g. numbers or letters) in mind and have to indicate items that have been already presented one, two or three items before. A typical episodic memory task is to ask participants to encode pictorial or verbal material and to freely recall this material after a break of several minutes or hours.

To fulfill the criteria of a suitable cognitive endophenotype, memory performance has to be heritable, has to be impaired in disease and has to show a neuronal correlate that can be further investigated (Papassotiropoulos & de Quervain, 2015). For both memory domains, working and episodic memory it could be shown that they are heritable traits (h^2 = 0.15 – 0.74) (Ando, Ono, & Wright, 2001; Calkins et al., 2010; Greenwood et al., 2011; Lee et al., 2012; Vinkhuyzen, van der Sluis, Boomsma, de Geus, & Posthuma, 2010; Volk, McDermott, Roediger, & Todd, 2006; Wilson et al., 2011). The heritability estimates are comparable for men and women and are relatively stable across different age ranges of adults (Finkel & McGue, 1998; Finkel, Pedersen, & McGue, 1995; Johansson et al., 1999; Read et al., 2006).

Neuropsychiatric diseases typically go along with an impairment in cognitive functioning such as working and episodic memory (Barch & Smith, 2008; Elvevåg & Goldberg, 2000). A deficit in episodic memory performance is one of the key symptoms of dementia; however, impairments in episodic and working memory can also be observed in a variety of other neuropsychiatric diseases such as schizophrenia (Forbes, Carrick, McIntosh, & Lawrie, 2009; Mesholam-Gately, Giuliano, Goff, Faraone, & Seidman, 2009; Van Snellenberg, 2009), depression (Drevets, Price, & Furey, 2008; Marazziti, Consoli, Picchetti, Carlini, & Faravelli, 2010) and posttraumatic stress disorder (Samuelson, 2011). Especially for schizophrenia, there is evidence that also unaffected relatives show impairments in cognitive functioning (Conklin, Curtis, Calkins, & Iacono, 2005; Egan et al., 2001; Thompson, Watson, Steinhauer, Goldstein, & Pogue-Geile, 2005).

For both episodic and working memory well described neuroanatomical networks are known, with the medial temporal lobe region (including the hippocampus) as key anatomical structure for episodic memory (Eichenbaum, 2013), and a frontoparietal network as key anatomical structure for working memory (Owen, McMillan, Laird, & Bullmore, 2005). Findings from several studies suggest that also shared neuroanatomical networks exist for these two domains (Bird & Burgess, 2008; Lückmann, Jacobs, & Sack, 2014). Neuropsychiatric diseases and decline in cognitive functioning typically go along with an impairment of these key neuroanatomical regions (Drevets et al., 2008; Persson et al., 2012; Sperling et al., 2010).

Within the scope of this thesis, the meta-analysis verified *KIBRA* as a promising target gene for human episodic memory performance, and resulted 1) in realistic effect size estimates for future studies and 2) provided a hint that the association between the investigated genetic variant rs17070145 and memory performance is not exclusive for episodic memory performance. The heritability estimation 1) validated that human working memory performance measured with the N-Back task is a suitable phenotype for genetic association studies and 2) delivered realistic heritability estimates for the share of heritability that is captured by common single-nucleotide polymorphisms (SNPs).

2.3 DNA methylation

Epigenetic regulation is known to be centrally involved in trait variability and disease etiology (Petronis, 2010). The DNA sequence itself as blueprint of the genetic code is relatively uniform and fixed for nearly all cells of an individual. Epigenetic mechanisms serve as filter for the genetic information to achieve differentiation on a phenotypic level between cells and over time. Adrian Bird defined epigenetics as "the structural adaptation of chromosomal regions so as to register, signal or perpetuate altered activity states" (Bird, 2007). Hence, from an intermediate phenotype perspective epigenetic modification is positioned between the genetic code and the phenotypic outcome and is also physically located very closely to the DNA sequence (see Figure 4). Despite this close connection to the DNA (Schalkwyk et al., 2010; Shoemaker, Deng, Wang, & Zhang, 2010), the epigenetic inheritance is less stable than the genetic counterpart and hence allows more rapid adaptation to the environment (Sollars et al., 2003).

DNA CpG methylation is one epigenetic mechanism that can be investigated at both high resolution and throughput (Bibikova et al., 2011). Epigenome-wide association studies allow to investigate more in detail the link between structural genetic variation and complex phenotypes (Tan, Christiansen, von Bornemann Hjelmborg, & Christensen, 2015). However, due to the temporal dynamics of the signal, the tissue specificity and the reactivity to environmental changes (Bock, 2009; Ladd-Acosta et al., 2015; Lister et al., 2013; Spiers et al., 2015), the complexity of the CpG methylation signal is a challenge from both the analytical and the interpretational perspective (Bock, 2012). This is especially the case for studies investigating neuropsychiatric brain diseases or human cognition. These studies typically have to rely on blood as proxy tissue instead of directly investigating brain tissue (Hannon, Lunnon, Schalkwyk, & Mill, 2015).

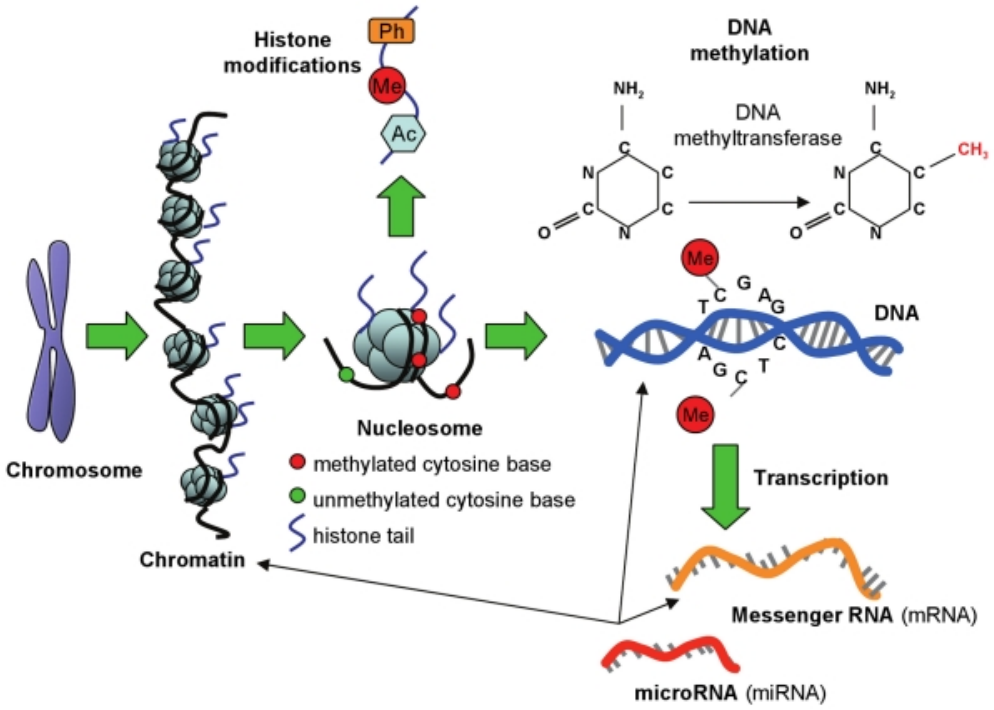


Figure 4 Depicted are the epigenetic mechanisms histone modification and DNA methylation that are involved in the regulation of transcription from DNA to messenger RNA (mRNA). The mRNA gets translated into a protein product. Another regulator of the mRNA translation is microRNA (miRNA) that might also interact with DNA methylation and histone modification (text is adapted from Relton & Davey Smith, 2010). The graphic is taken from Relton & Davey Smith, 2010.

There is evidence that epigenetic modification is involved in several neuropsychiatric diseases such as Alzheimer's disease or schizophrenia (Lardenoije et al., 2015; Lunnon et al., 2014; Pidsley et al., 2014). Especially for the glucocorticoid receptor, there is convincing evidence that stress leads to an epigenetic modification of the receptor and that this is linked to stress-related neuropsychiatric diseases (Bockmühl et al., 2015; Jawahar, Murgatroyd, Harrison, & Baune, 2015; Palma-Gudiel, Córdova-Palomera, Eixarch, Deuschle, & Fañanás, 2015; Vukojevic et al., 2014). Another important feature of the CpG methylation signal is its robust association with age (Hannum et al., 2013; Horvath, 2013; Horvath et al., 2012). Both, the stress and age related alterations in CpG methylation point to the fact that the epigenome might be able to bridge the gap between the fixed genetic blueprint and the variability in cognition and disease status over time.

However, when investigating healthy young adults, we are focused on a rather homogenous and healthy population with a narrow age-range. Within this homogenous population it is important to evaluate in how far the signal derived from DNA CpG methylation still comprises biologically or environmentally driven variability that can be used for analyses with further complex phenotypes. Replication analyses are a suitable tool to differentiate between the biologically and environmentally driven variability (summarized as common variance), technical variability and random noise. Hence, before associating the CpG methylation signal with further complex phenotypes, like memory performance, it is necessary to investigate and validate the signal itself. Within the scope of this thesis, the analyses of the methylation data provided realistic estimates of the expected common variation of each methylation site in a population of healthy young adults. This information can be used e.g. to preselect promising methylation sites for further association studies.

3 Methods

3.1 Meta-analysis

The issues of a suitable sample size and the lack of replication in molecular and cognitive neuroscience have been widely discussed (Button et al., 2013; IntHout, Ioannidis, Borm, & Goeman, 2015; Ioannidis, 2008; Ioannidis, 2014). These issues especially emerge in the field of genetics because of the very small effect sizes that can be expected for associations with complex traits (Lango Allen et al., 2010; Ripke et al., 2014; Visscher, 2008). Meta-analyses can be used to derive a summary statistic that accumulates the evidences from several studies and analyses that deal with a related research question. Hence, a meta-analysis is a suitable tool to verify a finding that was initially based on a rather small sample size and to provide more realistic effect size estimates for an association. Contrary to primary research, the analytical procedure of a meta-analysis is rather fixed and guidelines have been developed that should be followed when performing such an analysis (Moher, Liberati, Tetzlaff, & Altman, 2009).

The typical analytical workflow comprises an in-depth literature search best based on different search engines to identify relevant publications (see Figure 5). Via a systematic review, these publications have to be screened iteratively to exclude irrelevant publications. The reasons for exclusion and inclusion of specific publications have to be explicitly stated in a reproducible manner. Depending on the research question, it might also be necessary to further classify the analytical outcome within each published finding to be able to perform sub-meta analyses. As an example, we classified the reported task-performances for the memory related meta-analyses in working memory and episodic memory related tasks. Because this process contains subjective decision processes of the researcher and might introduce unwanted biases, all these steps should be done completely independent by two persons. Ambiguous classifications should be discussed with a third independent person. The relevant information of each study needs to be extracted (see Figure 6) to assess the quality of the research and to retrieve the relevant statistics for the meta-analysis. Again, to reduce any bias, two persons should independently follow through these steps.

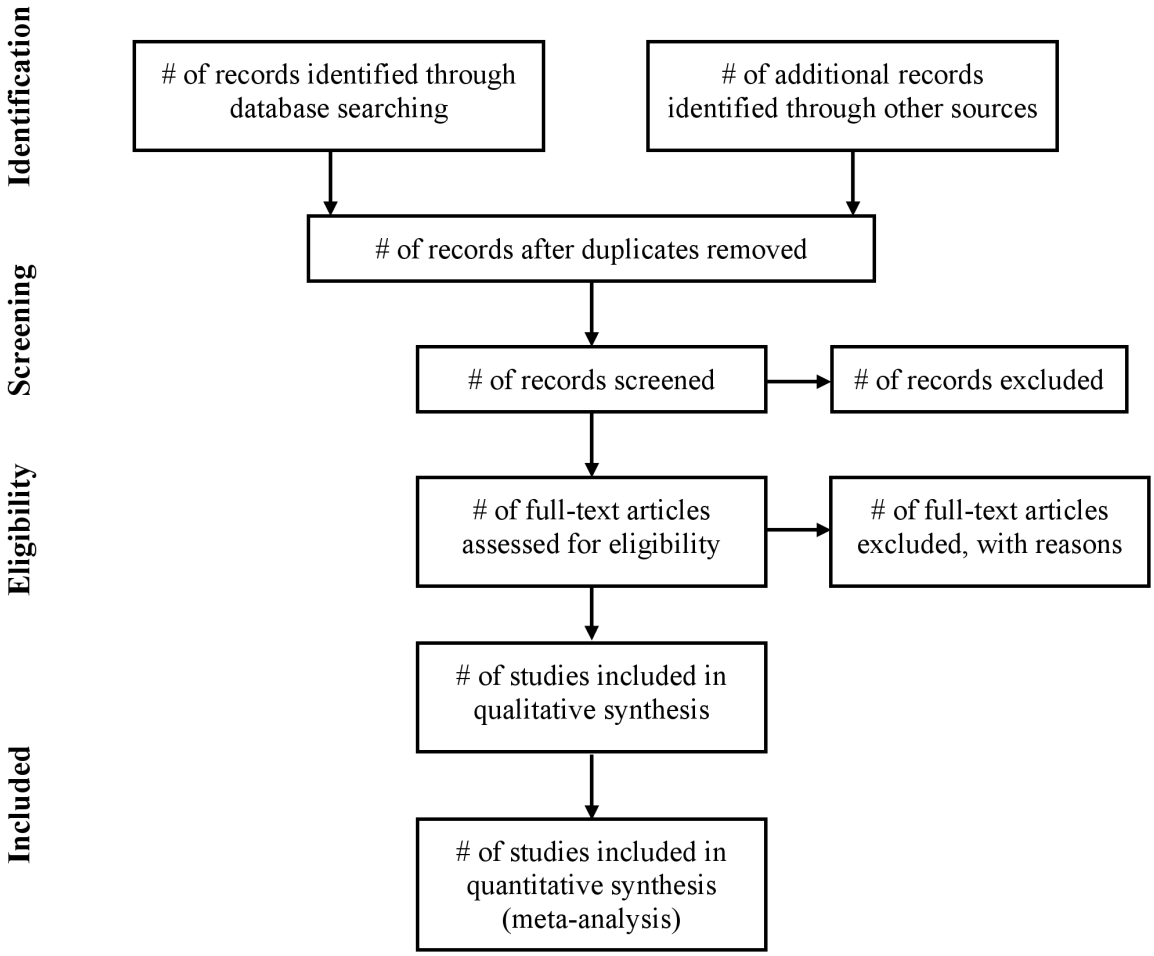


Figure 5 Flow of information through the different phases of a systematic review. Text and graphic are taken from Moher et al., 2009.

There are two main options for the meta-analysis itself named random-effect and fixed-effect meta-analysis. Random-effect but not fixed-effect meta-analysis can incorporate unequal variances between groups. This is especially important for genetic studies, where the included study populations might come from distinct ethnical groups or mixtures of ethnical groups. In this situation, a random-effect meta-analysis appears more appropriate because it will yield more conservative results in the presence of between-study heterogeneity (Hedges & Olkin, 1985; Ioannidis, Patsopoulos, & Evangelou, 2007). The choice of estimate to summarize the statistical effects is less critical, since these estimates can be transferred from one to another (Nakagawa & Cuthill, 2007). A standard algorithm for random-effect meta-analysis is the DerSimonian-Laird algorithm (DerSimonian & Laird, 1986), although it has been criticized that this algorithm is suboptimal when the number of studies included in the analysis is low (Hedges & Olkin, 1985).

Section/Topic	#	Checklist Item	Reported on Page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome-level assessment (see Item 12).	
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group and (b) effect estimates and confidence intervals, ideally with a forest plot.	
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., health care providers, users, and policy makers).	
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review level (e.g., incomplete retrieval of identified research, reporting bias).	
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	

doi:10.1371/journal.pmed.1000097.t001

Figure 6 Checklist of items to include when reporting a systematic review or meta-analysis. Text and graphic are taken from Moher et al., 2009.

Although a meta-analysis accumulates findings from different publications, and hence also different research groups, populations and experimental settings, it is not free from biases like e.g. the winner's curse phenomenon or publication bias. The scientific knowledge building cannot be seen as an independent process. Significant findings in a hot-topic field will stimulate adjacent research dealing with the same research questions. Unfortunately, the scientific publication procedure favors significant above non-significant findings and hence one cannot expect to accumulate all knowledge of a given association when performing a meta-analysis, but that there will be a bias towards positive findings especially at the beginning, which is called "proteus phenomenon" (Pfeiffer, Bertram, & Ioannidis, 2011).

3.2 Heritability estimation

Heritability of a complex trait is a prerequisite to perform genetic association studies with common genetic marker. Heritability estimates depend on the population under investigation as well as the environment (Visscher, Hill, & Wray, 2008) and the method used to estimate the heritability (Mousseau & Roff, 1987). Typically, heritability estimates for phenotypic data in humans have been estimated by comparing correlations between relatives, where the extent of genetic relatedness is derived from the degree of relationship. One method is the parent-offspring regression (Mousseau & Roff, 1987), but it is also possible to estimate heritability by applying mixed-model approaches to pedigree data (Akesson, Bensch, Hasselquist, Tarka, & Hansson, 2008; Kruuk, 2004). To reduce confounding effects of the environment, a preferred method is to compare monozygotic with dizygotic twins. Most-recently, based on 14'558'903 twin pairs, also meta-heritability estimates have been estimated that accumulate heritability estimates of 17'804 traits from 2'748 publications published between 1958 and 2012 (Polderman et al., 2015).

Due to the larger sample-sizes that are now available for complex-trait analyses, it is also possible to derive heritability estimates from unrelated individuals based on common SNP arrays, the so-called SNP-chip heritability (Yang, Lee, Goddard, & Visscher, 2011). The basis of this method is also a linear mixed model, however this method estimates genetic variation captured by all SNPs. In contrast to that approach the methods applied to parent-offspring pairs and siblings estimate the

genetic variance based on the phenotypic correlation for these pairs. Hence, these methods capture the contribution of all variants within the genome and not only of common SNPs (Yang et al., 2011). For the SNP-chip heritability, it is important to note that the sample-size has to be large enough to retrieve 1) accurate enough heritability estimates and 2) standard errors that are small enough for testing of significance. This is especially important if the overall heritability is not that high. From the analyses performed for this thesis, a rough approximation is that the sample-size should be 2'000 subjects or more for traits with a medium (defined as h^2 0.3-0.5) heritability.

Heritability estimates reflect the proportion of total variance of a phenotype that is attributable to genetic variation (e.g. in total or for common SNPs). If it comes to association studies and to the estimation of expected effect sizes, this is however a rather imprecise information (Visscher et al., 2008). Only in combination with the genetic architecture of a trait it is possible to make inferences about the expected effect sizes. Simply spoken, the higher the polygenicity is, the less contribution can be attributed to a single genetic factor.

Most recently, a complementary method has been developed, called stratified LD (linkage disequilibrium) score regression. Here, by adding annotation information to the genetic variants it is further possible to test whether groups of genetic variants disproportionally contribute to the heritability estimates (Finucane et al., 2015). In their analyses, especially conserved variants contributed most to the heritability estimates of 17 complex diseases. The stratified LD score regression method is a very nice example about how annotation information can be used as an additional analytical layer to gain more insight into a scientific field.

3.3 Replication studies

Large-scale array platforms exist e.g. for SNP-data, CpG methylation data or gene expression data, which measure hundred thousands of single-site signals for each dataset. The basic principle of an array is that the signal of each site, like one SNP, one CpG or one gene-expression value, are measured with several probes on one array, with the probe being the smallest entity on the array (see Figure 7). All probes that measure one site are called probe-set. Different normalization and

summarization methods exist to retrieve an interpretable signal for all sites that are measured on one chip. These site-signals typically comprise variability based on biological or environmental factors, technical variability and noise. Replication analyses can be used to dissect the different variance components of array data (Allison, Cui, Page, & Sabripour, 2006; Altman, 2005).

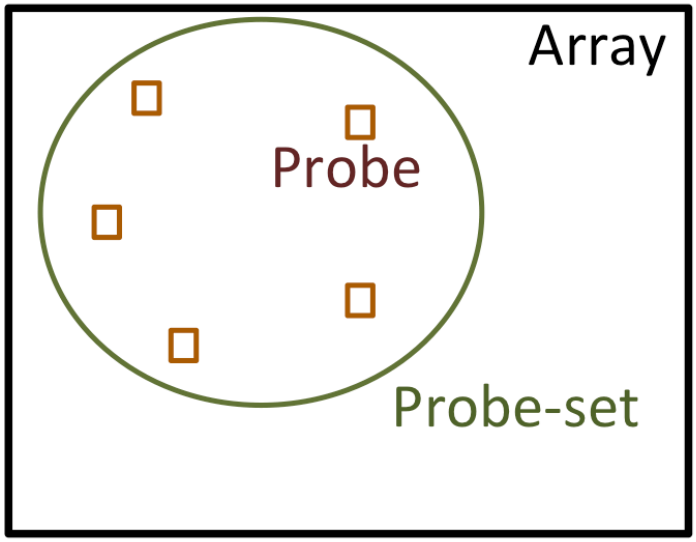


Figure 7 Generic and simplified concept of an array to measure e.g. SNPs, CpGs or gene-expression values. The smallest unit on the array is called a probe. Several probes together form a probe-set that measures the smallest interpretable unit of an array, like e.g. one specific SNP, CpG or expression values of one gene, which is called site. Normalization and summarization methods exist to generate for each site one interpretable signal that can be used for the downstream analyses.

Replication analysis with array data can be done with different material, with different arrays, with different subjects and on different levels. In a technical replication, the same material from the same time-point and the same subject is processed independently at least two times with the same array, whereas in biological replication the same procedure is performed but with the material of multiple subjects (Allison et al., 2006). Biological replicates are necessary to draw inferences from a study population.

Technical replication analyses are typically done on array level (Tylee, Kawaguchi, & Glatt, 2013), meaning that the signal of all sites is compared at once between two repeated measurements of an array. This method can also be used to compare signals between different platforms (for an example see Figure 8) and typically leads to very high confidence rates for DNA CpG methylation data (Bibikova et al., 2011; Sandoval et al., 2011). Technical replication analyses on array level

correspond to the comparison of profiles over all sites. Similar to a heritability estimate, this is a very general measurement and it is not possible to draw conclusions for each single site included in the analysis. However, high confidence rates on array level within but also between platforms are a prerequisite to perform more in-depth analyses of reproducibility e.g. on single-site level.

Association studies are typically done with single sites and inferences are drawn from the whole group of biological replicates. Based on a suitable number of biological replicates for which also technical replicates are available, it is also possible to assess the reliability on single-site level. With this approach, we compare the two repeated measurements of one site between all biological replicates. The number of biological replicates in the analysis determines the power to detect sites that can be measured reliably at least to some degree (assuming an alpha-error of 5 % and the power to be 80%): Based on 48 biological replicates it is possible to detect an *r* of 0.39 which corresponds to an explained variability of 15.3 %, whereas based on 96 biological replicates it is possible to detect an *r* of 0.28 (explained variability of 7.9 %). The reliability estimates derived from this analysis reflect the share of common variability in the signal that is based on differences in the biological or environmental background of the investigated population, but not due to technical variance or noise.

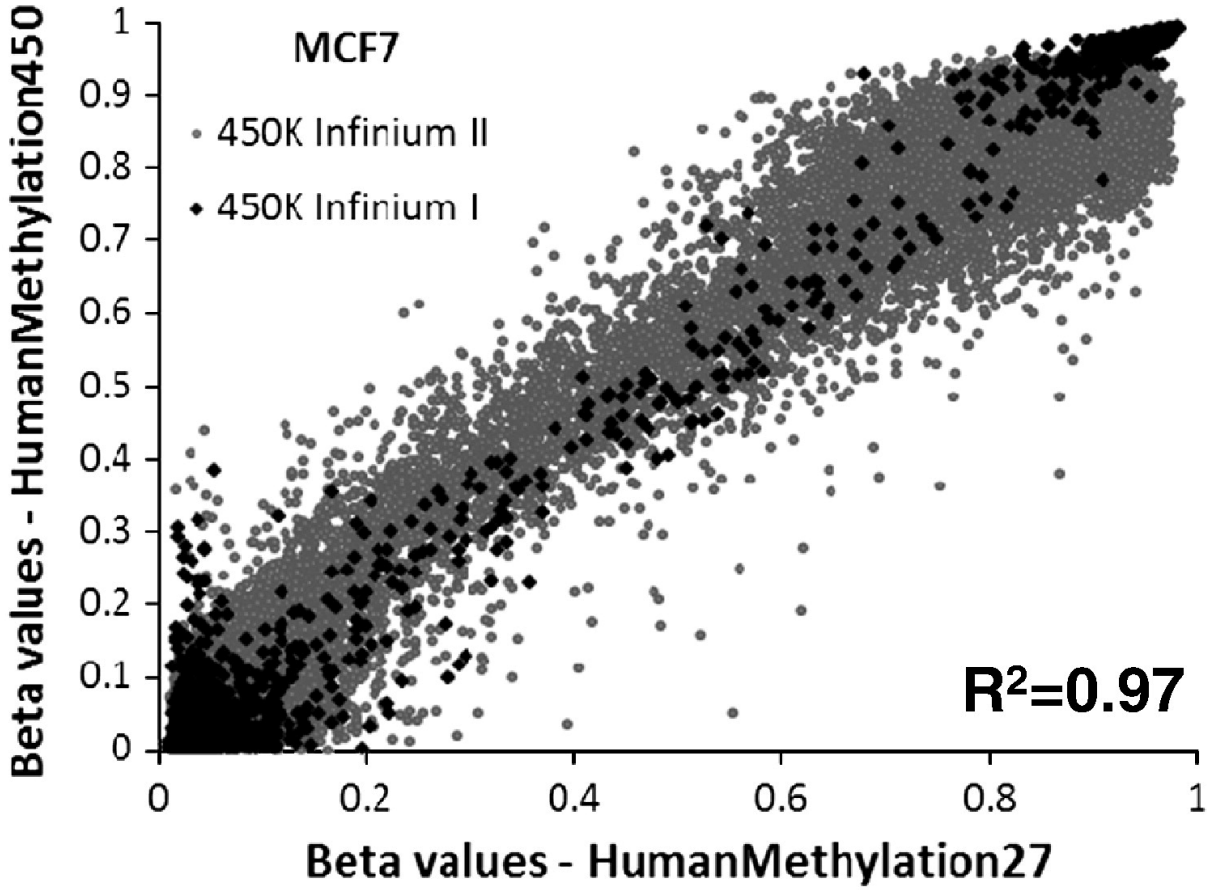


Figure 8 Correlation between HumanMethylation450 and HumanMethylation27 arrays. The plot illustrates the correlation of beta values between HumanMethylation450 and HumanMethylation27 arrays across 25'978 different CpGs (Bibikova et al., 2011).[2]

[2] Reprinted from Genomics, 98(4), Fig. 4, Bibikova M. et. al., High density DNA methylation array with single CpG site resolution, 288–295, 2011, with permission from Elsevier.

4 Original Research Papers

4.1 Association of KIBRA with episodic and working memory: a meta-analysis

RESEARCH ARTICLE

AMERICAN JOURNAL OF medical genetics PART B
Neuropsychiatric Genetics

Association of *KIBRA* With Episodic and Working Memory: A Meta-Analysis

Annette Milnik,[1,2]* Angela Heck,[1,3,4] Christian Vogler,[1,3,4] Hans-Jochen Heinze,[2,5,6] Dominique J.-F. de Quervain,[4,7] and Andreas Papassotiropoulos[1,3,4]

[1]Division of Molecular Neuroscience, Department of Psychology, University of Basel, Basel, Switzerland

[2]Department of Neurology, University of Magdeburg, Magdeburg, Germany

[3]Life Sciences Training Facility, Department Biozentrum, University of Basel, Basel, Switzerland

[4]University Psychiatric Clinics, University of Basel, Basel, Switzerland

[5]Leibniz Institute for Neurobiology, Magdeburg, Germany

[6]Helmholtz Center for Neurodegenerative Diseases, Magdeburg, Germany

[7]Division of Cognitive Neuroscience, Department of Psychology, University of Basel, Basel, Switzerland

Manuscript Received: 2 May 2012; Manuscript Accepted: 10 September 2012

WWC1 was first implicated in human cognition through a genome wide association study in 2006 that reported an association of the intronic single nucleotide polymorphism (SNP) rs17070145 with episodic memory performance. *WWC1* encodes the protein KIBRA, which is almost ubiquitously expressed. Together with its binding partners, KIBRA is assumed to play a role in synaptic plasticity. T-allele carriers of SNP rs17070145 have been reported to outperform individuals that are homozygous for the C-allele in episodic memory tasks. Here we report two random effects meta-analyses testing the association of rs17070145 with episodic and working memory. All currently available population-based association studies that investigated effects of rs17070145 on episodic or working memory were included in the analyses. Where performance measures for multiple domain-specific tasks were available for a given study population, averaged effect size estimates were calculated. The performed meta-analyses relied on 17 samples that were tested for episodic memory performance (N = 8,909) and 9 samples that had performed working memory tasks (N = 4,696). We report a significant association of rs17070145 with both episodic (r = 0.068, *P* = 0.001) and working memory (r = 0.035, *P* = 0.018). In summary, our findings indicate that SNP rs17070145 located within *KIBRA* explains 0.5% of the variance for episodic memory tasks and 0.1% of the variance for working memory tasks in samples of primarily Caucasian background.

Key words: WWC1; rs17070145; single nucleotide polymorphism; common variant; human

How to Cite this Article:
Milnik A, Heck A, Vogler C, Heinze H-J, de Quervain, DJ-F, Papassotiropoulos A. 2012. Association of *KIBRA* With Episodic and Working Memory: A Meta-Analysis.
Am J Med Genet Part B 159B:958–969.

INTRODUCTION

Genome-wide association studies (GWAS) offer a hypothesis-free approach to identify new genotype-phenotype associations by simultaneously testing hundred thousands of genetic markers. Yet, one of the key issues GWAS have to deal with is the multiple testing burden that easily gives rise to false positive results. In order to downsize the probability of false positive results, two major strategies are commonly applied: A, conservative multiple testing correction as for example Bonferroni correction and B, even more important, replication of the observed association in independent samples. In highly polygenic traits, single genetic variations are expected to have small impact on the total variance. Given the scenario of small to medium sample sizes with concurring true and small effect sizes, the necessary correction for type I error inflation comes along with a higher probability of overestimating the real

Additional supporting information may be found in the online version of this article.

Grant sponsor: The Swiss National Science Foundation (Sinergia Grants); Grant numbers: CRSIK0_122691, CRSI33_130080; Grant sponsor: The 7th framework programme of the European Union (ADAMS Project); Grant number: HEALTH-F4-2009-242257.
The authors declare no conflict of interest.

*Correspondence to:
Annette Milnik, Division of Molecular Neuroscience, University of Basel, Birmannsgasse 8, 4055 Basel, Switzerland.
E-mail: annette.milnik@med.ovgu.de

Article first published online in Wiley Online Library (wileyonlinelibrary.com): 12 October 2012
DOI 10.1002/ajmg.b.32101

effect size. Thus replication studies are hampered by the fact that they must be sufficiently powered in order to be capable of detecting also smaller effects. Despite the fact that recruiting large sample sizes is challenging, the GWAS approach paves the way for an unbiased view on the impact of the human genome on polygenic traits.

Heritability estimates for both episodic and working memory are substantial (h^2 0.37–0.74%) [Volk et al., 2006; Greenwood et al., 2011] and have not only been observed in healthy individuals but also in patients with memory disorders, such as late onset Alzheimer's disease (AD; h^2 0.42–0.47%) [Wilson et al., 2011]. Heritability for memory performance has also been estimated in an African [Calkins et al., 2010] and a Japanese [Ando et al., 2001] population with similar results (h^2 0.32–0.5%) as in Caucasian samples. The genetic effect is comparable for men and women [Read et al., 2006], relatively constant across different adult age groups [Finkel et al., 1995; Finkel and McGue, 1998], and also present in the oldest old [Johansson et al., 1999].

In 2006, Papassotiropoulos et al. [2006] reported an association of single nucleotide polymorphism (SNP) rs17070145 (T → C substitution) with episodic memory (initial sample N = 333, 5 min delayed free recall of words $P = 4 \times 10^{-6}$, 24 hr delayed free recall of words $P = 0.002$, averaged r = 0.22). This finding was replicated in the frame of the initial study in two independent samples (Europe and USA) of Caucasian genetic background (European sample N = 424, 10 min delayed free recall of pictures $P = 0.006$, r = 0.13; US sample N = 256, 30 min delayed free recall of words $P = 0.004$, Buschke Selective Reminding Test $P = 5 \times 10^{-5}$, averaged r = 0.22). Triggered by the initial report, further studies investigated a potential association of rs17070145 with a large variety of psychometric measures [Almeida et al., 2008; Nacmias et al., 2008; Need et al., 2008; Schaper et al., 2008; Bates et al., 2009; Jacobsen et al., 2009; Zhang et al., 2009; Hayashi et al., 2010; Preuschhof et al., 2010; Vassos et al., 2010; Yasuda et al., 2010; Burgess et al., 2011; Kauppi et al., 2011; Sédille-Mostafaie et al., 2011; Wersching et al., 2011]. Although being heterogeneous in respect to sample size and genetic background of sampled populations, the association of rs17070145 with memory performance could be replicated in independent studies and samples [Schaper et al., 2008; Almeida et al., 2008; Bates et al., 2009; Preuschhof et al., 2010; Vassos et al., 2010; Yasuda et al., 2010; Hayashi et al., 2010; Kauppi et al., 2011]. Non-replications were also reported [Need et al., 2008; Nacmias et al., 2008; Bates et al., 2009; Burgess et al., 2011; Wersching et al., 2011; Sédille-Mostafaie et al., 2011].

In addition to the effect of rs17070145 on episodic memory, some studies also detected a significant association between *WWC1* and working memory performance, or tasks which feature an overlap between episodic and working memory [Schaper et al., 2008; Vassos et al., 2010; Yasuda et al., 2010; Kauppi et al., 2011]. Recent research provides evidence for the involvement of the medial temporal lobe in working memory processes, besides its long known critical role for episodic memory [Bird and Burgess, 2008; Ranganath, 2010]. Given the important role of *KIBRA* for medial temporal lobe and hippocampal functioning as discussed in the initial study [Papassotiropoulos et al., 2006], a potential role for *KIBRA* in working memory can be considered as well. Behavioral studies on memory in Caucasian populations suggest that T-allele carriers outperform individuals that are homozygous for the C-allele.

SNP rs17070145 resides within the ninth intron of *WWC1*, which is located on chromosome 5q34 and encodes the KIBRA protein. *KIBRA* is almost ubiquitously expressed and is also present in memory-related brain regions [Papassotiropoulos et al., 2006; Johannsen et al., 2008]. Together with its binding partners, KIBRA is assumed to play a role in synaptic plasticity [Schneider et al., 2010]. For example, KIBRA interacts with PKMzeta, the brain-specific atypical isoform of PKCzeta [Büther et al., 2004; Schneider et al., 2010], which is strongly associated with memory formation and the maintenance of long-term potentiation (LTP). Furthermore, the WW domain allows KIBRA to interact with dendrin and synaptopodin [Kremerskothen et al., 2003; Duning et al., 2008], which play a role in synaptic plasticity, signal transduction and the organization of the cytoskeleton [Herb et al., 1997; Kremerskothen et al., 2006; Duning et al., 2008]. Additional interaction partners of KIBRA are implicated in vesicle-based transport processes, cell polarity, cell migration, and transcriptional regulation [Schneider et al., 2010]. Recent in vitro and animal studies demonstrated that KIBRA is part of an α-amino-3-hydroxy-5-methyl-4-isoxazole-propionic acid receptor (AMPAR) protein complex, which regulates AMPAR membrane trafficking. Adult KIBRA knockout mice were viable, but exhibited deficits in hippocampal LTP and long-term depression along with an impairment of learning and memory [Makuch et al., 2011].

While deficits in episodic memory are a key symptom of dementia, impairments in episodic and working memory tasks can be observed in a variety of neuropsychiatric disorders comprising schizophrenia [Forbes et al., 2009; Mesholam-Gately et al., 2009; Van Snellenberg, 2009], depression [Drevets et al., 2008; Marazziti et al., 2010], and posttraumatic stress disorder [Samuelson, 2011]. Having been associated with memory performance in healthy subjects, *WWC1* thus also became a target for the investigation of such disease-related phenotypes as mild cognitive impairment [Almeida et al., 2008] and dementia [Rodríguez-Rodríguez et al., 2009; Corneveaux et al., 2010; Hayashi et al., 2010; Burgess et al., 2011]. In a meta-analytic approach that included 16 independent samples, Burgess et al. [2011] could show a statistical trend for a decreased risk of T-allele carriers for dementia (4,436 cases and 4,334 controls, protective Odds Ratio for T-allele 0.94, $P = 0.07$).

Taken together, the initial findings yielded by the GWAS approach in healthy individuals served as a starting point to further investigate a potential association of *KIBRA* with susceptibility to memory-related diseases [Burgess et al., 2011]. Recent studies also suggested possible mechanisms [Schneider et al., 2010; Makuch et al., 2011] through which KIBRA acts on memory.

The present study aimed at estimating a general effect size for the association of the KIBRA SNP rs17070145 with episodic memory. In addition, we investigated whether rs17070145 is also associated with working memory. In order to achieve a task-independent, but domain-specific measure for the effect sizes of the *WWC1* SNP on episodic and working memory, we averaged over all available domain-specific performance-tests per sample. We included all available samples that provide sufficient information for estimating effect sizes. Since the included samples were heterogeneous with respect to genetic background, health status, age range, and

accomplished tasks, we performed random effects meta-analyses, because this approach incorporates unequal variances between groups. In the case of large between-study heterogeneity, random effects meta-analysis will yield more conservative results in comparison to fixed effects meta-analysis [Ioannidis et al., 2007].

MATERIALS AND METHODS

Literature Search and Study Selection

PubMed- and Medline-based literature search was done in November 2011 with "KIBRA" or "WWC1" as search terms (Supplementary Fig. 1). Studies were included, if they provided (I) a population-based association analysis that (II) comprised SNP rs17070145 and episodic or working memory tasks and (III) applied a T-dominant or additive genetic model. Furthermore the estimation of effect sizes and confidence intervals required the following information to be available (IV): Sample size, statistical model, test statistic or *P*-value and direction of effect.

Data Extraction

All studies were screened by two authors (A.M., A.H.). Data were independently extracted. For each study, we recorded name of first author, publication year and number of independent samples reported. For each sample, the following information was extracted or calculated: Sample size, age-range and mean age, sample sizes of genotype sub-groups (CC vs. CT vs. TT), genotyping method, ethnicity, health status, collected psychometric measurements, *P*-value for sex and age differences between genotype groups (CC vs. CT/TT), *P*-value for Hardy–Weinberg equilibrium (HWE). For each episodic and working memory task, the following information was extracted: Mean and standard deviation of performance for the two genotype groups (CC vs. CT/TT), calculated statistic model with included covariates, test statistics and *P*-value, direction of effect. If necessary, corresponding authors were contacted to provide missing information or to clarify inconclusive information.

Selection of Psychometric Measurements

Two authors (A.M. and A.H.) independently assigned the reported psychometric measurements to the following domains: Episodic memory tasks, working memory tasks that feature a potential overlap with episodic memory, pure working memory tasks, and memory-unrelated tasks. Since the set-up of working memory tasks has a potential influence on the involvement of different memory components, we further subdivided tests which feature a potential overlap with episodic memory in tasks with a marked working memory component (below referred to as narrow-sense working memory tasks), and tasks with a weaker working memory component (below referred to as broad-sense working memory tasks). Assignment to episodic memory, broad-sense working memory and narrow-sense working memory domains was done as follows: if the tasks comprised an interval of at least 3 min or a distraction between encoding and recall, it was allocated to episodic memory. If no delay between encoding and recall was present, items to be remembered were presented only once and stimulus material was not requiring multimodal processing (as would be the case for story recall or tasks involving the performance of actions, termed "subject-performed tasks"), then the task was allocated to narrow-sense working memory. If delay between encoding and recall was less than 3 min, but stimulus material was presented and recalled repeatedly or if the task included multimodal processing, then the task was allocated to broad-sense working memory (see Supplementary Table I). Allocation to the specific memory domains was additionally discussed by three of the authors (A.M., A.H., and C.V.).

Quality Assessment

Given the reported effects of sex and age on cognitive performance [Park et al., 2002; Read et al., 2006; Andreano and Cahill, 2009], most studies included sex and age as covariates in their analysis. Wherever reported sex or age effects between genotype groups (CC vs. CC/TT) were far from significance ($P > .25$), we rather used the uncorrected values for effect size estimation. In all other cases, we preferred to use the statistics that were corrected for sex and age. Wherever more than one linear model was calculated in the scope of one analysis, we used those models recommended in the respective studies (if recommendations were given).

Data Analysis

We used r as effect statistic estimate. The effect statistic r was preferred over d, because r^2 denotes the portion of phenotypic variance that can be explained by the genotypic variation. Wherever only an additive genetic model was reported, r was directly estimated from the regression coefficients of the linear model or from *P*-values. Wherever a two-group comparison was reported, we estimated Hedges d [Hedges and Olkin, 1985], since it is corrected for a potential bias caused by small sample size. Hedges d was preferentially estimated from standardized mean differences, *t*-values of two-sample *t*-tests, or *P*-values. If raw value information was missing, or if relevant confounding effects of sex or age were reported, Hedges d was estimated based on the following statistics: (M)AN(C)OVA *F*-value, linear regression *t*-value, regression coefficient or *P*-value. We corrected for the number of covariates used and for unequal sample sizes between genotype groups, wherever possible. Yet, given the allelic distribution in Caucasians, the effect of correction for unequal sample sizes is nearly negligible, if the T-dominant model (CC vs. CT/TT) was calculated. After estimation and correction, we transformed Hedges d to the test statistic r. If enough information was provided to calculate the estimates from the additive or the T-dominant genetic model, we preferred the T-dominant genetic model.

If one sample performed more than one episodic or working memory task, we averaged the effect size over all episodic or working memory tasks, respectively. Wherever sample sizes differed between tasks or tasks were performed in sub-samples only, we calculated a mean weighted by the size of the sub-samples. Thus no bias favoring the larger sample was introduced and the more accurate effect size estimation for larger samples was taken into account. In this scenario, we set the number of subjects equal to the N of the larger sample. If follow-up measurements using the exact same task were

reported, only the performance at the first time point was considered.

We calculated two main random effects meta-analyses, one for episodic memory and one for pure and narrow-sense working memory combined. A separate analysis of solely pure working memory was not feasible due to the low number of available tasks in that category. Homogeneity was tested using the Q-test for between-study heterogeneity. We additionally report I^2 as a measurement of heterogeneity, which varies between 0 and 100%. I^2 is a measure of the percentage of total variation across studies due to heterogeneity beyond chance, values over 50% indicate large heterogeneity [Ioannidis et al., 2007]. Potential publication bias was assessed by visual inspection of funnel plots and by regression tests for asymmetry. Furthermore, we included additional analyses that feature unpublished data of three samples that performed episodic and mainly pure working memory tasks to adjust for a potential publication bias. We also excluded the three samples of the initial study from the episodic memory analysis to correct for the winner's curse phenomenon. Additional analyses were calculated to determine ethnicity-specific estimates, but due to the low number of studies involving non-Caucasian samples, it was not feasible to separately conduct meta-analyses for these sub-samples. We also assessed whether the different definitions of working memory influenced the results of the meta-analyses.

Data analysis was done with R 2.14.0 [R Development Core Team, 2011]. Estimation of Hedges d and r was done as recommended by Nakagawa and Cuthill [2007]. The random effects model was calculated using the metafor package [Viechtbauer, 2010] and the DerSimonian-Laird algorithm [DerSimonian and Laird, 1986]. Reported *P*-values are two-tailed.

To illustrate the sample sizes needed to detect effects in the range as estimated from our meta-analyses, we calculated power analyses using the pwr package [Champely, 2009]. Alpha error rate was set to 0.05 in the case of one independent test, which represents the scenario of a single SNP replication ($P < 0.05$). In addition, we simulated the GWAS scenario using a Bonferroni correction to adjust for alpha-inflation in the case of multiple comparisons, assuming 1×10^6 independent tests per GWAS ($P < 5 \times 10^{-8}$).

RESULTS

Search Results

Out of 47 hits that were yielded by the literature search with the search terms "KIBRA" and "WWC1" (Supplementary Fig. 1), we selected all 16 studies, which gave information on association between KIBRA and a cognitive phenotype in the title or the abstract. All 16 studies provide genotype information for SNP rs17070175. The 31 not-further screened studies investigated molecular mechanism of KIBRA [Duning et al., 2008], KIBRA and disease status without investigating cognitive phenotypes [Galecki et al., 2010], or were review articles of KIBRA [Schneider et al., 2010].

Excluded Studies and Samples

Out of the 16 studies that underwent the full text search, a total of four studies had to be excluded completely with all provided samples. Reasons for exclusion were as follows: Missing memory phenotypes [Zhang et al., 2009], calculation of a C-dominant genetic model (CC/CT vs. TT) [Hayashi et al., 2010], study design (family-based association instead of a population-based association) [Vassos et al., 2010], or missing statistics for effect size calculation [Jacobsen et al., 2009]. One study provided information on three independent samples out of which one had to be excluded. For this specific sample (termed RS sample in the original study) reported sample size information in combination with health status was not traceable by means of the published data [Burgess et al., 2011]. Follow-up measurements of the identical task reported in two of the included studies were not considered in the analysis [Kauppi et al., 2011; Sédille-Mostafaie et al., 2011]. Sample and task information, estimation of effect size and direction of effect for all excluded studies, samples and follow-up measurements, as far as conclusive, are summarized in Supplementary Table II.

Included Studies

A total of 12 studies comprising 17 independent samples met our inclusion criteria (Table I). All studies were published between 2006 and 2011. Episodic memory tasks were performed in all 17 samples and for a total of 9 samples pure and narrow-sense working memory tasks were reported. In the additional working memory meta-analysis, we incorporated a total of 13 samples, allowing also broad-sense working memory tasks in the analysis. We also added three independent unpublished samples (Supplementary Table III) taken from our ongoing studies, which all provide episodic and mainly pure working memory tasks. Where reported or calculable, *P*-values for HWE tests of rs17070145 were $P > 0.03$. For all but two samples [Burgess et al., 2011], we estimated the effect sizes from a T-dominant model. Age range varied from 18 to 100. Taking into account all studies that provided information on the mean age of their study populations, the over all mean age was 64. Inclusion of the unpublished samples decreased the mean age to 57. Apart from the two non-Caucasian samples, a Japanese sample [Yasuda et al., 2010] and an African American sample [Burgess et al., 2011], ethnic background of the study populations was mainly Caucasian. Considered covariates varied widely between studies. It is important to note that effect size estimation from uncorrected values was not always possible because of missing values or confounding covariates.

Episodic Memory Random Effects Meta-Analysis

All published studies together comprise 17 samples that performed episodic memory tasks and overall sample size totaled to N = 8,909. Effect size estimates per sample and the overall estimated effect size for the main meta-analysis are shown in Figure 1. I^2 designates high between-study heterogeneity (66.70%). We obtained a significant result for the Q-test for between-studies heterogeneity ($Q = 48.05$, $df = 16$, $P = 5 \times 10^{-5}$), which requires the calculation of a random effects meta-analysis. Neither funnel plot (Supplementary Fig. 2) nor regression testing ($t = 0.87$, $df = 15$, $P = 0.40$) indicated a significant publication bias. Estimated r for episodic memory is 0.068 ($Z = 3.30$, $P = 0.001$, 95% CI 0.028–0.109; corresponding

TABLE I. Overview of the Included Samples, Which Are Published

Author	Year	Independent sample	Health and cognition status, sample information	Age range (mean)	Ancestry	Allele frequency				Original genetic model	Geno-type method	Task category: included tasks	GT × Sex (*P*)	GT × Age (*P*)
						CC	CT	TT	HWE					
Almeida	2008	–	Healthy; MMSE ≥ 24; CDR ≤ 0.5; DSM-IV no dementia; partially MCI probable	>50 (71.2[b])	Mainly Caucasian	138	143	31	0.52	T-dom	PCR–RFLP	1: CERAD delayed recall; CERAD recognition; 2c: CERAD immediate recall same word list three runs	0.49	0.01
Bates	2009	AAA	Aspirin for asymptomatic atherosclerosis randomized controlled trial; MMSE ≥ 24	55–82 (67.3)	Scotland	1,040	1,020	203	0.04[b]	T-dom	PCR (KASPar) or Taq-Man	1: AVLT delayed recall	NA	NA
		LBC	Lothian birth cohort	79 (79)	Scotland	254	218	57	0.31[b]	T-dom	TaqMan	1: LMT delayed story recall 30 min; 2c: LMT immediate story recall	NA	–
Burgess	2011	AA	CDR = 0; controls only	60–90[b] (73.2[b])	African-American	35[b]	110[b]	97[b]	0.68[b]	Additive	Sequenom	1: AVLT delayed recall; LMT delayed story recall 30 min	NA	NA
		JS	CDR = 0; Controls only	60–100[b] (78[b])	Caucasian-American	270[b]	251[b]	67[b]	0.45[b]	Additive	Sequenom	1: AVLT delayed recall; LMT delayed story recall 30 min	NA	NA
Kauppi	2011	–	Non-demented subjects; Betula project second wave	35–85 (58.4)	Sweden	1,040	932	258	0.03[b]	T-dom	Sequenom	1: SPT delayed cued recall; VT delayed cued recall; 2b: VT immediate free recall; immediate free recall of words composite score; 2c: SPT immediate free recall	0.5	0.5
Nacmias	2008	–	Subjective memory complaints with intact cognitive performance	42–78[a] (60.7)	Italy	30	33	7	0.79[b]	T-dom or additive	PCR–RFLP	1: Words free recall 10 min delay; words free recall 24 h delay; paired words free recall 10 min delay; paired words free recall 24 h delay; short story delayed recall; Rey Osterrieth Complex Figure Test recall; 2a: Corsi block-tapping test; Digit span; Trail-Making test B-A; 2b: Words acquisition; paired words acquisition; 2c: Short story immediate recall	0.58[a]	0.58[a]
Need	2008	Duke	Duke Genetics of Memory cohort	18–77 (md 23)	United States, non-European excluded	NA	NA	NA	NA	T-dom[a]	NA	1: Prose passage delayed recall 30 min; 2b: VRM CANTAB; 2c: Prose passage immediate recall	NA	NA
		German	NA	22–75 (md 51)	German	NA	NA	NA	NA	T-dom[a]	NA	1: AVLT delayed recall; 2b: AVLT immediate recall	NA	NA

(*Continued*)

TABLE I. (*Continued*)

Author	Year	Independent sample	Health and cognition status, sample information	Age range (mean)	Ancestry	Allele frequency CC	CT	TT	HWE	Original genetic model	Geno-type method	Task category: included tasks	GT × Sex (*P*)	GT × Age (*P*)
Papassotiropoulos	2006	Swiss1	Healthy subjects	18–48 (md 22)	Switzerland	164	135	34	0.44[b]	T-dom	Pyro-sequencing	1: Verbal memory task free recall 5 min delay; verbal memory task free recall 24 h delay; 2b: Verbal memory task immediate free recall	NA	NA
		Swiss2	Healthy subjects	18–28 (md 21)	Switzerland	NA	NA	NA	NA	T-dom	Pyro-sequencing	1: Visual episodic memory task free recall 10 min delay; 2a: Digit span	NA	NA
		USA	Cognitive normal subjects	20–81 (md 55)	United states (93% European ancestry)	NA	NA	NA	NA	T-dom	Pyro-sequencing	1: AVLT 30 min delay; SRT free recall; 2b: AVLT 1 immediate recall	NA	NA
Preuschhof	2010	–	Healthy young subjects	20–31 (25.7)	Caucasian	169	158	56	0.06	T-dom	TaqMan	1: Item-paired memory task	0.46[a]	NA
Schaper	2008	–	Healthy subjects; intact cognitive performance; CERAD within 1.5sd	NA (67[b])	Caucasian	29	33	2	0.07[b]	T-dom	Pyro-sequencing	1: VLMT free recall 30 min delay; VLMT recognition; VLMT free recall after interference; 2c: VLMT total immediate recall 1–5	0.56[a]	0.9[a]
Sedille-Mostafaie	2011	–	Vienna Transdanube Aging Study baseline phase	75–76 (–)	Austria	NA	NA	NA	NA	T-dom	TaqMan	1: FOME; IDSR-7; CERAD world list delayed recall; figure copy recall saving score; 2c: CERAD word list immediate recall	NA	–
Wersching	2011	–	Randomly selected subjects; no severe neurological or psychiatric disease	50–85 (63.5)	Germany	235	234	76	0.15	T-dom	PCR–RFLP	1: AVLT 7 long-term retrieval; AVLT 8 recognition; AVLT 6 short-term retrieval; 2b: AVLT 1 immediate recall; 2c: AVLT 1–5 immediate recall; AVLT 5 immediate recall	0.73	0.24
Yasuda	2010	–	Healthy subjects	20–65 (35.9)	Japan	7	56	124	0.83	T-dom and additive	TaqMan	1: WMS-R delayed recall; 2b: WMS-R verbal memory; WMS-R visual memory	0.05	0.77

1: Episodic memory task; 2a: Pure working memory; 2b: Narrow-sense working memory task; 2c: Broad-sense working memory task; a: Personal communication; b: Calculated from raw data (weighted means for age, exact test for HWE, sum for genotypes); HWE: Hardy–Weinberg equilibrium; GT: Genotype; md: Median; MMSE: Mini-Mental State Examination; CDR: Clinical Dementia Rating; DSM: Diagnostic and Statistical Manual of Mental Disorders; MCI: Mild cognitive impairment; CERAD: Consortium to Establish a Registry for Alzheimer's Disease; AVLT: Auditory-Verbal Learning Test; LMT: Wechsler Logical Memory Test; SPT: Subject Performed Task; VT: Verbal Task; VRM: Verbal Recognition Memory; CANTAB: Cambridge Neuropsychological Test Automated Battery; SRT: Buschke Selective Reminding Test; VLMT: Verbal Learning and Memory Test-German Version of AVLT; FOME: Fuld Object Memory Evaluation; IDSR-7: Intracategorical Delayed Selective Reminding Test; WMS-R: Wechsler Memory Scale-Revised Edition.

d = 0.14), which results in 0.5% explained variance. Inclusion of three independent unpublished samples (overall N = 10,263, 1,354 Caucasian subjects added) yielded an estimated r of 0.058 (Z = 3.15, $P = 0.002$, 95% CI 0.022–0.093, $I^2 = 63.52\%$). Exclusion of the three samples of the initial study (overall N = 7.896, 1,013 Caucasian subjects excluded) yielded an estimated r of 0.04 (Z = 2.19, $P = 0.03$, 95% CI 0.004–0.076, $I^2 = 49.37\%$). Exclusion of non-Caucasian subjects (N = 411, African American and Japanese samples) from the published data results in an estimated r of 0.067 (overall N = 8,498, Z = 3.28, $P = 0.001$, 95% CI 0.027–0.107, $I^2 = 64.07\%$).

Working Memory Random Effects Meta-Analysis

All published studies together comprise nine samples that performed pure or narrow-sense working memory tasks and overall sample size totaled to N = 4,696. Effect size estimates per sample and the overall estimated effect size for the main meta-analysis are shown in Figure 2. There is no hint for between-study heterogeneity ($I^2 = 0\%$, Q = 7.02, df = 8, $P = 0.53$). Neither funnel plot (Supplementary Fig. 3) nor regression testing (t = −0.43, df = 7, $P = 0.68$) revealed a significant publication bias. Estimated r for pure and narrow-sense working memory together is 0.035 (Z = 2.36, $P = 0.018$, 95% CI 0.006–0.063; corresponding d = 0.07) which results in 0.1% explained variance. Exclusion of the pure working memory tasks (overall N = 4,272, Papassotiropoulos Swiss sample 2 excluded) yielded an estimated r of 0.037 (Z = 2.09, $P = 0.04$, 95% CI 0.002–0.073, $I^2 = 12.48\%$). Inclusion of three independent unpublished samples (overall N = 6,050; 1,354 Caucasian subjects added) yielded an estimated r of 0.040 (Z = 3.09, $P = 0.002$, 95% CI 0.015–0.065, $I^2 = 0\%$). All three independent unpublished samples provided mainly pure and in one case an additional narrow-sense working memory task. Exclusion of non-Caucasian subjects (N = 187, Japanese sample) from the published data results in an estimated r of 0.030 (overall N = 4,509, Z = 2.02, $P = 0.043$, 95% CI 0.001–0.059, $I^2 = 0\%$). Combined analysis of pure, narrow-sense and broad-sense working memory tasks from 13 published samples with an overall sample size of N = 6,186 results in an estimated r of 0.039 (Z = 2.21, $P = 0.027$, 95% CI 0.004–0.073, $I^2 = 35.11\%$).

We also performed a post hoc analysis combining all available data (pure, narrow-sense and broad-sense memory tasks, published and unpublished data), since inclusion of the mainly pure working memory tasks performed in the unpublished samples had the same direction of effect as inclusion of samples which provide broad-sense working memory tasks. This resulted in 16 independent samples with an overall sample size of N = 7,540. Between-study heterogeneity is present ($I^2 = 23.36\%$), but not significant (Q = 19.57, df = 15, $P = 0.19$). Estimated effect size for all 16 samples together results in r = 0.041 (Z = 2.90, $P = 0.004$, 95% CI 0.013–0.069; corresponding d = 0.08), explaining 0.2% variance.

Power and Sample Size Analysis

Detection of the estimated overall effect sizes with 80% probability at an alpha error rate of 5% in the case of one single test would require sample sizes of 2,000 subjects for an episodic memory task and 5,000 subjects for working memory tasks, respectively (Fig. 3A). In order to reach a multiple testing corrected significance threshold assuming 1×10^6 independent tests ($P < 5 \times 10^{-8}$ as an arbitrary, but often used significance threshold for GWAS) more than 10,000 subjects are required to detect the effect on episodic memory (Fig. 3B).

DISCUSSION

The present meta-analysis suggests a significant effect of *KIBRA* SNP rs17070145 on episodic and working memory performance and supports the importance of the GWAS approach for the identification of molecules related to cognition in humans. The meta-analysis-based effect size estimation for the association of rs17070145 with episodic memory is substantially lower than expected from the initial findings [Papassotiropoulos et al., 2006], a phenomenon often referred to as ''winner's curse'' [Kraft, 2008]. This phenomenon, that is overestimation of the effect from the initial study, affects many research fields [Ioannidis, 2008] including genetic association studies [Ioannidis et al., 2001].

Wherever we are dealing with small sample sizes in relation to the expected effect sizes and a stringent significance threshold, an overestimation of the detected effect becomes very likely. This phenomenon noticeably vanishes when samples with adequate power are studied [Ioannidis, 2008]. A recent meta-analysis of GWAS investigating height reports that cumulated effect sizes of 180 SNPs do not explain more than 17% of the observed variance using a sample size of more than 180,000 subjects [Lango Allen et al., 2010]. Single SNPs yielding the most substantial association signals with human height do not explain more than 0.3–0.5% of the observed variance [Weedon et al., 2007; Sanna et al., 2008; Visscher, 2008]. Thus, the estimated explained variances by rs17070145 on episodic and working memory performance lie within the expected range for a quantitative trait locus.

The expected moderate effect sizes in statistical genetics underline the importance of recruiting populations of adequate size in order to be able to draw valid conclusions from genetic studies. Thus it should be avoided to derive effect size estimates from the initial study population. In contrast, the estimation of effect size ranges should be based on sufficiently powered replication samples and meta-analysis. Updating of effect-size estimations when more data becomes available is a necessary means to obtain more reliable estimates. This course of action also counteracts the so-called ''proteus phenomenon'', that is early replications tend to be biased towards the initial findings [Pfeiffer et al., 2011], which often accompanies the ''winner's curse'' phenomenon.

In the present meta-analysis, we included two additional sub-analyses to account for the winner's curse and publication bias phenomena for the episodic memory task. To address the publication bias, we included data of three unpublished samples. In order to adjust for the winner's curse phenomenon, we excluded the initial study samples. Both analyses yielded smaller effect-size estimates for episodic memory, but the association remained

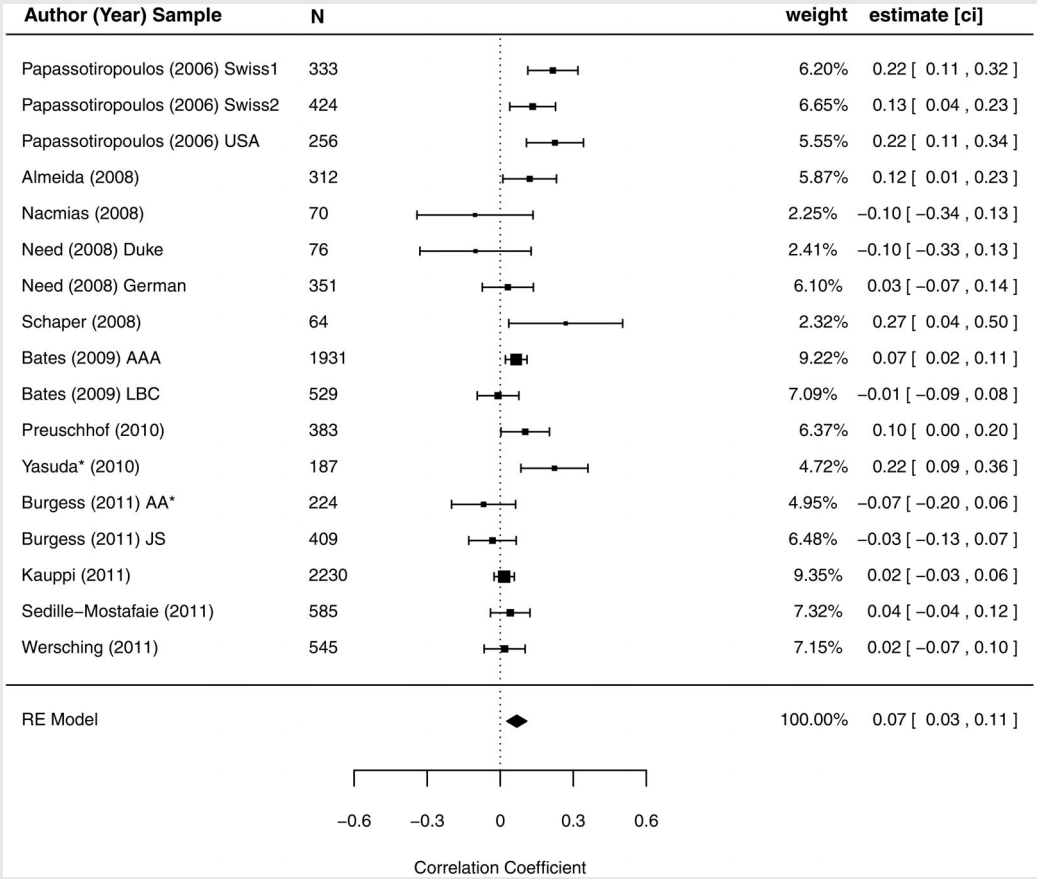


FIG. 1. Forest plot for the episodic memory tasks. Weight represents the percentage given to the observed outcome during model fitting. Stars (*) indicate samples with non-Caucasian ethnic background.

significant. Since random effects meta-analysis adjusts for large between study heterogeneity resulting in more conservative results and larger confidence intervals, the estimated effect sizes can only give hints for the range of the true effect sizes. Hence the provided power plots depict ranges of effect sizes.

Generally, between-study heterogeneity tends to increase for meta-analyses that comprise more samples [Ioannidis et al., 2007]. In the present meta-analysis, the between-study heterogeneity differs for the episodic and working memory meta-analyses, being higher for episodic memory. This difference diminishes yet remains present, if numbers of accounted studies become more equal. Study heterogeneity, which hampers replication [Moonesinghe et al., 2008], can be caused by a variety of factors, e.g. publication bias, selective report of calculated statistical models, possible presence of multiple functional variants in one gene, varying ethnic background including different linkage disequilibrium structure, phenotype correlation, differences in phenotype definition and experimental set up or gene-environment interactions. With the available data it was not possible to discriminate, whether the difference in homogeneity is a casual finding or has a systematic underlying effect. The most notable reduction of between-study heterogeneity could be achieved through the exclusion of the initial study samples, which reflects the Winner's Curse phenomenon. Exclusion of non-Caucasian samples did not markedly influence the level of heterogeneity. Given the amount of currently available data, we set aside further stratification of the samples. With more samples and larger sample sizes it might be feasible not only to test for possible sources of heterogeneity by sample stratification, but also to estimate sub-domain-specific effects as for example visual versus verbal memory. However, quantifying more general effects (e.g., episodic memory, working memory) comes with the appealing advantage that a more generalized interpretation of the underlying effect can be given.

The reported overall effect size estimates for *KIBRA* SNP rs17070145 may seem negligible considering the explained variance that can be attributed to this single variation. Yet it is important to stress here that the objective target of the GWAS approach is not to identify genetic variations that can explain the lion's share of phenotypic variation. Rather than that, this research approach offers the possibility to capitalize on naturally occurring common genetic variation to identify novel target molecules for a given complex trait of interest. Hence GWAS comes with the inherent

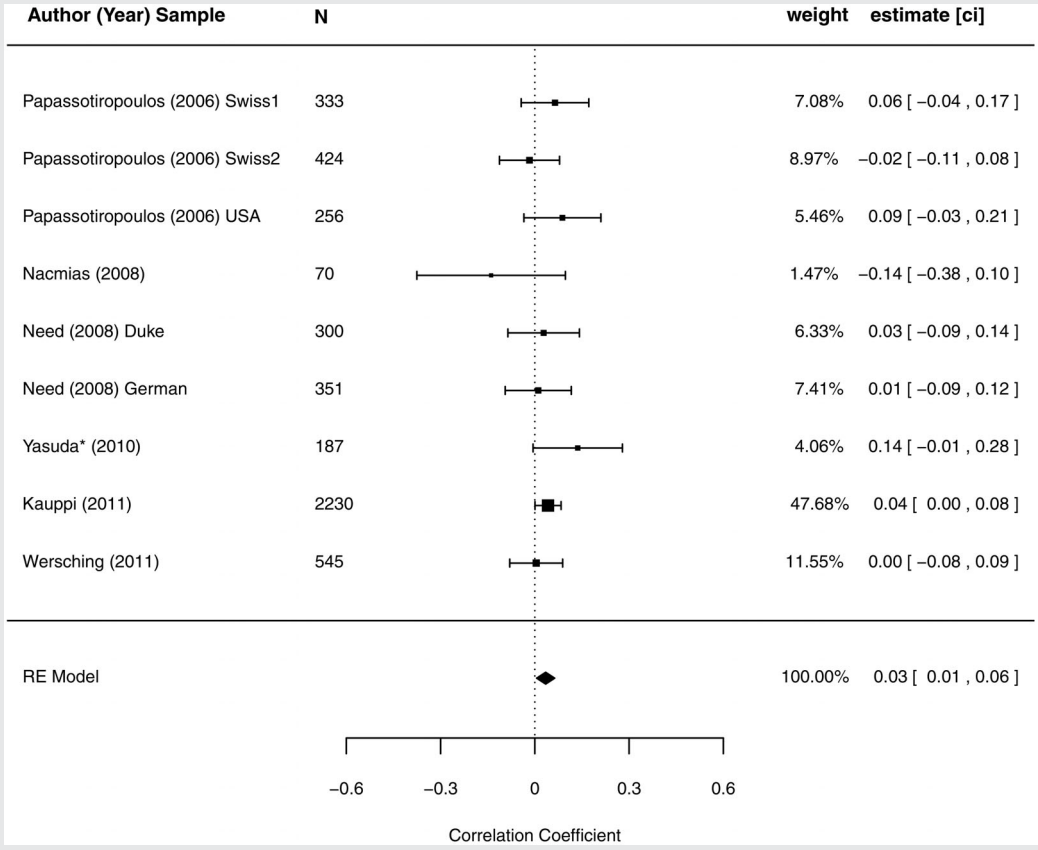


FIG. 2. Forest plot for the pure and narrow-sense working memory tasks. Weight represents the percentage given to the observed outcome during model fitting. Stars (*) indicate samples with non-Caucasian ethnic background.

strengths of a hypothesis-free, non-invasive and extremely flexible approach that is suitable to study the biology of complex human behavior. The scientific profit of GWAS is yielded by its potential to provide valuable starting points for a variety of research disciplines, which struggle together to elucidate the molecular underpinnings of human cognition and heritable disease. In the case of *KIBRA*, this is illustrated, for example, by the studies of Makuch et al. [2011], Corneveaux et al. [2010], and Burgess et al. [2011]. Makuch et al. [2011] described the important role of KIBRA for learning and memory in a knockout mouse model. Corneveaux et al. [2010] combined three different lines of evidence to extend the association of *KIBRA* from episodic memory performance to AD risk: Results from a gene expression study in brain tissue of AD-affected individuals, a brain imaging study (fluordesoxyglucose positron emission tomography) and an association analysis with AD disease status implicated a role for *KIBRA* in AD. In a meta-analytic approach, Burgess et al. [2011] further supported the finding of a potential effect on KIBRA on the risk to develop dementia. Thus these studies broaden the initial finding of an association with memory performance in healthy subjects to disease status. The regulatory role of KIBRA in AMPAR membrane trafficking may offer an explanation connecting the two findings, since it is known, for example, that beta-amyloid induces disruption of the AMPA signaling in AD [Shepherd and Huganir, 2007; Keifer and Zheng, 2010].

In summary, using a random effects meta-analytic approach, we find a significant association between SNP rs17070145 with episodic and working memory performance. *KIBRA* can be linked to learning and memory, for example, through its interaction with PKMzeta or its role in the regulation of AMPAR membrane trafficking [Büther et al., 2004; Schneider et al., 2010; Makuch et al., 2011], which is a key mechanism underlying synaptic plasticity. The estimated effect sizes that explain 0.1% of the variance for working memory and 0.5% for episodic memory are in the range of expected effect sizes for quantitative trait loci. Taken together the example of *KIBRA* illustrates the power of the GWAS approach to identify new target molecules related to human memory and to contribute to the elucidation of the molecular underpinnings of human cognition.

ACKNOWLEDGMENTS

This work was funded by the Swiss National Science Foundation (Sinergia Grants CRSIK0_122691 and CRSI33_130080 to D.Q. and

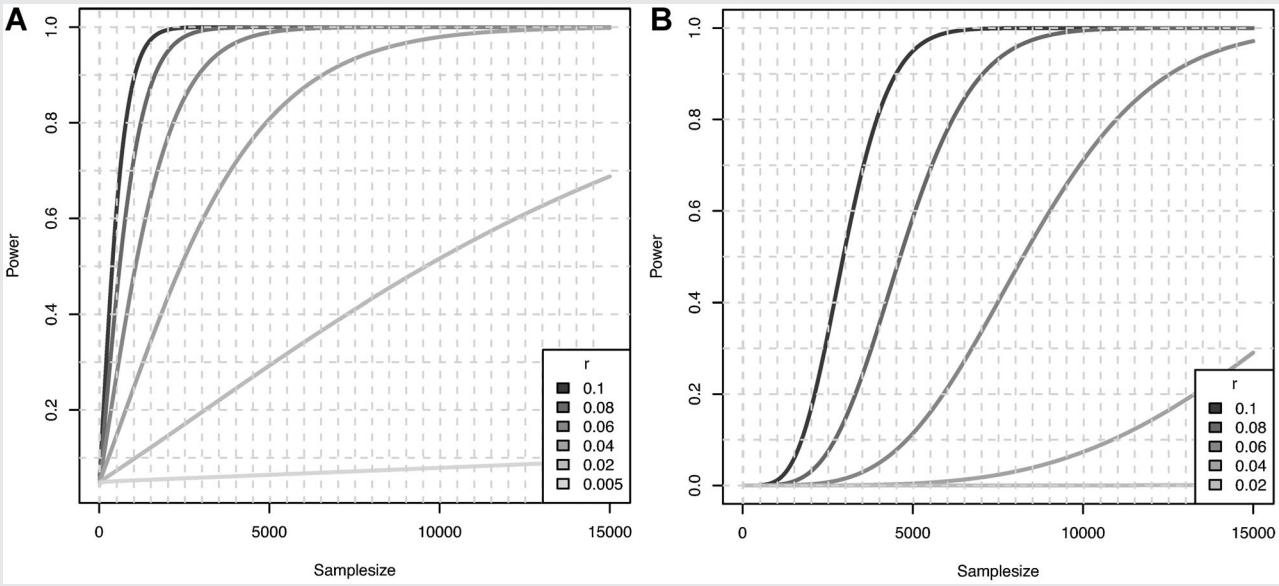


FIG. 3. Power analysis for the expected range of effect sizes, which illustrates the relation of sample sizes and expected effect size to power. A: Significance level is set to $P = 0.05$, which would be sufficient in the scenario of testing a single SNP. B: Significance level is set to $P = 5 \times 10^{-8}$ to correct for alpha-inflation in the case of 1×10^6 independent tests to keep a maximum alpha error rate of 5%.

A.P.) and by the 7th framework programme of the European Union (ADAMS Project, HEALTH-F4-2009-242257).

REFERENCES

Almeida OP, Schwab SG, Lautenschlager NT, Morar B, Greenop KR, Flicker L, Wildenauer D. 2008. KIBRA genetic polymorphism influences episodic memory in later life, but does not increase the risk of mild cognitive impairment. J Cell Mol Med 12:1672–1676.

Ando J, Ono Y, Wright MJ. 2001. Genetic structure of spatial and verbal working memory. Behav Genet 31:615–624.

Andreano JM, Cahill L. 2009. Sex influences on the neurobiology of learning and memory. Learn Mem 16:248–266.

Bates TC, Price JF, Harris SE, Marioni RE, Fowkes FGR, Stewart MC, Murray GD, Whalley LJ, Starr JM, Deary IJ. 2009. Association of KIBRA and memory. Neurosci Lett 458:140–143.

Bird CM, Burgess N. 2008. The hippocampus and memory: Insights from spatial processing. Nat Rev Neurosci 9:182–194.

Burgess JD, Pedraza O, Graff-Radford NR, Hirpa M, Zou F, Miles R, Nguyen T, Li M, Lucas JA, Ivnik RJ, Crook J, Pankratz VS, Dickson DW, Petersen RC, Younkin SG, Ertekin-Taner N. 2011. Association of common KIBRA variants with episodic memory and AD risk. Neurobiol Aging 32:557.e1–557.e9.

Büther K, Plaas C, Barnekow A, Kremerskothen J. 2004. KIBRA is a novel substrate for protein kinase Czeta. Biochem Biophys Res Commun 317: 703–707.

Calkins ME, Tepper P, Gur RC, Ragland JD, Klei L, Wiener HW, Richard J, Savage RM, Allen TB, O'Jile J, Devlin B, Kwentus J, Aliyu MH, Bradford LD, Edwards N, Lyons PD, Nimgaonkar VL, Santos AB, Go RCP, Gur RE. 2010. Project among African-Americans to explore risks for schizophrenia (PAARTNERS): Evidence for impairment and heritability of neurocognitive functioning in families of schizophrenia patients. Am J Psychiatry 167:459–472.

Champely S. 2009. pwr: Basic functions for power analysis. Available at. http://cran.r-project.org/package=pwr.

Corneveaux JJ, Liang WS, Reiman EM, Webster JA, Myers AJ, Zismann VL, Joshipura KD, Pearson JV, Hu-Lince D, Craig DW, Coon KD, Dunckley T, Bandy D, Lee W, Chen K, Beach TG, Mastroeni D, Grover A, Ravid R, Sando SB, Aasly JO, Heun R, Jessen F, Kölsch H, Rogers J, Hutton ML, Melquist S, Petersen RC, Alexander GE, Caselli RJ, Papassotiropoulos A, Stephan DA, Huentelman MJ. 2010. Evidence for an association between KIBRA and late-onset Alzheimer's disease. Neurobiol Aging 31:901–909.

DerSimonian R, Laird N. 1986. Meta-analysis in clinical trials. Control Clin Trials 7:177–188.

Drevets WC, Price JL, Furey ML. 2008. Brain structural and functional abnormalities in mood disorders: Implications for neurocircuitry models of depression. Brain Struct Funct 213:93–118.

Duning K, Schurek E-M, Schlüter M, Bayer M, Reinhardt H-C, Schwab A, Schaefer L, Benzing T, Schermer B, Saleem MA, Huber TB, Bachmann S, Kremerskothen J, Weide T, Pavenstädt H. 2008. KIBRA modulates directional migration of podocytes. J Am Soc Nephrol 19:1891–1903.

Finkel D, McGue M. 1998. Age differences in the nature and origin of individual differences in memory: A behavior genetic analysis. Int J Aging Hum Dev 47:217–239.

Finkel D, Pedersen N, McGue M. 1995. Genetic influences on memory performance in adulthood: Comparison of Minnesota and Swedish twin data. Psychol Aging 10:437–446.

Forbes NF, Carrick LA, McIntosh AM, Lawrie SM. 2009. Working memory in schizophrenia: A meta-analysis. Psychol Med 39:889–905.

Galecki P, Szemraj J, Florkowski A, Talarowska M, Bienkiewicz M, Galecka E, Lewinski A. 2010. Single nucleotide polymorphism of the KIBRA gene in recurrent depressive disorders. Neuro Endocrinol Lett 31:97–102.

Greenwood TA, Beeri MS, Schmeidler J, Valerio D, Raventós H, Mora-Villalobos L, Camacho K, Carrión-Baralt JR, Angelo G, Almasy L, Sano M, Silverman JM. 2011. Heritability of cognitive functions in families of successful cognitive aging probands from the Central Valley of Costa Rica. J Alzheimers Dis 27:897–907.

Hayashi N, Kazui H, Kamino K, Tokunaga H, Takaya M, Yokokoji M, Kimura R, Kito Y, Wada T, Nomura K, Sugiyama H, Yamamoto D, Yoshida T, Currais A, Soriano S, Hamasaki T, Yamamoto M, Yasuda Y, Hashimoto R, Tanimukai H, Tagami S, Okochi M, Tanaka T, Kudo T, Morihara T, Takeda M. 2010. KIBRA genetic polymorphism influences episodic memory in Alzheimer's disease, but does not show association with disease in a Japanese cohort. Dement Geriatr Cogn Disord 30: 302–308.

Hedges LV, Olkin I. 1985. Statistical methods for meta-analysis. Stat Med 20:825–840.

Herb A, Wisden W, Catania MV, Maréchal D, Dresse A, Seeburg PH. 1997. Prominent dendritic localization in forebrain neurons of a novel mRNA and its product, dendrin. Mol Cell Neurosci 8:367–374.

Ioannidis JPA. 2008. Why most discovered true associations are inflated. Epidemiology 19:640–648.

Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG. 2001. Replication validity of genetic association studies. Nat Genet 29:306–309.

Ioannidis JPA, Patsopoulos NA, Evangelou E. 2007. Heterogeneity in meta-analyses of genome-wide association investigations. PLoS ONE 2:e841.

Jacobsen LK, Picciotto MR, Heath CJ, Mencl WE, Gelernter J. 2009. Allelic variation of calsyntenin 2 (CLSTN2) modulates the impact of developmental tobacco smoke exposure on mnemonic processing in adolescents. Biol Psychiatry 65:671–679.

Johannsen S, Duning K, Pavenstädt H, Kremerskothen J, Boeckers TM. 2008. Temporal-spatial expression and novel biochemical properties of the memory-related protein KIBRA. Neuroscience 155:1165–1173.

Johansson B, Whitfield K, Pedersen NL, Hofer SM, Ahern F, McClearn GE. 1999. Origins of individual differences in episodic memory in the oldest-old: A population-based study of identical and same-sex fraternal twins aged 80 and older. J Gerontol B Psychol Sci Soc Sci 54:173–179.

Kauppi K, Nilsson L-G, Adolfsson R, Eriksson E, Nyberg L. 2011. KIBRA polymorphism is related to enhanced memory and elevated hippocampal processing. J Neurosci 31:14218–14222.

Keifer J, Zheng Z. 2010. AMPA receptor trafficking and learning. Eur J Neurosci 32:269–277.

Kraft P. 2008. Curses-winner's and otherwise-in genetic epidemiology. Epidemiology 19:649–651.

Kremerskothen J, Plaas C, Büther K, Finger I, Veltel S, Matanis T, Liedtke T, Barnekow A. 2003. Characterization of KIBRA, a novel WW domain-containing protein. Biochem Biophys Res Commun 300:862–867.

Kremerskothen J, Kindler S, Finger I, Veltel S, Barnekow A. 2006. Postsynaptic recruitment of dendrin depends on both dendritic mRNA transport and synaptic anchoring. J Neurochem 96:1659–1666.

Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, Willer CJ, Jackson AU, Vedantam S, Raychaudhuri S, Ferreira T, Wood AR, Weyant RJ, Segrè AV, Speliotes EK, Wheeler E, Soranzo N, Park J-H, Yang J, Gudbjartsson D, Heard-Costa NL, Randall JC, Qi L, Vernon Smith A, Mägi R, Pastinen T, Liang L, Heid IM, Luan J, Thorleifsson G, Winkler TW, Goddard ME, Sin Lo K, Palmer C, Workalemahu T, Aulchenko YS, Johansson A, Zillikens MC, Feitosa MF, Esko T, Johnson T, Ketkar S, Kraft P, Mangino M, Prokopenko I, Absher D, Albrecht E, Ernst F, Glazer NL, Hayward C, Hottenga J-J, Jacobs KB, Knowles JW, Kutalik Z, Monda KL, Polasek O, Preuss M, Rayner NW, Robertson NR, Steinthorsdottir V, Tyrer JP, Voight BF, Wiklund F, Xu J, Zhao JH, Nyholt DR, Pellikka N, Perola M, Perry JRB, Surakka I, Tammesoo M-L, Altmaier EL, Amin N, Aspelund T, Bhangale T, Boucher G, Chasman DI, Chen C, Coin L, Cooper MN, Dixon AL, Gibson Q, Grundberg E, Hao K, Juhani Junttila M, Kaplan LM, Kettunen J, König IR, Kwan T, Lawrence RW, Levinson DF, Lorentzon M, McKnight B, Morris AP, Müller M, Suh Ngwa J, Purcell S, Rafelt S, et al. 2010. Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature 467:832–838.

Makuch L, Volk L, Anggono V, Johnson RC, Yu Y, Duning K, Kremerskothen J, Xia J, Takamiya K, Huganir RL. 2011. Regulation of AMPA receptor function by the human memory-associated gene KIBRA. Neuron 71:1022–1029.

Marazziti D, Consoli G, Picchetti M, Carlini M, Faravelli L. 2010. Cognitive impairment in major depression. Eur J Pharmacol 626: 83–86.

Mesholam-Gately RI, Giuliano AJ, Goff KP, Faraone SV, Seidman LJ. 2009. Neurocognition in first-episode schizophrenia: A meta-analytic review. Neuropsychology 23:315–336.

Moonesinghe R, Khoury MJ, Liu T, Ioannidis JPA. 2008. Required sample size and nonreplicability thresholds for heterogeneous genetic associations. Proc Natl Acad Sci USA 105:617–622.

Nacmias B, Bessi V, Bagnoli S, Tedde A, Cellini E, Piccini C, Sorbi S, Bracco L. 2008. KIBRA gene variants are associated with episodic memory performance in subjective memory complaints. Neurosci Lett 436:145–147.

Nakagawa S, Cuthill IC. 2007. Effect size, confidence interval and statistical significance: A practical guide for biologists. Biol Rev Camb Philos Soc 82:591–605.

Need AC, Attix DK, McEvoy JM, Cirulli ET, Linney KN, Wagoner AP, Gumbs CE, Giegling I, Möller H-J, Francks C, Muglia P, Roses A, Gibson G, Weale ME, Rujescu D, Goldstein DB. 2008. Failure to replicate effect of Kibra on human memory in two large cohorts of European origin. Am J Med Genet B Neuropsychiatr Genet 147B:667–668.

Papassotiropoulos A, Stephan DA, Huentelman MJ, Hoerndli FJ, Craig DW, Pearson JV, Huynh K-D, Brunner F, Corneveaux J, Osborne D, Wollmer MA, Aerni A, Coluccia D, Hänggi J, Mondadori CRA, Buchmann A, Reiman EM, Caselli RJ, Henke K, de Quervain DJ-F. 2006. Common Kibra alleles are associated with human memory performance. Science 314:475–478.

Park DC, Lautenschlager G, Hedden T, Davidson NS, Smith AD, Smith PK. 2002. Models of visuospatial and verbal memory across the adult life span. Psychol Aging 17:299–320.

Pfeiffer T, Bertram L, Ioannidis JPA. 2011. Quantifying selective reporting and the proteus phenomenon for multiple datasets with similar bias. PLoS ONE 6:e18362.

Preuschhof C, Heekeren HR, Li S-C, Sander T, Lindenberger U, Bäckman L. 2010. KIBRA and CLSTN2 polymorphisms exert interactive effects on human episodic memory. Neuropsychologia 48:402–408.

R Development Core Team. 2011. R: A language and environment for statistical computing. Available at. http://www.r-project.org/.

Ranganath C. 2010. A unified framework for the functional organization of the medial temporal lobes and the phenomenology of episodic memory. Hippocampus 20:1263–1290.

Read S, Pedersen NL, Gatz M, Berg S, Vuoksimaa E, Malmberg B, Johansson B, McClearn GE. 2006. Sex differences after all those years? Heritability of cognitive abilities in old age. J Gerontol B Psychol Sci Soc Sci 61:137–143.

Rodríguez-Rodríguez E, Infante J, Llorca J, Mateo I, Sánchez-Quintana C, García-Gorostiaga I, Sánchez-Juan P, Berciano J, Combarros O. 2009. Age-dependent association of KIBRA genetic variation and Alzheimer's disease risk. Neurobiol Aging 30:322–324.

Samuelson KW. 2011. Post-traumatic stress disorder and declarative memory functioning: A review. Dialogues Clin Neurosci 13:346–351.

Sanna S, Jackson AU, Nagaraja R, Willer CJ, Chen W-M, Bonnycastle LL, Shen H, Timpson N, Lettre G, Usala G, Chines PS, Stringham HM, Scott LJ, Dei M, Lai S, Albai G, Crisponi L, Naitza S, Doheny KF, Pugh EW, Ben-Shlomo Y, Ebrahim S, Lawlor DA, Bergman RN, Watanabe RM, Uda M, Tuomilehto J, Coresh J, Hirschhorn JN, Shuldiner AR, Schlessinger D, Collins FS, Davey Smith G, Boerwinkle E, Cao A, Boehnke M, Abecasis GR, Mohlke KL. 2008. Common variants in the GDF5-UQCC region are associated with variation in human height. Nat Genet 40: 198–203.

Schaper K, Kolsch H, Popp J, Wagner M, Jessen F. 2008. KIBRA gene variants are associated with episodic memory in healthy elderly. Neurobiol Aging 29:1123–1125.

Schneider A, Huentelman MJ, Kremerskothen J, Duning K, Spoelgen R, Nikolich K. 2010. KIBRA: A new gateway to learning and memory? Front Aging Neurosci 2:2–4.

Sédille-Mostafaie N, Sebesta C, Huber KR, Zehetmayer S, Jungwirth S, Tragl KH, Fischer P, Krugluger W. 2011. The role of memory-related gene polymorphisms, KIBRA and CLSTN2, on replicate memory assessment in the elderly. J Neural Transm 119:77–80.

Shepherd JD, Huganir RL. 2007. The cell biology of synaptic plasticity: AMPA receptor trafficking. Annu Rev Cell Dev Biol 23:613–643.

Van Snellenberg JX. 2009. Working memory and long-term memory deficits in schizophrenia: Is there a common substrate? Psychiatry Res 174:89–96.

Vassos E, Bramon E, Picchioni M, Walshe M, Filbey FM, Kravariti E, McDonald C, Murray RM, Collier DA, Toulopoulou T. 2010. Evidence of association of KIBRA genotype with episodic memory in families of psychotic patients and controls. J Psychiatr Res 44: 795–798.

Viechtbauer MW. 2010. Conducting meta-analyses in R with the metafor package. J Stat Softw 36:1–48.

Visscher PM. 2008. Sizing up human height variation. Nat Genet 40: 489–490.

Volk HE, McDermott KB, Roediger HL, Todd RD. 2006. Genetic influences on free and cued recall in long-term memory tasks. Twin Res Hum Genet 9:623–631.

Weedon MN, Lettre G, Freathy RM, Lindgren CM, Voight BF, Perry JRB, Elliott KS, Hackett R, Guiducci C, Shields B, Zeggini E, Lango H, Lyssenko V, Timpson NJ, Burtt NP, Rayner NW, Saxena R, Ardlie K, Tobias JH, Ness AR, Ring SM, Palmer CNA, Morris AD, Peltonen L, Salomaa V, Davey Smith G, Groop LC, Hattersley AT, McCarthy MI, Hirschhorn JN, Frayling TM. 2007. A common variant of HMGA2 is associated with adult and childhood height in the general population. Nat Genet 39:1245–1250.

Wersching H, Guske K, Hasenkamp S, Hagedorn C, Schiwek S, Jansen S, Witte V, Wellmann J, Lohmann H, Duning K, Kremerskothen J, Knecht S, Brand E, Floel A. 2011. Impact of common KIBRA allele on human cognitive functions. Neuropsychopharmacology 36:1296–1304.

Wilson RS, Barral S, Lee JH, Leurgans SE, Foroud TM, Sweet RA, Graff-Radford N, Bird TD, Mayeux R, Bennett DA. 2011. Heritability of different forms of memory in the Late Onset Alzheimer's Disease Family Study. J Alzheimers Dis 23:249–255.

Yasuda Y, Hashimoto R, Ohi K, Fukumoto M, Takamura H, Iike N, Yoshida T, Hayashi N, Takahashi H, Yamamori H, Morihara T, Tagami S, Okochi M, Tanaka T, Kudo T, Kamino K, Ishii R, Iwase M, Kazui H, Takeda M. 2010. Association study of KIBRA gene with memory performance in a Japanese population. Biol Psychiatry 11:852–857.

Zhang H, Kranzler HR, Poling J, Gruen JR, Gelernter J. 2009. Cognitive flexibility is associated with KIBRA variant and modulated by recent tobacco use. Neuropsychopharmacology 34:2508–2516.

4.2 Substantial SNP-based heritability estimates for working memory performance

OPEN

Citation: Transl Psychiatry (2014) 4, e438; doi:10.1038/tp.2014.81

www.nature.com/tp

ORIGINAL ARTICLE

Substantial SNP-based heritability estimates for working memory performance

C Vogler[1,2,3,4,6], L Gschwind[1,5,6], D Coynel[5], V Freytag[1], A Milnik[1,2], T Egli[1], A Heck[1,2,3], DJ-F de Quervain[2,3,5] and A Papassotiropoulos[1,2,3,4]

Working memory (WM) is an important endophenotype in neuropsychiatric research and its use in genetic association studies is thought to be a promising approach to increase our understanding of psychiatric disease. As for any genetically complex trait, demonstration of sufficient heritability within the specific study context is a prerequisite for conducting genetic studies of that trait. Recently developed methods allow estimating trait heritability using sets of common genetic markers from genome-wide association study (GWAS) data in samples of unrelated individuals. Here we present single-nucleotide polymorphism (SNP)-based heritability estimates (h^2_{SNP}) for a WM phenotype. A Caucasian sample comprising a total of $N = 2298$ healthy and young individuals was subjected to an *N*-back WM task. We calculated the genetic relationship between all individuals on the basis of genome-wide SNP data and performed restricted maximum likelihood analyses for variance component estimation to derive the h^2_{SNP} estimates. Heritability estimates for three 2-back derived WM performance measures based on all autosomal chromosomes ranged between 31 and 41%, indicating a substantial SNP-based heritability for WM traits. These results indicate that common genetic factors account for a prominent part of the phenotypic variation in WM performance. Hence, the application of GWAS on WM phenotypes is a valid method to identify the molecular underpinnings of WM.

Translational Psychiatry (2014) **4,** e438; doi:10.1038/tp.2014.81; published online 9 September 2014

INTRODUCTION

Working memory (WM) has a pivotal role in human cognition and cognitive performance, allowing the integration of information from instantly perceived stimuli, long-term-memory and thought-processes. A wide variety of psychological testing procedures have been developed to render WM performance quantifiable.[1–4] Assessment of WM performance along with brain imaging techniques have been used to elucidate and delimit its different components, aiming at a more circumscribed understanding of the neuronal processes at the basis of WM.[5] The integration of extensive research on WM yields a common consensus that it constitutes a complex trait and that the buffer size for the transiently stored content varies interindividually, which is partly due to genetic factors. WM is an important endophenotype in neuropsychiatric research and its use in genetic association studies is thought to be a promising approach to increase our understanding of psychiatric disease.[6–11] Namely, two recent studies have corroborated the genetic link between schizophrenia and WM, demonstrating the validity of this endophenotype for schizophrenia research: Stefansson *et al.*[12] report that control subjects carrying Copy Number Variants, which predispose to schizophrenia and autism, perform at a level that is in between patients and population controls in cognitive tasks including a spatial WM test. A genome-wide gene set enrichment study conducted by our group, identified a set of voltage-gated cation channel activity genes that were robustly linked to performance in WM tasks in healthy individuals and also to risk for schizophrenia in a large case–control sample.[13] Both these studies suggest that the findings of cognitive deficits are translatable between healthy subjects and cohorts of psychiatric patients. Improving our understanding of the molecular basis of this endophenotype might be key for future drug discovery and treatment options in psychiatry.[14,15] As for any genetically complex trait, demonstration of sufficient heritability within the specific study context is a prerequisite for conducting genetic studies of that trait. Heritability is a concept that summarizes how much of the phenotypic variation in a given trait is attributable to heritable factors, the majority of them being genetic. Since inter-individual trait differences attributable to genetic variability are a prerequisite for quantitative trait loci mapping, estimation of trait heritability is important to demonstrate the validity of a quantitative trait loci study of a given trait. Conventionally, heritability estimates in humans are derived from phenotypic data by comparing correlations between relatives, where the extent of genetic relatedness is derived from the degree of relationship. Results from previous twin and family studies that used a variety of tasks to measure WM performance have shown heritability estimates ranging between 15 and 72%.[16–19] Recently developed methods propose inferring genetic identity from high-throughput SNP data and correlating these estimates with phenotypic resemblance among unrelated individuals.[20,21] This allows the estimation of an SNP-based heritability measure (h^2_{SNP}) for any specific genome-wide association study data set. Importantly, the h^2_{SNP} value quantifies the amount of phenotypic variation that can be

[1]Division of Molecular Neuroscience, Department of Psychology, University of Basel, Basel, Switzerland; [2]Psychiatric University Clinics, University of Basel, Basel, Switzerland; [3]Transfaculty Research Platform, University of Basel, Basel, Switzerland; [4]Department Biozentrum, Life Sciences Training Facility, University of Basel, Basel, Switzerland and [5]Division of Cognitive Neuroscience, Department of Psychology, University of Basel, Basel, Switzerland. Correspondence: Dr C Vogler, Division of Molecular Neuroscience, Department of Psychology, University of Basel, Birmannsgasse 8, CH-4055 Basel, Switzerland.
E-mail: (christian.vogler@unibas.ch)
[6]These two authors contributed equally to this work.
Received 23 May 2014; revised 11 July 2014; accepted 26 July 2014

Table 1. Descriptive statistics of task performance measures

	2-Back – 0-back	*2-Back*	*d-Prime*	*MRT 2-back*
Total ($N = 2298$)	−0.077 (0.07)	0.891 (0.07)	2.436 (0.84)	558.224 (128.12)
Females ($N = 1471$)	− 0.075 (0.07)	0.895 (0.07)	2.485 (0.83)	568.890 (131.09)
Males ($N = 827$)	− 0.082 (0.07)	0.885 (0.07)	2.347 (0.85)	539.252 (120.43)

Abbreviations: MRT, mean reaction time; WM, working memory. Performance metrics for WM (*N*-back derived) and RT in milliseconds: mean (s.d.).

explained by the common SNPs represented in genome-wide association study data sets. Of note, h^2_{SNP} hereby represents a lower bound heritability estimate, because causal variants that are neither genotyped nor tagged by the used set of markers are completely disregarded. In the present study we show that a substantial part of phenotypic variance of WM performance can be explained by using the common marker set represented on the Affymetrix 6.0 Human SNP array to estimate h^2_{SNP} values for *N*-back-derived WM phenotypes.

MATERIALS AND METHODS

Ethics statement

The experiments were approved by the ethics committee of the Cantons of Basel-City and Basel-Country. Written informed consent was obtained from all subjects before participation.

Participants and genotyping procedure

A sample of 2703 young healthy Swiss individuals (1721 females, 982 males; mean age: 22.5 years; median age: 22 years; range: 18–38 years) was assessed for WM performance using an *N*-back task. Saliva samples were obtained from each study participant, using an Oragene DNA Self-Collection Kit (DNA Genotek, Ottawa, Ontario, Canada). DNA was extracted from saliva using standard protocols. All subjects were individually genotyped using the Affymetrix Human SNP assay 6.0 according to the manufacturer's recommendation.

N-back task

All subjects completed the 0-back and 2-back version of the *N*-back task, after they were instructed and trained on the task. The 0-back condition served as a non-memory-guided control condition, measuring general attention, concentration and reaction time. The *N*-back task was presented on a computer screen and consisted of 12 blocks (six 0-back and six 2-back blocks). Per block, 14 stimuli consisting of 3 targets and 11 non-targets, were presented in a randomized order. Each stimulus was presented for 500 ms, followed by a black screen for 1500 ms. For each stimulus participants had to indicate as fast as possible whether it was a target or a non-target by pressing the corresponding button. At the beginning of each block the instruction was shown for 5 s. There was a 20 s break after every second block. The two blocks between the breaks consisted of a 0-back and a 2-back block, which were presented in a randomized order. Each block lasted for 33 s. The *N*-back task had a total duration of roughly 10 min, including instructions presented at the beginning of the task. Performance in the 0-back and the 2-back tasks was assessed by the mean correct response accuracy. The difference in mean accuracy between 2-back and the 0-back condition served as main phenotype (2-back mean accuracy attention corrected: 2-back – 0-back). We also tested the uncorrected 2-back mean accuracy and the d-prime for the 2-back condition. The d-prime was calculated according to the following formula:

$$d' = \text{Probit}(\text{hit rate}) - \text{Probit}(\text{false alarm rate})$$

The maximal number of hits (that is, correct responses to targets) was $N = 36$ and the maximal number of false alarms (that is, incorrect responses to non-targets) was $N = 132$. To allow for a probit-transformation, the hit and false-alarm-rates of 0 or 1 had to be avoided by replacing these extreme values, which was done as follows: 100% of hits were set to 1 – 1/36 and 0% were set to 1/36. Accordingly, 100% of false alarms were set to 1 – 1/132 and 0% of false alarms were set to 1/132.[22] In addition, we also estimated h^2_{SNP} for the mean reaction time of the 2-back task (2-back mean reaction time). See Table 1 for descriptive statistics of uncorrected task performance measures.

Genotypic outlier analysis

We identified and subsequently excluded subjects with technical biases or non-Caucasian ancestry. A Bayesian clustering algorithm was applied to genome-wide summary statistics to identify and exclude samples affected by a potential technical bias. Briefly, considering a combination of genome-wide call rate and heterozygosity rate, the algorithm infers each sample's posterior probability to belong to the outliers' class, as extreme values for these summary statistics may be indicative of a genotyping bias. Similarly, the Bayesian clustering algorithm was applied on the coordinates of the subjects on the first two PCA components inferred from HapMap data using the YRI, CEU and CHB-JPT populations. Thus, samples of non-Caucasian ancestry were identified and removed, excluding $N = 272$ individuals.[23]

Phenotypic and genetic quality controls

To adjust for sex and age, all *N*-back derived performance measurements (accuracy and reaction time) entered a linear model and residuals were extracted, which were used for all further calculations. As the presence of outliers in the phenotypic data can affect heritability estimation for complex traits,[24] all subjects whose performance measurements in any of the four *N*-back performance measurements differed more than 3.5 standard deviations from the sample mean were excluded ($N_{\text{excluded}} = 40$).

Due to this procedure, the exact same number of subjects was included in all subsequent analyses for all investigated phenotypes.

The genotype data set was filtered (MAF $\geqslant 0.01$; HWE *pFisher* $\geqslant 0.001$; per SNP call rate $\geqslant 95\%$; per individual call rate $\geqslant 95\%$) to obtain a set of 702'744 autosomal and 27'265 X-chromosomal SNPs for further analysis.

Estimation of the pairwise genetic similarity using all autosomal markers and subsequently removing one of each pairs showing genetic relatedness > 0.025 (~ second-degree cousins) led to the exclusion of $N = 93$ subjects, yielding a sample size of $N = 2298$ for h^2_{SNP} estimation.

h^2_{SNP} Estimation

To obtain the genome-wide heritability estimates h^2_{SNP}, we calculated the genetic relationship between all individuals on the basis of the autosomal SNP data and performed restricted maximum likelihood analyses for variance component estimation using the GCTA software package.[20] In a second step, h^2_{SNP} estimates were calculated for all chromosomes separately including the X-chromosome. The genetic relationship matrix for the X-chromosome was obtained using the designated option in the GCTA toolset.

RESULTS

Heritability estimates for the three 2-back derived WM performance measures on the basis of autosomal chromosomes ranged between 31 and 41%, indicating substantial SNP-based heritabilities (Figure 1). The proportion of phenotypic variance explained by all autosomal SNPs for difference in response accuracy between the 2-back and the 0-back condition was 41% (s.e. = 0.139; *P*-value = 0.0008). h^2_{SNP} for the mean response accuracy in the 2-back alone (that is, without correction for 0-back performance) and the false-response corrected d-prime condition were 31% with s.e. = 0.138; *P*-value = 0.006 and s.e. = 0.140; *P*-value = 0.01, respectively. Due to several reports linking processing speed to WM-related cognitive abilities,[25,26] we also investigated heritability of the mean reaction time in the 2-back condition (h^2_{SNP}: 24%; s.e. = 0.142; *P*-value = 0.04). A

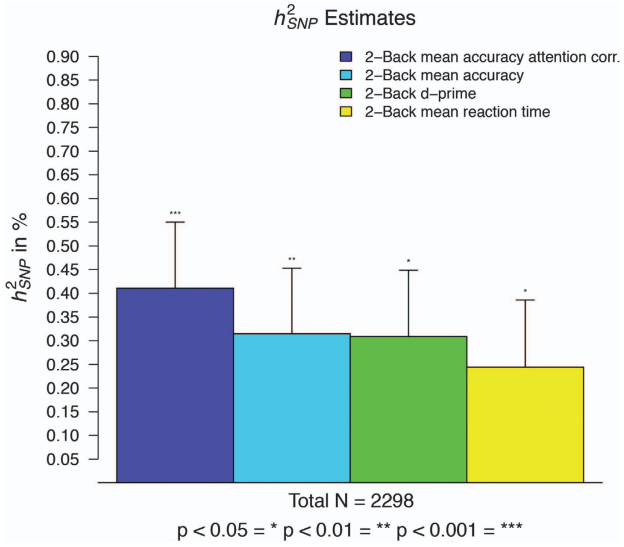


Figure 1. h^2_{SNP} estimates for 2-back derived WM measures in percent for a total of $N = 2298$. The error bars represent standard errors. 2-Back mean accuracy attention corrected (corr.) h^2_{SNP} 41%, s.e. 14%, P-value = 0.0008; 2-back mean accuracy h^2_{SNP} 31%, s.e. 14%, P-value = 0.006; 2-back d-prime h^2_{SNP} 31%, s.e. 14%, P-value = 0.01; 2-back mean reaction time h^2_{SNP} 24%, s.e. 14%, P-value = 0.04. SNP, single-nucleotide polymorphism; WM, working memory.

descriptive analysis of correlations between phenotypes shows strong intercorrelations between the three accuracy-related 2-back WM phenotypes ($r^2 > 0.65$), whereas correlation coefficients between the mean reaction time in the 2-back condition and the WM phenotypes are negligible ($r^2 < 0.01$) by means of shared phenotypic variance (Table 2).

Chromosomewise h^2_{SNP} estimates

In a further step, we partitioned heritability into contributions from single chromosomes, including the X-chromosome. Figure 2 shows the h^2_{SNP} values per chromosome, presenting an approximation to the amount of phenotypic variance explained by the single chromosomes. For each phenotype, we derived a vector containing h^2_{SNP} estimates per chromosome. To investigate whether the correlational pattern of the phenotypes is reflected also in a similar chromosomal heritability profile, we calculated the Pearson's correlation coefficients for these vectors. The highly correlated WM phenotypes show a similar profile for the chromosomewise h^2_{SNP} estimates. Regarding the single chromosome h^2_{SNP} values for the mean reaction time in the 2-back condition, we observe a different distribution of h^2_{SNP} estimates per chromosome. Table 3 depicts the descriptive Pearson's correlation coefficients between the chromosomewise h^2_{SNP} estimates for all phenotypes. We also investigated whether the chromosomal length and the number of SNPs are correlated with the phenotypic variance that each chromosome explains. We observed significant correlations (Pearson's r in the range between 0.53 and 0.68) for the WM phenotypes with the chromosomal length and number of SNPs, whereas no significant correlation was observed for the mean reaction time of the 2-back task (See Table 4).

DISCUSSION

The present data demonstrate that a substantial proportion of variance in WM performance is captured by common genome-wide association study SNPs in a sample of healthy young unrelated individuals. The h^2_{SNP} estimates for the N-back-derived WM phenotypes range from 31 to 41% with a mean standard error of 0.14. Thus, the h^2_{SNP} estimates are consistent with the previously reported heritability for WM on the basis of twin-studies.[27] Taken together, these findings add further support for the hypothesis that WM is a highly heritable trait. The chromosomal partitioning analysis (Figure 2), which depicts h^2_{SNP} estimates for all chromosomes analyzed individually, provides a clear hint for the pronounced polygenicity of the investigated WM phenotype. This finding is in line with the previously reported ubiquitious polygenicity of human complex traits.[28] The amount of genetic variation that explains the variance of WM performance is proportional to the chromosomal length and the number of investigated SNPs per chromosome (see Table 4). The correlational pattern for the investigated phenotypes (see Table 2) indicates that the WM performance measures and the mean reaction time are independent behavioral metrics. The h^2_{SNP} estimate based on genome-wide data (see Figure 1) are lower for the mean reaction time and the h^2_{SNP} estimates per chromosome also show a different distribution compared with the WM performance derived estimates (see Figure 2). Furthermore, although there is an observable trend for the proportionality of h^2_{SNP} estimates with chromosomal length, the correlation fails to reach statistical significance (see Table 2), which may be due to a lower overall heritability of this trait. Hence, we conclude that the WM performance measured with the 2-back task is independent of the mean response time under the cognitive load of performing the 2-back test.

Table 2. Correlation of phenotypes

	2-Back – 0-back	*2-Back*	*d-Prime*
2-Back – 0-back			
2-Back	0.91***		
d-Prime	0.83***	0.94***	
MRT 2-back	– 0.08***	0.01	0.06**

Abbreviation: MRT, mean reaction time. Pearson's correlation coefficients, df = 2296; *$P < 0.05$, **$P < 0.005$, ***$P < 0.0005$.

Of note, the heritability estimation for complex traits using genome-wide data sets is rather a complement than a substitution to studies on twin- and family-based heritability. Marker-based heritability estimation represents a lower bound for the true trait heritability as it relies only on the effects of common variants assuming an additive variance model. On one hand, the investigation of large pools of unrelated individuals allows to assess heritability without the undermining effects due to shared environment or familiality. Yet, on the other hand, it will not take the potential effects of rare variants into account. The SNP-based heritability estimates for WM suggest that a large share of phenotypic variance can be explained by common SNPs rendering well-powered genome-wide association study data sets a promising tool to discover molecular players that act in concert to form this complex trait. It has been repeatedly shown that familial risk for psychiatric diseases like schizophrenia is often accompanied by reduced cognitive abilities.[29–31] Results from a long-term study suggest that people who reported psychotic-like experiences in late adulthood performed poorer in cognitive tasks during childhood and adolescence.[32] In a recent twin study investigating healthy twins, Goldberg *et al.*[6] report that the phenotypic correlation between intelligence quotient and WM can be almost entirely attributed to shared genetic variance. In an effort to investigate dimensions of observable behavior and neurobiological measures that can be used to classify psychopathology, the National Institute of Mental Health established the Research Domain Criteria project. The N-back task used in the present study

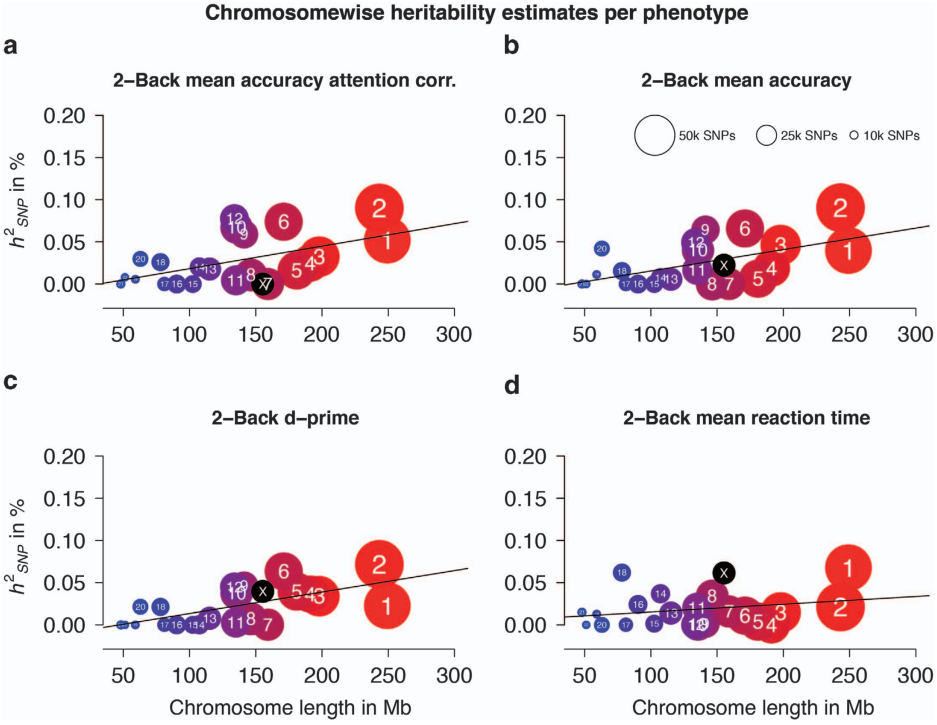


Figure 2. Chromosomewise h^2_{SNP} estimates for *N*-back phenotypes and mean reaction time h^2_{SNP} estimates per chromosome are plotted on the y axis. The x axis depicts the chromosomal length in Mb. Circle size indicates the number of SNPs used in the analysis after quality control. Color scheme from red to blue reflects the chromosomal ordering from chromosome 1 to 22, with the X-chromosome colored in black. The black line depicts the regression line for chromosomal length and explained variance (regression coefficients (df = 21): (**a**) $r = 0.53$, *P*-value = 0.009, (**b**) $r = 0.56$, *P*-value = 0.006, (**c**) $r = 0.65$, *P*-value = 0.0009, (**d**) $r = 0.25$, *P*-value = 0.3). SNP, single-nucleotide polymorphism.

Table 3. Correlation of chromosome h^2_{SNP} vectors

	2-Back – 0-back	*2-Back*	*d-Prime*
2-Back – 0-back			
2-Back	0.91***		
d-Prime	0.83***	0.87***	
MRT 2-back	– 0.09	– 0.04	–0.04

Abbreviations: MRT, mean reaction time; SNP, single-nucleotide polymorphism. Pearson's correlation coefficients, df = 21; *$P < 0.05$, **$P < 0.005$, ***$P < 0.0005$.

Table 4. Correlation of chromosomal h^2_{SNP} estimates with chromosomal length and number of SNPs per chromosome

	2-Back – 0-back	*2-Back*	*d-Prime*	*MRT 2-back*
Chromosomal length	0.53*	0.56*	0.65**	0.25
Number of SNPs	0.61**	0.60**	0.68***	0.15

Abbreviations: MRT, mean reaction time; SNP, single-nucleotide polymorphism. Pearson's correlation coefficients, df = 21; *$P < 0.05$, **$P < 0.005$, ***$P < 0.0005$.

is deemed appropriate to measure the sub-constructs of WM of active maintenance, limited capacity and with some, albeit not definitive evidence, for flexible updating according to the Research Domain Criteria project. Given the implication of cognitive and especially WM deficits in schizophrenia and other psychiatric disorders, the genetically guided decomposition of WM-related molecular pathways might pave the way for a better understanding of the genetic architecture implicated in these mental disorders.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

This work was supported by the Swiss National Science Foundation (Sinergia Grants CRSI33_130080 and CRSII1_136227 to AP and DJ-FdeQ.).

REFERENCES

1 Barch DM, Moore H, Nee DE, Manoach DS, Luck SJ. CNTRICS imaging biomarkers selection: working memory. *Schizophr Bull* 2012; **38**: 43–52.
2 Daneman M, Carpenter P. Individual differences in working memory and reading. *J Verbal Learn Verbal Behav* 1980; **19**: 450–466.
3 Gur RC, Richard J, Hughett P, Calkins ME, Macy L, Bilker WB *et al.* A cognitive neuroscience-based computerized battery for efficient measurement of individual differences: standardization and initial construct validation. *J Neurosci Methods* 2010; **187**: 254–262.
4 Kirchner WK. Age differences in short-term retention of rapidly changing information. *J Exp Psychol* 1958; **55**: 352–358.
5 Rissman J, Wagner AD. Distributed representations in memory: insights from functional brain imaging. *Annu Rev Psychol* 2012; **63**: 101–128.
6 Goldberg X, Alemany S, Rosa A, Picchioni M, Nenadic I, Owens SF *et al.* Substantial genetic link between iq and working memory: implications for molecular genetic studies on schizophrenia. the european twin study of schizophrenia (EUTwinsS). *Am J Med Genet B Neuropsychiatr Genet* 2013; **162**: 413–418.
7 Snitz BE, Macdonald AW, Carter CS. Cognitive deficits in unaffected first-degree relatives of schizophrenia patients: a meta-analytic review of putative endophenotypes. *Schizophr Bull* 2006; **32**: 179–194.
8 Gottesman II, Gould TD. The endophenotype concept in psychiatry: etymology and strategic intentions. *Am J Psychiatry* 2003; **160**: 636–645.
9 Gould TD. Gottesman II. Psychiatric endophenotypes and the development of valid animal models. *Genes Brain Behav* 2006; **5**: 113–119.
10 Dick DM, Jones K, Saccone N, Hinrichs A, Wang JC, Goate A *et al.* Endophenotypes successfully lead to gene identification: results from the collaborative study on the genetics of alcoholism. *Behav Genet* 2006; **36**: 112–126.

11 Murrough JW, Iacoviello B, Neumeister A, Charney DS, Iosifescu D V. Cognitive dysfunction in depression: neurocircuitry and new therapeutic strategies. *Neurobiol Learn Mem* 2011; **96**: 553–563.
12 Stefansson H, Meyer-Lindenberg A, Steinberg S, Magnusdottir B, Morgen K, Arnarsdottir S *et al.* CNVs conferring risk of autism or schizophrenia affect cognition in controls. *Nature* 2014; **505**: 361–366.
13 Heck A, Fastenrath M, Ackermann S, Auschra B, Bickel H, Coynel D *et al.* Converging genetic and functional brain imaging evidence links neuronal excitability to working memory, psychiatric disease, and brain activity. *Neuron* 2014; **81**: 1203–1213.
14 Papassotiropoulos A, de Quervain DJ-F. Genetics of human episodic memory: dealing with complexity. *Trends Cogn Sci* 2011; **15**: 381–387.
15 Hyman SE. Psychiatric drug development: diagnosing a crisis. *Cerebrum* 2013; **2013**: 5.
16 Vinkhuyzen AAE, van der Sluis S, Boomsma DI, de Geus EJC, Posthuma D. Individual differences in processing speed and working memory speed as assessed with the Sternberg memory scanning task. *Behav Genet* 2010; **40**: 315–326.
17 Lee T, Mosing MA, Henry JD, Trollor JN, Ames D, Martin NG *et al.* Genetic influences on four measures of executive functions and their covariation with general cognitive ability: the Older Australian Twins Study. *Behav Genet* 2012; **42**: 528–538.
18 Karlsgodt KH, Kochunov P, Winkler AM, Laird AR, Almasy L, Duggirala R *et al.* A multimodal assessment of the genetic control over working memory. *J Neurosci* 2010; **30**: 8197–8202.
19 Wilson RS, Barral S, Lee JH, Leurgans SE, Foroud TM, Sweet RA *et al.* Heritability of different forms of memory in the Late Onset Alzheimer's Disease Family Study. *J Alzheimers Dis* 2011; **23**: 249–255.
20 Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 2011; **88**: 76–82.
21 Yang J, Lee SH, Goddard ME, Visscher PM. Genome-wide complex trait analysis (GCTA): methods, data analyses, and interpretations. *Methods Mol Biol* 2013; **1019**: 215–236.
22 Stanislaw H, Todorov N. Calculation of signal detection theory measures. *Behav Res Methods Instrum Comput* 1999; **31**: 137–149.
23 Bellenguez C, Strange A, Freeman C, Donnelly P, Spencer CC, Wellcome Trust Case Control Consortium. A robust clustering algorithm for identifying problematic samples in genome-wide association studies. *Bioinformatics* 2012; **28**: 134–135.
24 Hayat Y, Yang J, Xu H, Zhu J. Influence of outliers on QTL mapping for complex traits. *J Zhejiang Univ Sci B* 2008; **9**: 931–937.
25 Baker LA, Vernon PA, Ho HZ. The genetic correlation between intelligence and speed of information processing. *Behav Genet* 1991; **21**: 351–367.
26 Ogata S, Kato K, Honda C, Hayakawa K. Common genetic factors influence hand strength, processing speed, and working memory. *J Epidemiol* 2013; **24**: 31–38.
27 Ando J, Ono Y, Wright MJ. Genetic structure of spatial and verbal working memory. *Behav Genet* 2001; **31**: 615–624.
28 Yang J, Lee T, Kim J, Cho M-C, Han B-G, Lee J-*et al.* Ubiquitous polygenicity of human complex traits: genome-wide analysis of 49 traits in Koreans. *PLoS Genet* 2013; **9**: e1003355.
29 Wiener H, Klei L, Calkins M, Wood J, Nimgaonkar V, Gur R *et al.* Principal components of heritability from neurocognitive domains differ between families with schizophrenia and control subjects. *Schizophr Bull* 2013; **39**: 464–471.
30 Toulopoulou T, Goldberg TE, Mesa IR, Picchioni M, Rijsdijk F, Stahl D *et al.* Impaired intellect and memory: a missing link between genetic risk and schizophrenia? *Arch Gen Psychiatry* 2010; **67**: 905–913.
31 Conklin HM, Curtis CE, Calkins ME, Iacono WG. Working memory functioning in schizophrenia patients and their first-degree relatives: cognitive functioning shedding light on etiology. *Neuropsychologia* 2005; **43**: 930–942.
32 Barnett JH, McDougall F, Xu MK, Croudace TJ, Richards M, Jones PB. Childhood cognitive function and adult psychopathology: associations with psychotic and non-psychotic symptoms in the general population. *Br J Psychiatry* 2012; **201**: 124–130.

4.3 Common epigenetic variation in a European population of mentally healthy young adults

Title Page

Common epigenetic variation in a European population of mentally healthy young adults

Authors

Annette Milnik[1,2,3], Christian Vogler[1,2,3], Philippe Demougin[2,4], Tobias Egli[1,2], Virginie Freytag[1,2], Angela Heck[1,2,3], Fabian Peter[1,2,4], Klara Spalek[2,5], Dominique J.-F. de Quervain[2,3,5], Andreas Papassotiropoulos[1,2,3,4], Vanja Vukojevic[1,2,4]

Affiliations

1 Division of Molecular Neuroscience, Department of Psychology, University of Basel, CH-4055 Basel, Switzerland

2 Transfaculty Research Platform Molecular and Cognitive Neurosciences, University of Basel, CH-4055 Basel, Switzerland

3 Psychiatric University Clinics, University of Basel, CH-4055 Basel, Switzerland

4 Department Biozentrum, Life Sciences Training Facility, University of Basel, CH-4056 Basel, Switzerland

5 Division of Cognitive Neuroscience, Department of Psychology, University of Basel, CH-4055 Basel, Switzerland

Short title

Common epigenetic variation

Abstract

DNA methylation represents an important link between structural genetic variation and complex phenotypes. The study of genome-wide CpG methylation and its relation to traits relevant to psychiatry has become increasingly important. Here, we analyzed quality metrics of 396'833 CpG sites in a population of 568 mentally healthy young adults. For 25 % of all CpGs we observed medium to large common epigenetic variation. These CpGs were overrepresented in open sea and shore regions, as well as in intergenic regions. Furthermore, methylation quantitative trait loci (meQTLs) were strongly enriched for the most-variable CpGs. This suggests that a significant proportion of common DNA methylation is at least partially genetically driven and may be observed similarly across tissues. These findings are of particular relevance for studies of complex neuropsychiatric traits, which often rely on proxy tissues.

Text

Phenotypic differences between individuals are only partially explained by genetic differences. Additional sources of variation, including epigenetic regulation, are also centrally involved in trait variability and disease etiology (1). Of the various epigenetic mechanisms, DNA methylation is being studied most extensively, because the available technology allows for the investigation of methylation patterns at both high resolution and –throughput (2). DNA methylation represents an important connection between structural genetic variation and complex phenotypes (3), a link that can be investigated by epigenome-wide association studies (EWAS) (4; 5). However, EWAS are challenged by the high complexity of the methylation signal, which displays variability related to such factors as the population under study, tissue and temporal dynamics. Further such signals consist of varying amounts of measurement error and systematic variance of no interest (like technical artifacts) (6–9).

In a population under study, replication analyses allow to estimate a trait's naturally occurring variation (i.e., signals' variance exceeding the technical variance and random noise (10)). Within the scope of high-density biological data, replication analyses can be done on different levels. Typically, analysis of technical replicates for CpG methylation data is done on a methylome-wide level by comparing overall methylation profiles (11). On a methylome-wide level, high reproducibility within and between technologies has been repeatedly shown (2; 12). However, association studies such as EWAS are done on single-CpG level. Therefore, we focused herein on the assessment of epigenetic variation in each single-CpG by evaluating the replication analysis on single-CpG level.

Owing to the complexity of the CpG methylation signal and to the underlying dynamic regulation mechanisms, results from methylation association studies are challenged by interpretational difficulties, especially in cross-sectional designs (13; 14). With respect to that, the identification of most-variable and reproducible CpGs in a sample of healthy young adults may serve as a baseline for the assessment of common epigenetic variation and as a guiding reference for other studies. Importantly, such a reference contributes to the identification of

CpGs that show common epigenetic variation as opposed to CpGs for which epigenetic differences emerge in the course of disease (15).

Here we performed comprehensive reliability analyses of the CpG methylation signal by utilizing the widely used Illumina Infinium Human Methylation 450K array in a large sample of mentally healthy young adults. We also conducted a series of association studies with CpG methylation as the dependent variable for validation purpose. As expected, technical replication analyses on a methylome-wide level showed high reproducibility for human DNA derived from blood. On single-CpG level a considerable part of CpGs showed naturally occurring variation (see supplementary table 1). The most-variable CpGs tended to be also the most reliable ones. Furthermore, the most-variable CpGs were overrepresented in the open sea and the shore of CpG islands as well as in intergenic regions. They were also more likely to show significant association signals when investigating methylation quantitative trait loci (meQTL) and calculating associations with age and sex. This natural epigenetic variation may be further studied in EWAS, such as those related to neuropsychiatric traits.

Methods and Material

Subjects

The subjects included in this study (*N* = 568; mean age 23.8y, 18.3 – 38.8; 59 % females) represent a subset of an ongoing study, which was previously described e.g. in (16; 17). All participants received general information about the study and gave their written informed consent. The ethics committee of the Cantons of Basel-Stadt and Basel-Landschaft approved the experiment. For additional information regarding the general study design see the supplementary material.

Affymetrix SNP 6.0 based genotyping

SNP genotyping for all samples was performed as described in the Genome-Wide Human SNP Nsp/Sty 6.0 User Guide (Affymetrix, Santa Clara, CA USA; see supplemental material). The mean call-rate per subject was 98.7 % (90.1 % – 99.7 %). After outlier detection and pruning, we kept *N* = 533 subjects and *N* = 185,145 SNPs for the association analyses (see supplementary material).

HumanMethylation Infinium 450K BeadChip based methylation analyses

Microarray-based DNA methylation analysis was performed at ServiceXS (ServiceXS B.V., Leiden, The Netherlands) on the HumanMethylation450 BeadChip (Illumina, Inc., San Diego, CA, U.S.A). Preprocessing was done independently for the two batches of the main sample (*N* = 568) and the two batches of the technical replicates sample (*N* = 145; for additional information see supplementary material). Data were extracted and analyzed from the generated idat files using the R package RnBeads version 0.99.9 (21). CpG annotation was based on the manufactures annotation file (HumanMethylation450_15017482_v.1.2). During preprocessing, the background was subtracted using the "noob" method in the methylumi package (22). We used the SWAN algorithm (23) for normalization.

The following probe categories were excluded from the final data sets, based on the annotation provided within the RnBeads package: non-CpG context probes, probes with a SNP mapping directly to the target CpG site, as well as probes with three and more SNPs mapping within the 50mer probe (MAF threshold was set to 0.01; N = 18,998 CpGs; see Figure 3), gonosomal probes (N = 11,473 CpGs), non-specific probes. Using the Greedycut algorithm, we iteratively removed the probes and data sets of the highest impurity ($p < 0.05$).

Postprocessing was done for each sample (main sample and technical replicates sample) separately, combining the β-values of the preprocessed data for the two available batches per sample. The β-values were processed step-by-step in order to correct for further influential factors as follows: 1) using logit-transformation (M-value, (24); done with the R-package car (25)); 2) z-transformation per plate (correcting for plate and batch effects); 3) regressing out the first 8 axis of a principal component analysis (PCA). The PCA was based on all subjects per sample and all CpGs with no missing values (>95% of the included CpGs; done with the R-package pcaMethods (26)).

The accepted missing rate per CpG was set to < 5%. If not already excluded before, we further excluded probes associated to sex chromosomes, polymorphic probes or multi-mapping probes (N_{max} = 63'974) based on the annotation files provided from Price (27) and Chen (28). Only CpGs surviving all filtering steps in both samples were used (N = 396'833).

We performed a cross-platform validation for the genotyping and methylation data (see supplementary material).

Technical replication analyses

We used Pearson correlation coefficients for all analyses. On methylome level, we calculated pairwise comparisons for all available CpGs based on the β-values. These pairwise comparisons were done for DNA derived from the identical (one comparison for each subject, N = 145 comparisons in total; within-subject comparison) or from differing DNA (288 comparisons for each subject, N = 41'760 comparisons in total; between-subject comparison).

On single-CpG level we did pairwise comparisons separately for each CpG, which resulted in N = 396'833 correlation coefficients. This analysis was done after applying all pre and postprocessing steps. To further evaluate the distribution of the N = 396'833 r-values, we applied a Gaussian fit allowing up to 5 overlapping Gaussian distributions by using the optimx function in R (settings: method L-BFGS-B, ndeps 0.0001, maxit 40'000). The minimum and maximum values were restricted to $-1 \geq m \leq 1$, $0 \geq sd \leq 1$ and $0 \geq p \leq 1$. The starting values were as follows: m = c(0, 0.5, 0.6, 0.7, 0.8); sd = 0.12; p = c(0.6, 0.1, 0.1, 0.1, 0.1).

To obtain a random r-distribution for a sample size of N = 145 subjects, we repeatedly (N = 400'000 times) generated two standard normal random variables (length of 145 each) and calculated r between these two variables.

Association analyses

The CpG methylation data used for the association analyses was based on all pre and postprocessing steps. The SNP-association studies were performed on a genome-wide scale for all N = 396'833 CpGs separately assuming an additive genetic model and applying an epi-genome-wide Bonferroni correction ($\alpha = 0.1\%$, correcting for 1,000,000 x 396'833 tests, resulting in $p_{bonf} < 2.5 \times 10^{-15}$). For the *cis*-analyses, we used a less stringent significance threshold $p_{cis} < 1 \times 10^{-5}$ ($\alpha = 1\%$, correcting for at least 1,000 independent tests per CpG).

To evaluate the associations with sex and age, we calculated a single linear model for each CpG, including sex and age as independent variables (per independent variable $\alpha = 5\%$, 396'833 independent tests, $p < 1.3 \times 10^{-7}$).

Software

If not mentioned differently, analyses were conducted in R (version: 2.15.1 and higher, R Development Core Team 2012) or PLINK (18).

Results

In the current study, we investigated the methylome and single CpG signal in a large sample (N = 568; main sample) of healthy young adults. For N = 145 subjects, a technical replication was available, based on the identical DNA derived from blood (technical replicates sample). The two datasets of these subjects were used for all reliability analyses. The main sample and technical replicates sample were processed independently, starting from identical DNA specimens. Based on SNP-data, N = 533 subjects out of the main sample were identified as a homogenous sub-sample of European ancestry (association sample; see supplementary material), and thus were used for all association analyses that were performed for validation.

Methylome-wide reliability analyses

Based on the β-distribution on methylome-wide level (Figure 1A), we observed high signal reproducibility of technical replicates (N = 145 pairs, average r = 0.997, r_{min} = 0.990, r_{max} = 0.999; see Figure 1B for one example). This was in agreement with previous reports (2; 12). However, estimated signal similarity between DNA of different subjects (N = 41'905 pairs of between-subject comparisons), also suggested high signal consistency on the methylome-wide level (average r = 0.994, r_{min} = 0.985, r_{max} = 0.997; see Figure 1C for one example). For each of the datasets separately we tested, which of the remaining N = 289 datasets showed the highest similarity on methylome-wide level. In 91 % of these comparisons, the highest similarity was obtained for its technical replicate, thus the identical DNA.

Estimation of natural occurring variation on single CpG level

With the replication analysis on single-CpG level we aimed to detect the most-variable CpGs in the investigated population. Given the sample size of N = 145, we were adequately powered (95 %) on a nominal significance level (5 %) to detect a reliability r = 0.29, which corresponds to a natural epigenetic variation of 8.6 %.

Taking all N = 396'833 CpGs into account, the average r per single-CpG was 0.191 (r_{min} = -0.361, r_{max} = 0.988). Figure 2A shows the distribution of r-values across all N = 396'833 CpGs. To further examine this distribution, we applied a Gaussian mixture model allowing up to five Gaussian distributions (Figure 2A). The result of the Gaussian mixture model indicated that for approx. 53 % of the CpGs the observed signal variability was mainly based on random or technical variance in our study (red sub-distribution Figure 2A); the majority of these CpGs' r-values were below ~ 0.3. This was further corroborated by comparing the r distribution of these CpGs to a distribution of r-values, which were based on generated random data sets (see Figure 2B).

For CpGs with $r \geq 0.3$ (N = 100'564; for an example see Figure 1E) we assumed a significant underlying natural epigenetic variation. These CpGs exhibited a higher inter-individual variability in comparison to CpGs with an $r < 0.3$ (see Figure 3 and Figure 1D). Accordingly, we classified the CpGs with $r \geq 0.3$ as the most-variable CpGs. The signal of CpGs with an $r < 0.3$ most likely comprised a mixture of two different categories: random signals or signals with very low natural epigenetic variation (~2/3 of all CpGs below r = 0.3; for an example see Figure 1F) and signals with a medium to low natural epigenetic variation (~1/3 of all CpGs below r = 0.3; see light-green sub-distribution in Figure 2A).

Association analyses

Having detected CpGs that show considerable amounts of naturally occurring variation in the investigated population, we next assessed the phenotypic relevance of such variation by means of significant hits in association studies using CpG methylation as dependent variable. It has already been demonstrated that DNA methylation can be influenced by genetic variants (29; 30), as well as sex and age (31; 32). Hence, we calculated EWAS for genetic variants, sex and age in the association sample (N = 533).

meQTL analysis

To investigate the effect of SNPs on single-CpG methylation levels (meQTL), we performed genome-wide association studies on all 396'833 CpGs. By applying a stringent Bonferroni correction ($p < 2.5 \times 10^{-15}$) we identified a total of $N = 56'319$ significant CpG-SNP associations, based on $N = 25'140$ unique SNPs and $N = 31'135$ unique CpGs. As expected (29; 30), most of the top meQTLs were in *cis* of the investigated CpG (± 3.5 Mb 96.7 %). Of note, the more reliable CpGs were, the more likely they showed a significant meQTL (Figure 4A).

Given that most of the significant signal appeared in close chromosomal proximity, we restricted the subsequent analysis to a ± 3.5 Mb window surrounding the investigated CpG (*cis* analysis) and applied a less stringent *p*-value threshold ($p < 1 \times 10^{-5}$). This analysis yielded a total of $N = 228'559$ meQTLs, based on $N = 65'732$ unique SNPs and 72'310 unique CpGs. Again, the more reliable the CpG was, the more significant association signals could be revealed (see Figure 4B). In the CpGs with the highest r (average $r = 0.81$; see Fig 4B) 69.2 % of all CpGs revealed at least one significant meQTL in *cis*.

Association with sex and age

The analyses revealed significant associations with sex ($N = 6'827$ CpG) and age ($N = 405$ CpGs). Similar to the previous analyses, also here more reliable CpGs were more likely to show significant association results (see Figure 4C and Figure 4D). Of note, only autosomal CpGs were used for all analysis (see Methods).

To summarize, a well-powered replication analysis on single CpG level identified CpGs with substantial inter-individual variability, suggesting a medium to large natural epigenetic variation for the investigated population. Furthermore, the association analyses suggest that these most-variable CpGs are additionally more likely to yield significant association results.

Location of the most-variable CpGs

CpGs were classified with respect to the genomic location in the following categories: open sea, shore, shelf and island. As depicted in Figure 5, the most-variable CpGs were overrepresented in open sea and shore, and underrepresented in CpG islands. This is in accordance with the finding that the most-variable CpGs were also overrepresented in intergenic regions.

Discussion

In the present study we performed comprehensive reliability analyses of the HumanMethylation 450 K array in a large cohort of healthy young adults. The single-CpG analyses revealed reliable signals for 25% of the examined CpGs, suggesting a medium to large common epigenetic variation (> 8.6%) in the investigated population. These most-variable CpGs were more likely to reveal significant association signals. The majority of those CpGs was significantly associated with meQTLs, suggesting a considerable contribution of genetic variation to the CpGs' variability. The most-variable CpGs were additionally over-represented in open sea and shelf genomic regions, as well as in intergenic regions.

The technical replication analyses revealed a highly consistent signal on methylome-wide level. However, replication analysis on a single-CpG level can help identifying the most-variable CpGs that show common epigenetic variation in the investigated population. It is important to stress that we cannot draw final conclusions with regard to CpGs that showed no or very low variability. Variability of such CpGs may change as a function of sample size, tissue selection, samples' environmental background, genetic background, or disease status. By analyzing a large sample of healthy young adults in one tissue (i.e. blood) we most likely estimated a lower-bound variability for single-CpGs. Yet, these results may serve as a baseline reference for the naturally occurring epigenetic variation of specific CpGs. Of note, our meQTL results are in line with previous studies (33; 34) and point to the fact that genetic background significantly contributes to the most-variable CpG methylation signal.

Lack of reliability on a single-CpG level in our sample may reflect a truly invariant signal (e.g. the CpG is highly methylated in all subjects), for which the measured variability is only due to chance. However, failure to detect systematic variance can be caused also by the specifics of the microarray technology, including a fixed and limited signal resolution (2). Sequencing technologies can bypass this issue via customized signal resolution, e.g. by optimizing sequencing depth (35). Increasing sequencing depth, however, increases costs, while still being faced with the challenges of signal reliability.

Performing technical replication is a considerable cost factor. Therefore, it is crucial to discuss the relation between benefits and costs of such a replication approach (10; 36; 37). If the study design emphasizes the comparison of quite distinct states, like a case-control design in cancer research, it might prove more efficient to increase sample size rather than performing technical replications. Yet, if the aim is to perform quantitative trait analysis in a sample drawn from a homogenous population, determining the most-variable signals can considerably help optimizing the detection rate and the probability of observing true-positive findings.

Taken together, our results indicate that the Illumina Infinium Human Methylation 450K signals are reliable and valid, both on a methylome-wide and on a single-CpG level. Importantly, a relevant percentage (>25 %) of single CpGs shows a medium to strong common epigenetic variation in a homogenous sample of healthy young adults (see supplementary table 1). These findings could serve as a baseline-determination of CpGs that show natural epigenetic variation in healthy humans. A strong enrichment of meQTLs for the most-variable CpGs additionally suggests that a significant proportion of common DNA methylation may be shared across tissues. These findings could be of special relevance for studies of complex phenotypes, as in the case of neuropsychiatric disorders that often rely on proxy tissue.

Acknowledgment

This work was funded by the Swiss National Science Foundation (Sinergia grant CRSI33_130080 to D.Q. and A.P., individual grant 320030_159740 to D.Q.)

Financial Disclosures

All authors reported no biomedical financial interests or potential conflicts of interest.

References

1. Petronis A (2010): Epigenetics as a unifying principle in the aetiology of complex traits and diseases. *Nature*. 465: 721–7.

2. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, *et al.* (2011): High density DNA methylation array with single CpG site resolution. *Genomics*. 98: 288–95.

3. Tan Q, Christiansen L, von Bornemann Hjelmborg J, Christensen K (2015): Twin methodology in epigenetic studies. *J Exp Biol*. 218: 134–139.

4. Jones PA (2012): Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet*. 13: 484–92.

5. Rakyan VK, Down TA, Balding DJ, Beck S (2011): Epigenome-wide association studies for common human diseases. *Nat Rev Genet*. 12: 529–41.

6. Pidsley R, Viana J, Hannon E, Spiers HH, Troakes C, Al-Saraj S, *et al.* (2014): Methylomic profiling of human brain tissue supports a neurodevelopmental origin for schizophrenia. *Genome Biol*. 15: 483.

7. Lokk K, Modhukur V, Rajashekar B, Märtens K, Mägi R, Kolde R, *et al.* (2014): DNA methylome profiling of human tissues identifies global and tissue-specific methylation patterns. *Genome Biol*. 15: R54.

8. Horvath S, Zhang Y, Langfelder P, Kahn RS, Boks MP, van Eijk K, *et al.* (2012): Aging effects on DNA methylation modules in human brain and blood tissue. *Genome Biol*. 13: R97.

9. Davies MN, Volta M, Pidsley R, Lunnon K, Dixit A, Lovestone S, *et al.* (2012): Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. *Genome Biol*. 13: R43.

10. Altman N (2005): Replication, Variation and Normalisation in Microarray Experiments. *Appl Bioinformatics*. 4: 33–44.

11. Tylee DS, Kawaguchi DM, Glatt SJ (2013): On the outside, looking in: a review and evaluation of the comparability of blood and brain "-omes". *Am J Med Genet B Neuropsychiatr Genet*. 162B: 595–603.

12. Sandoval J, Heyn H, Moran S, Serra-Musach J, Pujana MA, Bibikova M, Esteller M (2011): Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics*. 6: 692–702.

13. Relton CL, Davey Smith G (2010): Epigenetic epidemiology of common complex disease: prospects for prediction, prevention, and treatment. *PLoS Med*. 7: e1000356.

14. Smith GD, Lawlor DA, Harbord R, Timpson N, Day I, Ebrahim S (2007): Clustered environments and randomized genes: a fundamental distinction between conventional and genetic epidemiology. *PLoS Med*. 4: e352.

15. Flanagan JM (2015): Epigenome-wide association studies (EWAS): past, present, and future. *Methods Mol Biol*. 1238: 51–63.

16. de Quervain DJ-F, Kolassa I-T, Ertl V, Onyut PL, Neuner F, Elbert T, Papassotiropoulos A

(2007): A deletion variant of the alpha2b-adrenoceptor is related to emotional memory in Europeans and Africans. *Nat Neurosci*. 10: 1137–9.

17. Heck A, Fastenrath M, Ackermann S, Auschra B, Bickel H, Coynel D, *et al.* (2014): Converging genetic and functional brain imaging evidence links neuronal excitability to working memory, psychiatric disease, and brain activity. *Neuron*. 81: 1203–13.

18. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, *et al.* (2007): PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 81: 559–75.

19. Bellenguez C, Strange A, Freeman C, Donnelly P, Spencer CCA (2012): A robust clustering algorithm for identifying problematic samples in genome-wide association studies. *Bioinformatics*. 28: 134–5.

20. Kristensen LS, Mikeska T, Krypuy M, Dobrovic A (2008): Sensitive Melting Analysis after Real Time- Methylation Specific PCR (SMART-MSP): high-throughput and probe-free quantitative DNA methylation detection. *Nucleic Acids Res*. 36: e42.

21. Assenov Y, Müller F, Lutsik P, Walter J, Lengauer T, Bock C (2014): Comprehensive analysis of DNA methylation data with RnBeads. *Nat Methods*. . doi: 10.1038/nmeth.3115.

22. Davis S, Du P, Bilke S, Triche T Jr, Bootwalla M (2014): methylumi: Handle Illumina methylation data. R package version 2.12.0. .

23. Maksimovic J, Gordon L, Oshlack A (2012): SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. *Genome Biol*. 13: R44.

24. Du P, Zhang X, Huang C-C, Jafari N, Kibbe WA, Hou L, Lin SM (2010): Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics*. 11: 587.

25. Fox J, Weisberg S (2011): *An R Companion to Applied Regression*, Second. Thousand Oaks {CA}: Sage.

26. Stacklies W, Redestig H, Scholz M, Walther D, Selbig J (2007): pcaMethods -- a Bioconductor package providing PCA methods for incomplete data. *Bioinformatics*. 23: 1164–1167.

27. Price ME, Cotton AM, Lam LL, Farré P, Emberly E, Brown CJ, *et al.* (2013): Additional annotation enhances potential for biologically-relevant analysis of the Illumina Infinium HumanMethylation450 BeadChip array. *Epigenetics Chromatin*. 6: 4.

28. Chen Y, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, *et al.* (2013): Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics*. 8: 203–9.

29. Schalkwyk LC, Meaburn EL, Smith R, Dempster EL, Jeffries AR, Davies MN, *et al.* (2010): Allelic skewing of DNA methylation is widespread across the genome. *Am J Hum Genet*. 86: 196–212.

30. Shoemaker R, Deng J, Wang W, Zhang K (2010): Allele-specific methylation is prevalent and is contributed by CpG-SNPs in the human genome. *Genome Res*. 20: 883–9.

31. Horvath S (2013): DNA methylation age of human tissues and cell types. *Genome Biol*. 14: R115.

32. Xu H, Wang F, Liu Y, Yu Y, Gelernter J, Zhang H (2014): Sex-biased methylome and transcriptome in human prefrontal cortex. *Hum Mol Genet*. 23: 1260–70.

33. McRae AF, Powell JE, Henders AK, Bowdler L, Hemani G, Shah S, *et al.* (2014): Contribution of genetic variation to transgenerational inheritance of DNA methylation. *Genome Biol*. 15: R73.

34. Fraser HB, Lam LL, Neumann SM, Kobor MS (2012): Population-specificity of human DNA methylation. *Genome Biol*. 13: R8.

35. Bock C (2012): Analysing and interpreting DNA methylation data. *Nat Rev Genet*. 13: 705–19.

36. Quackenbush J (2002): Microarray data normalization and transformation. *Nat Genet*. 32 Suppl: 496–501.

37. Ayroles JF, Gibson G (2006): Analysis of variance of microarray data. *Methods Enzymol*. 411: 214–33.

Figures

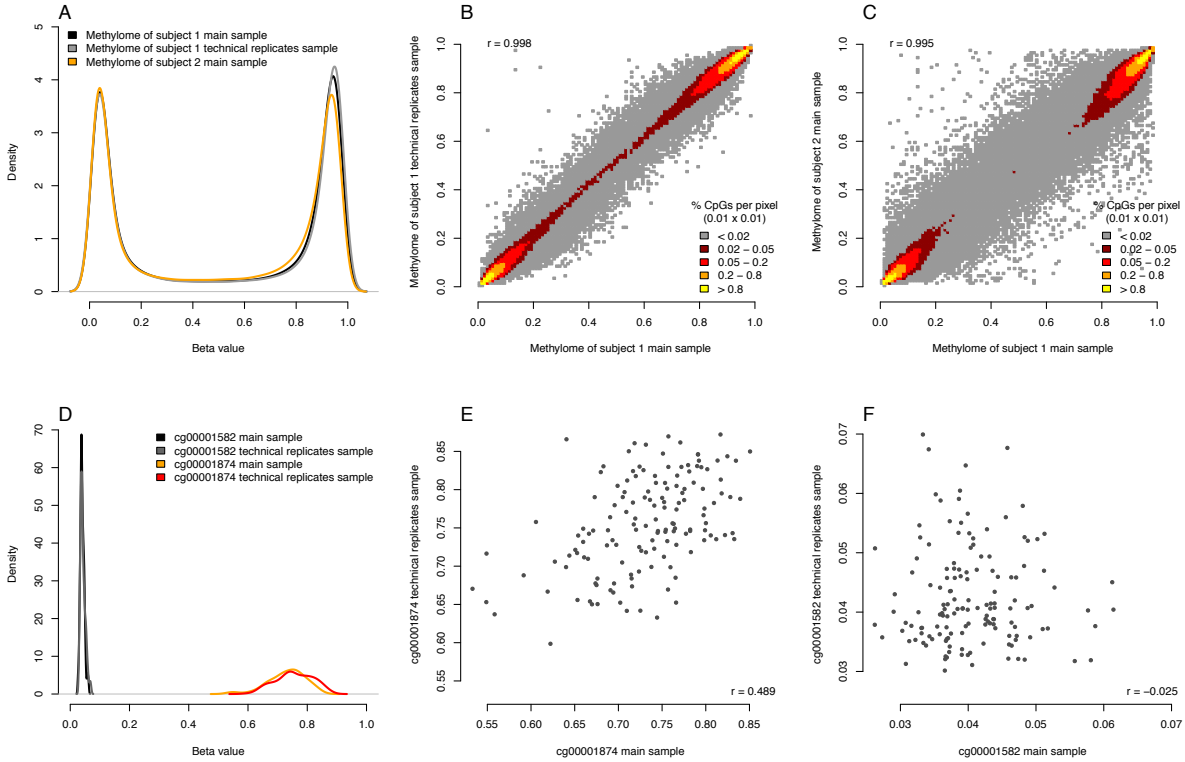


Figure 1: Methylation signal distributions and replication analyses. (A) Beta distribution on methylome level, shown for the DNA of two different subjects. For the first subject, both methylome data sets, from the main sample and from the technical replication sample, are shown. Beta values on methylome level plotted against each other from **(B)** the same subject and **(C)** two different subjects. **(D)** Beta distribution shown for two distinct CpGs, separately for the main and the technical replicates sample. Data of the main sample is restricted to the *N* = 145 subjects, which also appear in the technical replicates sample. **(E** and **F)** Beta values of two distinct CpGs plotted against each other for all *N* = 145 subjects with available data in the main sample and technical replication sample. *r*: Pearson correlation coefficient.

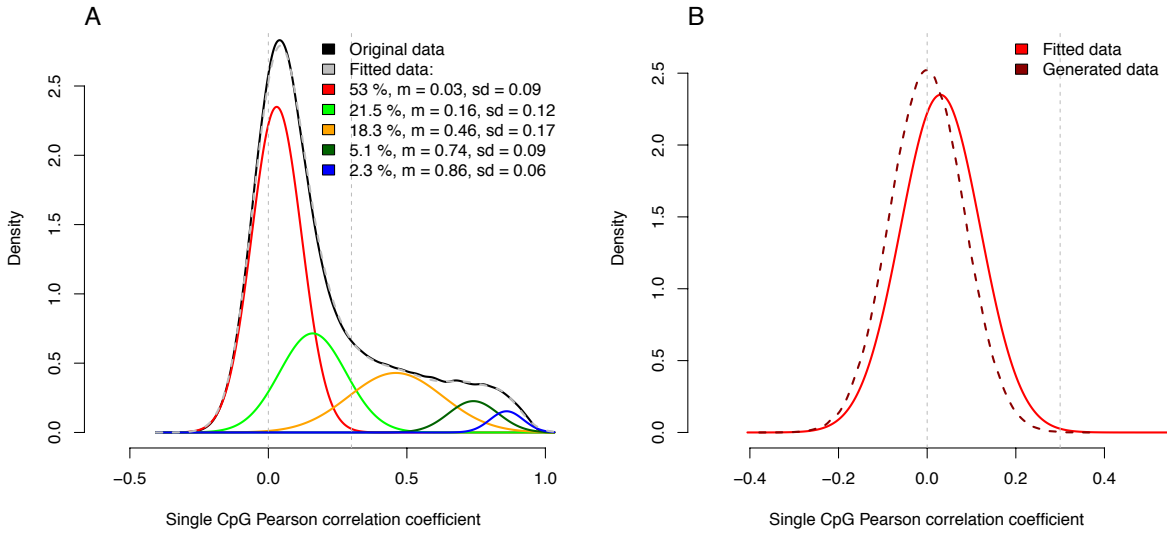


Figure 2: Replication analysis on single-CpG level. (A) The *r* distribution based on all *N* = 396'833 CpGs is depicted in black. Results of the Gaussian fit are depicted in grey and colored lines. The legend shows for each estimated sub-distribution the percentage of CpGs, the center (m) and the standard deviation (sd). **(B)** Superimposition of an *r* distribution based on random signals to the fitted Gaussian distribution with the mean closest to zero (as represented by the red curve in panel **A**). The randomly generated distribution is based on standard normal random variables (length *N* = 145) and was adjusted by the corresponding probability of random probes from the Gaussian fit (53 %). Vertical dotted gray lines depict the center of a random distribution (m = 0) and an *r* of 0.3. Above *r* = 0.3 it is unlikely that a CpG shows a signal variability based on random signals only.

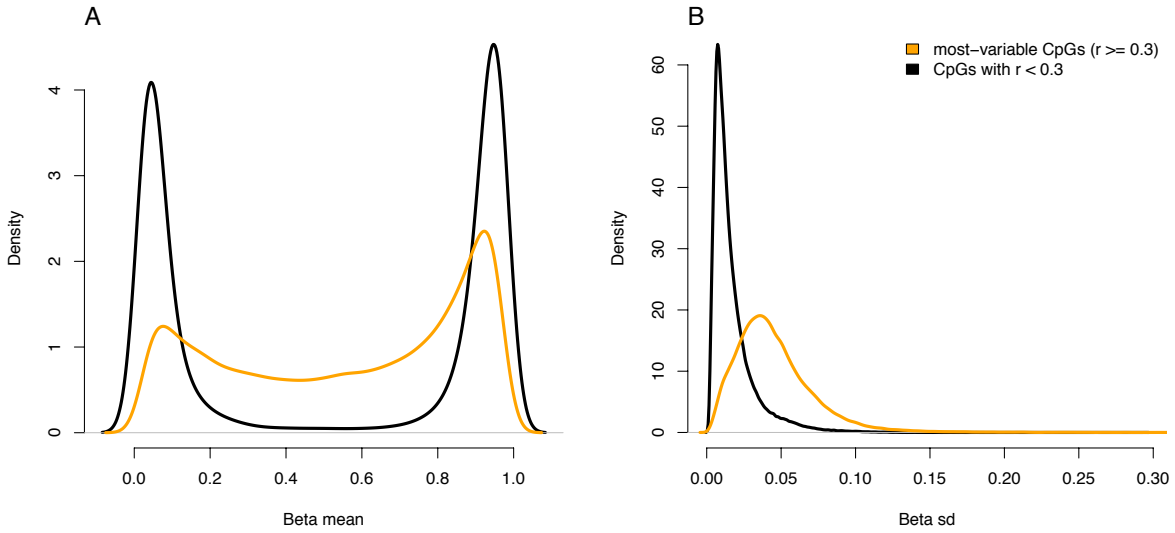


Figure 3: Density distributions of the beta's mean and standard deviation (sd), depending on the estimated variability of the CpGs. CpGs that were classified as most-variable in the replication analysis on single CpG level were less likely to show extreme beta-values **(A)** and showed a wider distribution **(B)**.

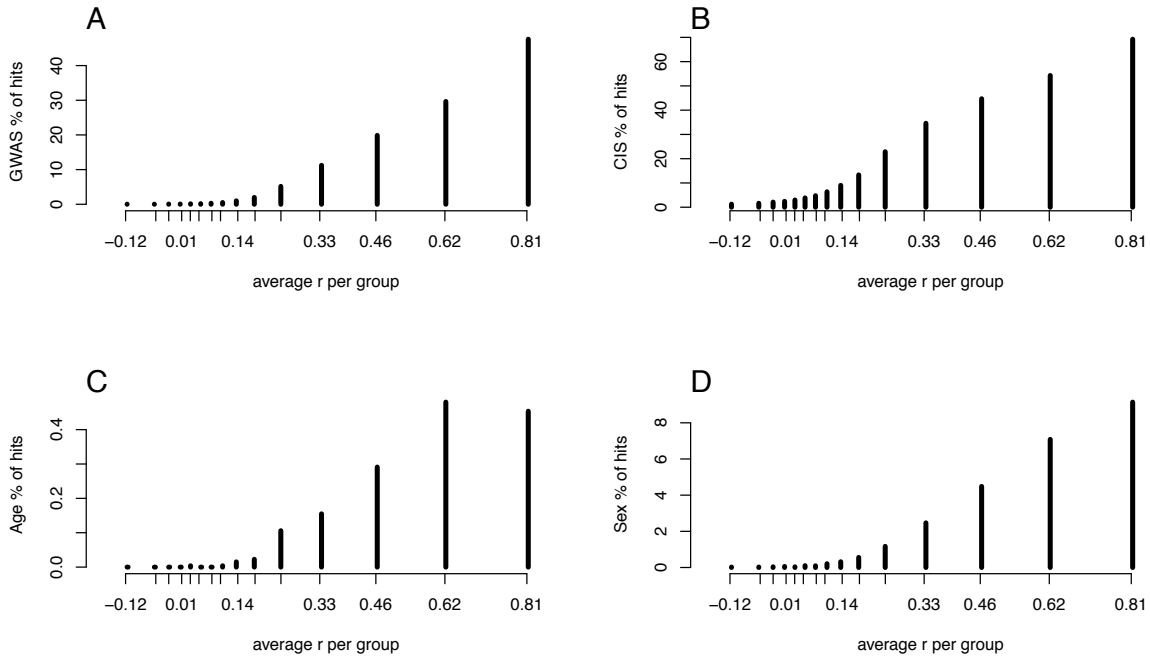


Figure 4: Association analyses results. CpGs were ordered with respect to their technical replicate *r*, and grouped into 15 equally sized bins of ~26'450 CpGs. On the x-axis, the bars are plotted at the average *r* per group. On the y-axis, the percentage of hits within the given group is depicted. **(A)** Percentage of CpGs, with at least one genome-wide Bonferroni-corrected ($p < 2.5 \times 10^{-15}$) genome-wide SNP-hit (meQTL). **(B)** Percentage of CpGs, with at least one meQTL in *cis* (± 3.5 Mb) upon a more liberal *p*-value threshold of $p < 1 \times 10^{-5}$. Percentage of CpGs that showed a significant association with age **(C)** or sex **(D).**

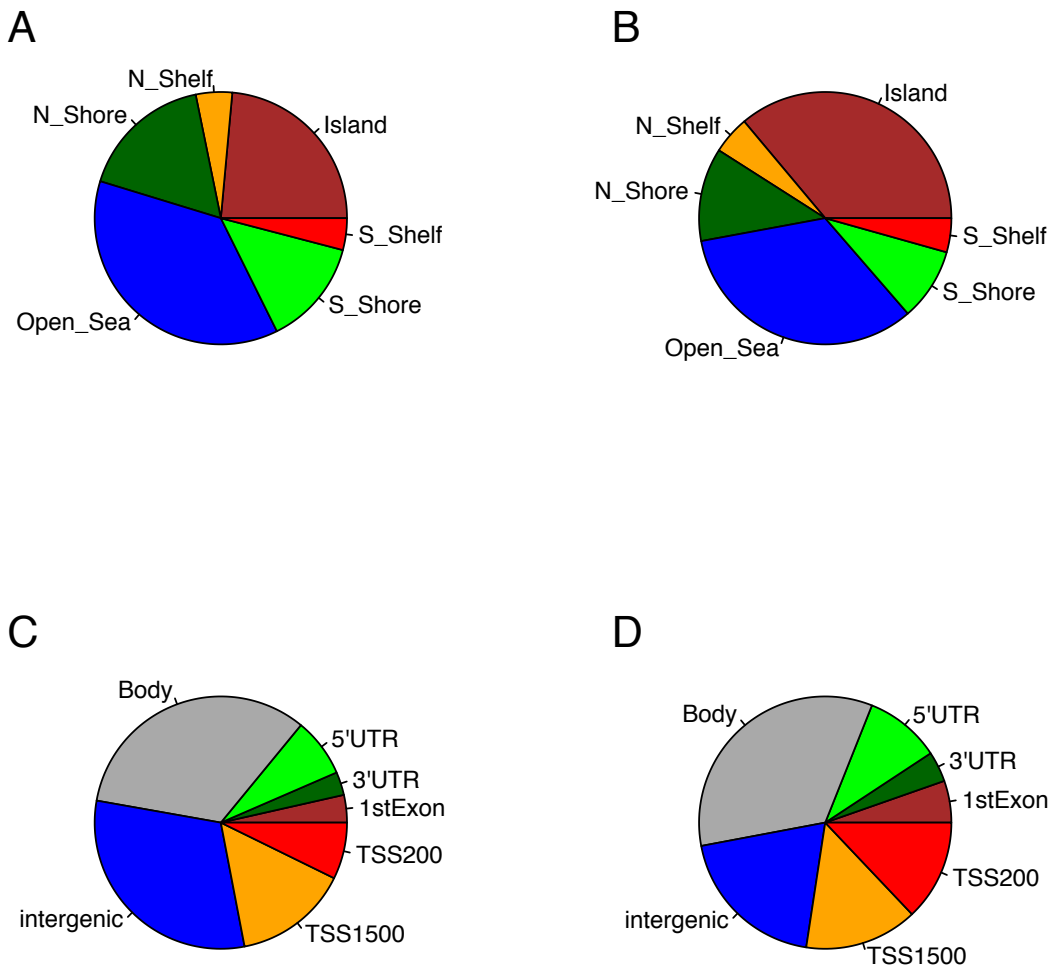


Figure 5: Genomic location of the most-variable CpGs. The most-variable CpGs ($r \geq 0.3$) **(A, C)** in comparison CpGs showing less common variation ($r < 0.3$) **(B, D)** were overrepresented in open sea and shore as well as in intergenic regions.

Supplemental Material

1 Study-design

Aim of the study is to recruit large samples of healthy young adults for assessing cognitive performance measurements, personality traits, functional and anatomical magnetic resonance imaging and genetics (based on saliva DNA) at the time-point of the main investigation. Subjects were re-invited to an additional blood and saliva sampling. The time point of this second investigation was on average 348 days (min 1 day; max 1384 days; median 314 days) after the main investigation: subject's DNA was collected between midday and evening (mean time = 2:30 p.m., range 1:00 p.m. – 8.00 p.m.); saliva and blood sampling took place within 1 hour. Subjects were free of any neurological or psychiatric illness, and did not take any medication (apart from oral contraception) at both time points of the study. Women using different methods of hormonal contraceptives (e.g., oral, spiral, patch) and naturally cycling women were included without restrictions.

2 Cross-dataset Validation

A per-subject crosscheck between phenotypic data, methylation data and genetics data was performed using the reported sex and the sex-predictions based on the arrays data, as well as matching of all SNPs represented on the Illumina 450K array to the corresponding Affymetrix SNP 6.0 or imputation derived genotype calls. This crosscheck allowed an unambiguous assignment of each methylation dataset to the corresponding genetic and phenotypic dataset.

3 Affymetrix SNP 6.0 based genotyping

SNP genotyping

Genotyping was performed using DNA obtained from saliva. Saliva samples were collected using the Oragene DNA Kit (DNA Genotek, Kanata, Canada). Saliva DNA was extracted from the Oragene DNA Kit using the standard precipitation protocol recommended by the producer. Genomic DNA concentration was adjusted to 50ng/µl in water. Digestion was done with 250ng of DNA, in parallel with 10 units of Sty I and Nsp I restriction enzymes (New England Biolabs, Knowl Piece, UK) for 2 hours at 37°C. Following, enzyme specific adaptor oligonucleotides were ligated onto the digested ends with T4 DNA Ligase for 3 hours at 16°C. After adjusting volume to 100µl with water, 10µl of the diluted ligation reactions were subjected to PCR. Three PCR reactions of 100µl were performed for Sty digested products and four PCR reactions for Nsp. PCR was performed with Titanium Taq DNA Polymerase (Clontech, Mountain View, CA) in the presence of 4.5 µM PCR primer 002 (Affymetrix, Santa Clara, CA), 350 µM each dNTP (Clontech), 1M G-C Melt (Clontech), and 1X Titanium Taq PCR Buffer (Clontech). Following cycling parameters were used: 94°C, 3 min; 30 x (94°C, 30 s; 60°C, 45 s; 68°C, 15 s); 68°C, 7 min. PCR products expected in the range size between 200-1100 bp were verified using 2% (weight/volume) TBE gel electrophoresis. PCR products were pooled and purified with the Filter Bottom Plate (Millipore, Billerica, MA; P/N MDRLN0410) using Agencourt AMPure XP Beads (Beckman Coulter, Fuillerton, CA). Quantification of purified PCR products was done on a Zenith 200rt microplate reader (Anthos-Labtec, Cambridge, UK), with average yield of 4 to 5µg/µl per sample. Subsequently, the SNP Nsp/Sty 5.0/6.0 Assay Kit (Affymetrix) was used. For the fragmentation around 250 µg of purified PCR products was digested using 0.5 units of DNAse I at 37°C for 35 minutes. Average size of fragmentation products less than 180 bps was verified using 4% (weight/volume) TBE gel electrophoresis. Following, the DNA was end labeled with 105 units of terminal deoxynucleotidyl transferase at 37°C for 4 hours. The labeled DNA was then hybridized onto Genome-Wide Human SNP 6.0 Array at 50°C for 18 hours at 60 rpm. The hybridized array was washed, stained, and scanned according to the manufacturer's

(Affymetrix) instructions using Affymetrix GeneChip Command Console (AGCC, version 3.2.0.1515). Generation of SNP calls and array quality control were performed using the command line programs of the Affymetrix Power Tools package (version: apt-1-14.4.1). According to the manufacturer's recommendation, Contrast QC was chosen as QC metric, using the default value of greater or equal than 0.4. All samples passing QC criteria were subsequently genotyped using the Birdseed (v2) algorithm.

Outlier detection and pruning

The genotypic data was projected on the two first PCA components inferred from HapMap reference populations (YRI, CEU and CHB-JPT populations). Outliers were identified using a Bayesian Clustering Algorithm (19); *N* = 35 subjects out of the main sample were identified as outlier and excluded from the association analyses. For computational efficiency of the SNP-association studies we used a reduced SNP-set: pruning of the Affymetrix-6.0 SNP-data was done with PLINK based on the *N* = 533 subjects of the association sample, with the following settings: sliding window of size 50bp shifting with steps of 5bp and a VIF threshold of 3; MAF > 0.02; HWE > 0.001; missing rate per SNP < 5%; *N* = 185,145 SNPs remaining.

4 HumanMethylation Infinium 450K BeadChip based methylation analyses

Blood samples were collected using BD Vaccutainer Push Button blood collection set and 10.0 mL BD Vacutainer Plus plastic whole blood tube, BD Hemogard™ closure with spray-coated K_2EDTA (Becton, Dickinson and Company , New Jersey, NJ). Hematological analysis, including blood cell counts, was performed at the collection time point with Sysmex pocH-100*i*™ Automated Hematology Analyzer (Sysmex Co, Kobe, JP.) DNA was isolated from the remaining fraction, upon plasma removal. The isolation was performed with QIAmp Blood Maxi Kit (Qiagen AG, Hilden, DE), using the recommended spin protocol.

In order to obtain high purity DNA prior to bisulfite conversion, isolated DNA samples were additionally re-purified. For this purpose, 2µg of DNA isolated with QIAmp/Oragene procedure, was incubated overnight at 50°C with proteinase K (Lysis buffer: 30 mM Tris·Cl; 10 mM EDTA; 1% SDS, pH 8.0; 150ng/µl Proteinase K), agitated by gentle orbital shaking. Next, the DNA was purified using Genomic DNA Clean & Concentrate Kit (Zymo Research, Irvine, CA). The quality and concentration of DNA were assessed using gel electrophoresis, NanoDrop ND- 1000 (Thermo Scientific, Waltham, MA) and fluorometry measurements (Qubit dsDNA BR Assay Kit; Invitrogen, Carlsbad, CA), respectively.

The bisulfite conversion was performed with 500 ng genomic DNA input using the EZ DNA Methylation Gold Kit (Zymo Research). A bisulfite conversion quality control on the probes was performed with DNA qPCR reaction and subsequent melting curve analysis (20). The bisulfite-converted DNA was processed and hybridized to the HumanMethylation450 BeadChip (Illumina, Inc, San Diego, CA), according to the manufacturer's instructions. The BeadChip images were scanned on the iScan system.

All DNA probes isolated from blood were randomized and processed in 8 plates, with 95 probes per plate and 1 additional control probe. The plates were processed in 3 batches, comprising 3, 4 and 1 plate, respectively. Main sample: The 568 probes were processed on 6 plates (batch 1: 2 plates; batch 2: 4 plates). Technical replication sample: The 145 probes were processed on 2 plates (batch 1: 1 plate; batch 3: 1 plate).

5 Discussion

Complex data analyses offer great opportunities to gain insight in complex human phenotypes. However, analysis and outcome are also challenging by their complexity. In the current thesis I presented three different works dealing with complex data analytical strategies. The main goal behind all three of them was not to build up comprehensive theoretical frameworks, but to perform more preparatory analytical steps. The meta-analysis validated and extended a genetic association finding of the SNP rs17070145 with human memory performance by accumulating information of about 6 years of research performed worldwide. The heritability analysis verified that a common SNP-chip array is an appropriate dataset to perform more complex genetic analyses with human working memory performance. The analysis of common epigenetic variation validated the DNA CpG methylation dataset in the context of complex analyses with mentally healthy young adults. Additionally, when comparing the outcomes of all three analyses, they also shed light on the varying complexity of putative endophenotypes in human research.

In 2006, Papassotiropoulos et al. reported an association of the SNP rs17070145 located in *KIBRA* with episodic memory performance (Papassotiropoulos et al., 2006), based on rather small sample sizes (initial sample $N = 333$; replication sample Europe $N = 424$; replication sample USA $N = 256$). Although replicated twice within the same study, this might still be a false positive finding. However, based on the meta-analytical approach that combined data of 17 independent studies, it was possible to verify and extend the initial finding: In the meta-analysis, the SNP rs17070145 was not only associated with episodic ($r = 0.068$) but also with working memory performance ($r = 0.035$). As expected, the estimated effect sizes were considerably smaller than the initially reported effect sizes ($r_{min} = 0.13$, $r_{max} = 0.22$).

The reported effect size estimates of the meta-analyses appear rather small, and might be called negligible. However, it is important to note that the strength of a large-scale cross-sectional study is to search for promising target molecules based on naturally occurring genetic variation. As a next step these promising findings should be accompanied by more targeted research that investigates potential causal mechanisms behind the identified covariance. For *KIBRA*, several promising studies based on different study designs have been performed since 2006. As one example, in

2011 Makuch et al. reported that *KIBRA* knockout mice show an impairment in learning and memory (Makuch et al., 2011). Later on Duning et al. (2013) could show, that the identified SNP is in linkage disequilibrium with two exonic SNPs (rs3822600, rs3822659) that affect KIBRAs C2 domain (Duning et al., 2013). Also based on knockout mice, Vogt-Eisele et al. (2014) additionally provided a potential mechanism how KIBRA act on memory via stabilization of the synaptic protein kinase M ζ (Vogt-Eisele et al., 2014). So from this global perspective, the initial finding successfully stimulated more targeted research on KIBRA as a quite promising target molecule for memory performances.

For the N-Back working memory task, it was further possible to identify a significant SNP-chip heritability based on N = 2'298 unrelated healthy individuals. As one would expect, there was no prominent contribution of e.g. one single chromosome only, but an equal distribution of heritability estimates over chromosomes, as expected for highly polygenetic traits. Heritability estimates for 2-Back derived working memory performance ranged between 31 % (s.e. 14 %) and 41 % (s.e. 14 %). This is slightly above the SNP-chip heritability estimates derived for related neuropsychiatric diseases (Lee et al., 2013). Recently, based on a bivariate analysis a shared genetic variability between schizophrenia and performance IQ of about 14 % could be revealed in a large (N = 11'853) UK schizophrenia sample (Hubbard et al., 2015). For future research, a comparable analysis of genetic covariance between schizophrenia and working memory performance would be of interest.

The results from the estimation of common variation for the DNA CpG methylation data point to the fact that a carful investigation of the primary data might help improving the signal-to-noise ratio in complex analytical situations. Only about 1/3 of the CpGs showed a considerable amount of common variation in the investigated population of healthy young adults. Importantly, the amount of variation was informative for the association studies that were performed for further validation. These results may serve as baseline information for further studies. If a CpG already shows a considerable part of common variation in a population of healthy young adults, and is additionally associated with disease status, this might be a hint that the variation in the CpG is a very general predisposition to the disease,

whereas when the variation appears in the time-course of the disease, this might rather be a causal epimutation.

When comparing the outcomes of all three works it becomes obvious that the phenotypes memory performance and DNA CpG methylation show a considerable difference in their polygenicity, although both are classified as putative intermediate phenotypes or endophenotypes. From a theoretical perspective these two have at least one considerable difference in nature: DNA methylation can be seen as a filter that dynamically modifies the output of the genetic blueprint. It is also known that environmental factors may change DNA methylation status. Nevertheless, DNA methylation itself has no direct impact on the environment but might modify other intermediate phenotypes like brain functioning or functioning of the immune system. In contrast to that, differences in memory performance rather reflect the sum of the human experiences in combination with brain and body functioning. Besides this theoretical view, the difference in the polygenicity of these two phenotypes can nicely be seen in the genetic analyses performed: For working-memory performance the heritability estimates per chromosome show a polygenetic picture with a large amount of common variants being involved. This picture is also confirmed by the meta-analysis, in which the estimated effect sizes for one common genetic variant are rather small. In contrast to that, for the CpG-methylation data we could show in the genetic association analysis that a considerable amount of CpGs has highly significant associations with SNPs in close proximity to the CpG only. Adjacent analyses (not included in the final paper) could show that in most cases there is only one main SNP-signal that is associated with the CpG methylation signal, with a medium to large effect size.

To summarize, complex phenotypes call for complex data analyses strategies. However, with these complex analytical approaches one has to find a good balance between the need for complexity and the benefit of simplicity. Simplifying the research design is a suitable approach to circumvent overwhelmingly complex analytical situations. Furthermore a thoughtful investigation of the primary signals is a key step before building up more comprehensive theoretical models. The analyses reported in this cumulative thesis aimed at performing preparatory analytical steps before building up comprehensive theoretical and analytical models. Here, it is important to note that neither large amounts of data nor complex data analyses can

replace the stimulating function of defining a suitable theoretical model. Bernard Forscher (1963) summarizes this in his analogy of a brickyard, where the brickyard represents data: What should be avoided is that “no effort was made even to maintain the distinction between a pile of bricks and a true edifice” (Forscher, 1963). Hence, to achieve knowledge from information in epidemiological oriented research, big data and complex analytical strategies should finally be seen as tools to build up suitable models of reality, which will be the future directive of my work.

6 References

Akesson, M., Bensch, S., Hasselquist, D., Tarka, M., & Hansson, B. (2008). Estimating heritabilities and genetic correlations: comparing the “animal model” with parent-offspring regression using data from a natural population. *PloS One*, *3*(3), e1739. doi:10.1371/journal.pone.0001739

Allison, D. B., Cui, X., Page, G. P., & Sabripour, M. (2006). Microarray data analysis: from disarray to consolidation and consensus. *Nature Reviews. Genetics*, *7*(1), 55–65. doi:10.1038/nrg1749

Altman, N. (2005). Replication, Variation and Normalisation in Microarray Experiments. *Applied Bioinformatics*, *4*(1), 33–44. doi:10.2165/00822942-200504010-00004

Ando, J., Ono, Y., & Wright, M. J. (2001). Genetic structure of spatial and verbal working memory. *Behav Genet*, *31*(6), 615–624.

Barch, D. M., & Smith, E. (2008). The cognitive neuroscience of working memory: relevance to CNTRICS and schizophrenia. *Biological Psychiatry*, *64*(1), 11–7. doi:10.1016/j.biopsych.2008.03.003

Besen, J., & Gan, S. D. (2014). A critical evaluation of clinical research study designs. *The Journal of Investigative Dermatology*, *134*(3), e18. doi:10.1038/jid.2013.545

Bibikova, M., Barnes, B., Tsan, C., Ho, V., Klotzle, B., Le, J. M., … Shen, R. (2011). High density DNA methylation array with single CpG site resolution. *Genomics*, *98*(4), 288–95. doi:10.1016/j.ygeno.2011.07.007

Bird, A. (2007). Perceptions of epigenetics. *Nature*, *447*(7143), 396–8. doi:10.1038/nature05913

Bird, C. M., & Burgess, N. (2008). The hippocampus and memory: insights from spatial processing. *Nat Rev Neurosci*, *9*(3), 182–194. doi:10.1038/nrn2335

Blalock, H. M. J. (1966). The Identification Problem and Theory Building: The Case of Status Inconsistency. *American Sociological Review*, *31*(1), 52–61.

Bock, C. (2009). Epigenetic biomarker development. *Epigenomics*, *1*(1), 99–110. doi:10.2217/epi.09.6

Bock, C. (2012). Analysing and interpreting DNA methylation data. *Nature Reviews. Genetics*, *13*(10), 705–19. doi:10.1038/nrg3273

Bockmühl, Y., Patchev, A. V, Madejska, A., Hoffmann, A., Sousa, J. C., Sousa, N., … Spengler, D. (2015). Methylation at the CpG island shore region upregulates Nr3c1 promoter activity after early-life stress. *Epigenetics*, *10*(3), 247–57. doi:10.1080/15592294.2015.1017199

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews. Neuroscience*, *14*(5), 365–76. doi:10.1038/nrn3475

Cahill, L., McGaugh, J. L., & Weinberger, N. M. (2001). The neurobiology of learning and memory: some reminders to remember. *Trends in Neurosciences*, *24*(10), 578–81.

Calkins, M. E., Tepper, P., Gur, R. C., Ragland, J. D., Klei, L., Wiener, H. W., … Gur, R. E. (2010). Project among African-Americans to explore risks for schizophrenia (PAARTNERS): evidence for impairment and heritability of neurocognitive functioning in families of schizophrenia patients. *Am J Psychiatry*, *167*(4), 459–472. doi:10.1176/appi.ajp.2009.08091351

Conklin, H. M., Curtis, C. E., Calkins, M. E., & Iacono, W. G. (2005). Working memory functioning in schizophrenia patients and their first-degree relatives: cognitive functioning shedding light on etiology. *Neuropsychologia*, *43*(6), 930–42. doi:10.1016/j.neuropsychologia.2004.09.013

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Control Clin Trials*, *7*(3), 177–188.

Drevets, W. C., Price, J. L., & Furey, M. L. (2008). Brain structural and functional abnormalities in mood disorders: implications for neurocircuitry models of depression. *Brain Struct Funct*, *213*(1-2), 93–118. doi:10.1007/s00429-008-0189-x

Duning, K., Wennmann, D. O., Bokemeyer, A., Reissner, C., Wersching, H., Thomas,

C., … Kremerskothen, J. (2013). Common exonic missense variants in the C2 domain of the human KIBRA protein modify lipid binding and cognitive performance. *Translational Psychiatry*, *3*, e272. doi:10.1038/tp.2013.49

Egan, M. F., Goldberg, T. E., Gscheidle, T., Weirich, M., Rawlings, R., Hyde, T. M., … Weinberger, D. R. (2001). Relative risk for cognitive impairments in siblings of patients with schizophrenia. *Biological Psychiatry*, *50*(2), 98–107.

Eichenbaum, H. (2013). What H.M. taught us. *Journal of Cognitive Neuroscience*, *25*(1), 14–21. doi:10.1162/jocn_a_00285

Elvevåg, B., & Goldberg, T. E. (2000). Cognitive impairment in schizophrenia is the core of the disorder. *Critical Reviews in Neurobiology*, *14*(1), 1–21.

Eustache, F., & Desgranges, B. (2008). MNESIS: towards the integration of current multisystem models of memory. *Neuropsychology Review*, *18*(1), 53–69. doi:10.1007/s11065-008-9052-3

Ferbinteanu, J., Kennedy, P. J., & Shapiro, M. L. (2006). Episodic memory--from brain to mind. *Hippocampus*, *16*(9), 691–703. doi:10.1002/hipo.20204

Finkel, D., & McGue, M. (1998). Age differences in the nature and origin of individual differences in memory: a behavior genetic analysis. *Int J Aging Hum Dev*, *47*(3), 217–39.

Finkel, D., Pedersen, N., & McGue, M. (1995). Genetic influences on memory performance in adulthood: comparison of Minnesota and Swedish twin data. *Psychol Aging*, *10*(3), 437–446.

Finucane, H. K., Bulik-Sullivan, B., Gusev, A., Trynka, G., Reshef, Y., Loh, P.-R., … Price, A. L. (2015). Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nature Genetics*. doi:10.1038/ng.3404

Fisher, R. A. (1918). The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Transactions of the Royal Society of Edinburgh*, *52*, 399–433.

Forbes, N. F., Carrick, L. A., McIntosh, A. M., & Lawrie, S. M. (2009). Working memory in schizophrenia: a meta-analysis. *Psychol Med*, *39*(06), 889–905.

Forscher, B. K. (1963). Chaos in the Brickyard. *Science (New York, N.Y.)*, *142*(3590),

339. doi:10.1126/science.142.3590.339

Gelernter, J. (2015). Genetics of complex traits in psychiatry. *Biological Psychiatry*, *77*(1), 36–42. doi:10.1016/j.biopsych.2014.08.005

Gottesman, I. I. (2003). The Endophenotype Concept in Psychiatry: Etymology and Strategic Intentions. *American Journal of Psychiatry*, *160*(4), 636–645. doi:10.1176/appi.ajp.160.4.636

Gottesman, I. I., & Shields, J. (1973). Genetic theorizing and schizophrenia. *The British Journal of Psychiatry : The Journal of Mental Science*, *122*(566), 15–30.

Greenwood, T. A., Beeri, M. S., Schmeidler, J., Valerio, D., Raventós, H., Mora-Villalobos, L., … Silverman, J. M. (2011). Heritability of Cognitive Functions in Families of Successful Cognitive Aging Probands from the Central Valley of Costa Rica. *J Alzheimers Dis*, *27*(4), 897–907. doi:10.3233/JAD-2011-110782

Hannon, E., Lunnon, K., Schalkwyk, L., & Mill, J. (2015). Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes. *Epigenetics*. doi:10.1080/15592294.2015.1100786

Hannum, G., Guinney, J., Zhao, L., Zhang, L., Hughes, G., Sadda, S., … Zhang, K. (2013). Genome-wide methylation profiles reveal quantitative views of human aging rates. *Molecular Cell*, *49*(2), 359–67. doi:10.1016/j.molcel.2012.10.016

Heck, A., Fastenrath, M., Ackermann, S., Auschra, B., Bickel, H., Coynel, D., … Papassotiropoulos, A. (2014). Converging genetic and functional brain imaging evidence links neuronal excitability to working memory, psychiatric disease, and brain activity. *Neuron*, *81*(5), 1203–13. doi:10.1016/j.neuron.2014.01.010

Heck, A., Fastenrath, M., Coynel, D., Auschra, B., Bickel, H., Freytag, V., … Papassotiropoulos, A. (2015). Genetic Analysis of Association Between Calcium Signaling and Hippocampal Activation, Memory Performance in the Young and Old, and Risk for Sporadic Alzheimer Disease. *JAMA Psychiatry*, *72*(10), 1029–36. doi:10.1001/jamapsychiatry.2015.1309

Hedges, L. V, & Olkin, I. (1985). Statistical methods for meta-analysis. *Stat Med*, *20*(6), 825–840. doi:10.1002/sim.650

Henke, K. (2010). A model for memory systems based on processing modes rather

than consciousness. *Nature Reviews. Neuroscience*, *11*(7), 523–32. doi:10.1038/nrn2850

Horvath, S. (2013). DNA methylation age of human tissues and cell types. *Genome Biology*, *14*(10), R115. doi:10.1186/gb-2013-14-10-r115

Horvath, S., Zhang, Y., Langfelder, P., Kahn, R. S., Boks, M. P., van Eijk, K., … Ophoff, R. A. (2012). Aging effects on DNA methylation modules in human brain and blood tissue. *Genome Biology*, *13*(10), R97. doi:10.1186/gb-2012-13-10-r97

Hubbard, L., Tansey, K. E., Rai, D., Jones, P., Ripke, S., Chambert, K. D., … Zammit, S. (2015). Evidence of Common Genetic Overlap Between Schizophrenia and Cognition. *Schizophrenia Bulletin*, sbv168. doi:10.1093/schbul/sbv168

IntHout, J., Ioannidis, J. P. A., Borm, G. F., & Goeman, J. J. (2015). Small studies are more heterogeneous than large ones: a meta-meta-analysis. *Journal of Clinical Epidemiology*, *68*(8), 860–9. doi:10.1016/j.jclinepi.2015.03.017

Ioannidis, J. P. A. (2008). Calibration of credibility of agnostic genome-wide associations. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics : The Official Publication of the International Society of Psychiatric Genetics*, *147B*(6), 964–72. doi:10.1002/ajmg.b.30721

Ioannidis, J. P. A. (2014). How to Make More Published Research True. *PLoS Medicine*, *11*(10), e1001747. doi:10.1371/journal.pmed.1001747

Ioannidis, J. P. A., Patsopoulos, N. A., & Evangelou, E. (2007). Heterogeneity in meta-analyses of genome-wide association investigations. *PloS One*, *2*(9), e841. doi:10.1371/journal.pone.0000841

Jawahar, M. C., Murgatroyd, C., Harrison, E. L., & Baune, B. T. (2015). Epigenetic alterations following early postnatal stress: a review on novel aetiological mechanisms of common psychiatric disorders. *Clinical Epigenetics*, *7*, 122. doi:10.1186/s13148-015-0156-3

Johansson, B., Whitfield, K., Pedersen, N. L., Hofer, S. M., Ahern, F., & McClearn, G. E. (1999). Origins of individual differences in episodic memory in the oldest-old: a population-based study of identical and same-sex fraternal twins aged 80 and older. *J Gerontol B Psychol Sci Soc Sci*, *54*(3), 173–179.

Kramer, M. S. (1998). A comparison of laboratory and epidemiologic approaches to

pediatric research. *Pediatric Research*, *43*(6), 711–8. doi:10.1203/00006450-199806000-00001

Kreft, I. G. G., Kreft, I., & de Leeuw, J. (1998). *Introducing Multilevel Modeling*. SAGE Publications.

Kruuk, L. E. B. (2004). Estimating genetic parameters in natural populations using the "animal model". *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *359*(1446), 873–90. doi:10.1098/rstb.2003.1437

Ladd-Acosta, C., Shu, C., Lee, B. K., Gidaya, N., Singer, A., Schieve, L. A., … Daniele Fallin, M. (2015). Presence of an epigenetic signature of prenatal cigarette smoke exposure in childhood. *Environmental Research*, *144*(Pt A), 139–148. doi:10.1016/j.envres.2015.11.014

Lango Allen, H., Estrada, K., Lettre, G., Berndt, S. I., Weedon, M. N., Rivadeneira, F., … Hirschhorn, J. N. (2010). Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature*, *467*(7317), 832–838. doi:10.1038/nature09410

Lardenoije, R., Iatrou, A., Kenis, G., Kompotis, K., Steinbusch, H. W. M., Mastroeni, D., … Rutten, B. P. F. (2015). The Epigenetics of Aging and Neurodegeneration. *Progress in Neurobiology*, *131*, 21–64. doi:10.1016/j.pneurobio.2015.05.002

Lee, S. H., Ripke, S., Neale, B. M., Faraone, S. V, Purcell, S. M., Perlis, R. H., … Wray, N. R. (2013). Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nature Genetics*, *45*(9), 984–94. doi:10.1038/ng.2711

Lee, T., Mosing, M. A., Henry, J. D., Trollor, J. N., Ames, D., Martin, N. G., … Sachdev, P. S. (2012). Genetic influences on four measures of executive functions and their covariation with general cognitive ability: the Older Australian Twins Study. *Behavior Genetics*, *42*(4), 528–38. doi:10.1007/s10519-012-9526-1

Leek, J. T., & Peng, R. D. (2015). What is the question? *Science*, *347*(6228), 1314–1315. doi:10.1126/science.aaa6146

Lister, R., Mukamel, E. A., Nery, J. R., Urich, M., Puddifoot, C. A., Johnson, N. D., … Ecker, J. R. (2013). Global Epigenomic Reconfiguration During Mammalian Brain Development. *Science*, science.1237905–. doi:10.1126/science.1237905

Lückmann, H. C., Jacobs, H. I. L., & Sack, A. T. (2014). The cross-functional role of

frontoparietal regions in cognition: internal attention as the overarching mechanism. *Progress in Neurobiology*, *116*, 66–86. doi:10.1016/j.pneurobio.2014.02.002

Lunnon, K., Smith, R., Hannon, E., De Jager, P. L., Srivastava, G., Volta, M., … Mill, J. (2014). Methylomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease. *Nature Neuroscience*, *17*(9), 1164–70. doi:10.1038/nn.3782

Lyon, A. (2013). Why are Normal Distributions Normal? *The British Journal for the Philosophy of Science*, *65*(3), 621–649. doi:10.1093/bjps/axs046

Makuch, L., Volk, L., Anggono, V., Johnson, R. C., Yu, Y., Duning, K., … Huganir, R. L. (2011). Regulation of AMPA Receptor Function by the Human Memory-Associated Gene KIBRA. *Neuron*, *71*(6), 1022–1029. doi:10.1016/j.neuron.2011.08.017

Marazziti, D., Consoli, G., Picchetti, M., Carlini, M., & Faravelli, L. (2010). Cognitive impairment in major depression. *Eur J Pharmacol*, *626*(1), 83–86. doi:10.1016/j.ejphar.2009.08.046

Mark, W., & Toulopoulou, T. (2015). Cognitive intermediate phenotype and genetic risk for psychosis. *Current Opinion in Neurobiology*, *36*, 23–30. doi:10.1016/j.conb.2015.08.008

Mayr, E. (1997). *Evolution and the Diversity of Life: Selected Essays*. Harvard Univ Pr; Auflage: Revised. (25. März 1997).

Mesholam-Gately, R. I., Giuliano, A. J., Goff, K. P., Faraone, S. V., & Seidman, L. J. (2009). Neurocognition in first-episode schizophrenia: A meta-analytic review. *Neuropsychology*, *23*, 315–336.

Miller, E. K. (2013). The "working" of working memory. *Dialogues in Clinical Neuroscience*, *15*(4), 411–8.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine*, *6*(7), e1000097. doi:10.1371/journal.pmed.1000097

Mousseau, T. A., & Roff, D. A. (1987). Natural selection and the heritability of fitness components. *Heredity*, *59 (Pt 2)*, 181–97.

Nakagawa, S., & Cuthill, I. C. (2007). Effect size, confidence interval and statistical

significance: a practical guide for biologists. *Biol Rev Camb Philos Soc*, *82*(4), 591–605. doi:10.1111/j.1469-185X.2007.00027.x

Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping*, *25*(1), 46–59. doi:10.1002/hbm.20131

Palma-Gudiel, H., Córdova-Palomera, A., Eixarch, E., Deuschle, M., & Fañanás, L. (2015). Maternal psychosocial stress during pregnancy alters the epigenetic signature of the glucocorticoid receptor gene promoter in their offspring: a meta-analysis. *Epigenetics*, *10*(10), 893–902. doi:10.1080/15592294.2015.1088630

Papassotiropoulos, A., & de Quervain, D. J.-F. (2011). Genetics of human episodic memory: dealing with complexity. *Trends in Cognitive Sciences*, *15*(9), 381–7. doi:10.1016/j.tics.2011.07.005

Papassotiropoulos, A., & de Quervain, D. J.-F. (2015). Failed drug discovery in psychiatry: time for human genome-guided solutions. *Trends in Cognitive Sciences*, *19*(4), 183–7. doi:10.1016/j.tics.2015.02.002

Papassotiropoulos, A., Gerhards, C., Heck, A., Ackermann, S., Aerni, A., Schicktanz, N., … de Quervain, D. J.-F. (2013). Human genome-guided identification of memory-modulating drugs. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(46), E4369–74. doi:10.1073/pnas.1314478110

Papassotiropoulos, A., Stephan, D. A., Huentelman, M. J., Hoerndli, F. J., Craig, D. W., Pearson, J. V, … de Quervain, D. J.-F. (2006). Common Kibra alleles are associated with human memory performance. *Science*, *314*(5798), 475–478. doi:10.1126/science.1129837

Persson, J., Pudas, S., Lind, J., Kauppi, K., Nilsson, L.-G., & Nyberg, L. (2012). Longitudinal structure-function correlates in elderly reveal MTL dysfunction with cognitive decline. *Cerebral Cortex (New York, N.Y. : 1991)*, *22*(10), 2297–304. doi:10.1093/cercor/bhr306

Petronis, A. (2010). Epigenetics as a unifying principle in the aetiology of complex traits and diseases. *Nature*, *465*(7299), 721–7. doi:10.1038/nature09230

Pfeiffer, T., Bertram, L., & Ioannidis, J. P. A. (2011). Quantifying selective reporting and the Proteus phenomenon for multiple datasets with similar bias. *PloS One*,

6(3), e18362. doi:10.1371/journal.pone.0018362

Pidsley, R., Viana, J., Hannon, E., Spiers, H. H., Troakes, C., Al-Saraj, S., … Mill, J. (2014). Methylomic profiling of human brain tissue supports a neurodevelopmental origin for schizophrenia. *Genome Biology*, *15*(10), 483. doi:10.1186/PREACCEPT-1621721621132088

Polderman, T. J. C., Benyamin, B., de Leeuw, C. A., Sullivan, P. F., van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics*, *47*(7), 702–709. doi:10.1038/ng.3285

Potter, J. D. (2001). At the interfaces of epidemiology, genetics and genomics. *Nature Reviews. Genetics*, *2*(2), 142–7. doi:10.1038/35052575

Preston, G. A., & Weinberger, D. R. (2005). Intermediate phenotypes in schizophrenia: a selective review. *Dialogues in Clinical Neuroscience*, *7*(2), 165–79.

Rasch, B., Papassotiropoulos, A., & de Quervain, D.-F. (2010). Imaging genetics of cognitive functions: Focus on episodic memory. *NeuroImage*, *53*(3), 870–7. doi:10.1016/j.neuroimage.2010.01.001

Read, S., Pedersen, N. L., Gatz, M., Berg, S., Vuoksimaa, E., Malmberg, B., … McClearn, G. E. (2006). Sex differences after all those years? Heritability of cognitive abilities in old age. *J Gerontol B Psychol Sci Soc Sci*, *61*(3), 137–143.

Relton, C. L., & Davey Smith, G. (2010). Epigenetic epidemiology of common complex disease: prospects for prediction, prevention, and treatment. *PLoS Medicine*, *7*(10), e1000356. doi:10.1371/journal.pmed.1000356

Ripke, S., Neale, B. M., Corvin, A., Walters, J. T. R., Farh, K.-H., Holmans, P. A., … O'Donovan, M. C. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature*, *511*(7510), 421–7. doi:10.1038/nature13595

Samuelson, K. W. (2011). Post-traumatic stress disorder and declarative memory functioning: a review. *Dialogues Clin Neurosci*, *13*(3), 346–351.

Sandoval, J., Heyn, H., Moran, S., Serra-Musach, J., Pujana, M. A., Bibikova, M., & Esteller, M. (2011). Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics : Official Journal of the DNA*

Methylation Society, *6*(6), 692–702.

Schacter, D. L. (2013). Memory: from the laboratory to everyday life. *Dialogues in Clinical Neuroscience*, *15*(4), 393–5.

Schadt, E. E., Lamb, J., Yang, X., Zhu, J., Edwards, S., Guhathakurta, D., … Lusis, A. J. (2005). An integrative genomics approach to infer causal associations between gene expression and disease. *Nature Genetics*, *37*(7), 710–7. doi:10.1038/ng1589

Schalkwyk, L. C., Meaburn, E. L., Smith, R., Dempster, E. L., Jeffries, A. R., Davies, M. N., … Mill, J. (2010). Allelic skewing of DNA methylation is widespread across the genome. *American Journal of Human Genetics*, *86*(2), 196–212. doi:10.1016/j.ajhg.2010.01.014

Shoemaker, R., Deng, J., Wang, W., & Zhang, K. (2010). Allele-specific methylation is prevalent and is contributed by CpG-SNPs in the human genome. *Genome Research*, *20*(7), 883–9. doi:10.1101/gr.104695.109

Smith, G. D., Lawlor, D. A., Harbord, R., Timpson, N., Day, I., & Ebrahim, S. (2007). Clustered environments and randomized genes: a fundamental distinction between conventional and genetic epidemiology. *PLoS Medicine*, *4*(12), e352. doi:10.1371/journal.pmed.0040352

Sollars, V., Lu, X., Xiao, L., Wang, X., Garfinkel, M. D., & Ruden, D. M. (2003). Evidence for an epigenetic mechanism by which Hsp90 acts as a capacitor for morphological evolution. *Nature Genetics*, *33*(1), 70–4. doi:10.1038/ng1067

Sperling, R. A., Dickerson, B. C., Pihlajamaki, M., Vannini, P., LaViolette, P. S., Vitolo, O. V, … Johnson, K. A. (2010). Functional alterations in memory networks in early Alzheimer's disease. *Neuromolecular Medicine*, *12*(1), 27–43. doi:10.1007/s12017-009-8109-7

Spiers, H., Hannon, E., Schalkwyk, L. C., Smith, R., Wong, C. C. Y., O'Donovan, M. C., … Mill, J. (2015). Methylomic trajectories across human fetal brain development. *Genome Research*, *25*(3), 338–52. doi:10.1101/gr.180273.114

Tan, Q., Christiansen, L., von Bornemann Hjelmborg, J., & Christensen, K. (2015). Twin methodology in epigenetic studies. *The Journal of Experimental Biology*, *218*(Pt 1), 134–139. doi:10.1242/jeb.107151

Thompson, J. L., Watson, J. R., Steinhauer, S. R., Goldstein, G., & Pogue-Geile, M. F.

(2005). Indicators of genetic liability to schizophrenia: a sibling study of neuropsychological performance. *Schizophrenia Bulletin*, *31*(1), 85–96. doi:10.1093/schbul/sbi009

Toulopoulou, T., van Haren, N., Zhang, X., Sham, P. C., Cherny, S. S., Campbell, D. D., … Kahn, R. S. (2015). Reciprocal causation models of cognitive vs volumetric cerebral intermediate phenotypes for schizophrenia in a pan-European twin cohort. *Molecular Psychiatry*, *20*(11), 1386–96. doi:10.1038/mp.2014.152

Tulving, E. (2002). Episodic memory: from mind to brain. *Annual Review of Psychology*, *53*, 1–25. doi:10.1146/annurev.psych.53.100901.135114

Tylee, D. S., Kawaguchi, D. M., & Glatt, S. J. (2013). On the outside, looking in: a review and evaluation of the comparability of blood and brain "-omes". *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics : The Official Publication of the International Society of Psychiatric Genetics*, *162B*(7), 595–603. doi:10.1002/ajmg.b.32150

van der Sijde, M. R., Ng, A., & Fu, J. (2014). Systems genetics: From GWAS to disease pathways. *Biochimica et Biophysica Acta*, *1842*(10), 1903–1909. doi:10.1016/j.bbadis.2014.04.025

Van Snellenberg, J. X. (2009). Working memory and long-term memory deficits in schizophrenia: is there a common substrate? *Psychiatry Res*, *174*(2), 89–96. doi:10.1016/j.pscychresns.2009.04.001

Vinkhuyzen, A. A. E., van der Sluis, S., Boomsma, D. I., de Geus, E. J. C., & Posthuma, D. (2010). Individual differences in processing speed and working memory speed as assessed with the Sternberg memory scanning task. *Behavior Genetics*, *40*(3), 315–26. doi:10.1007/s10519-009-9315-7

Visscher, P. M. (2008). Sizing up human height variation. *Nat Genet*, *40*(5), 489–490. doi:10.1038/ng0508-489

Visscher, P. M., Hill, W. G., & Wray, N. R. (2008). Heritability in the genomics era--concepts and misconceptions. *Nature Reviews. Genetics*, *9*(4), 255–66. doi:10.1038/nrg2322

Vogt-Eisele, A., Krüger, C., Duning, K., Weber, D., Spoelgen, R., Pitzer, C., … Schneider, A. (2014). KIBRA (KIdney/BRAin protein) regulates learning and

memory and stabilizes Protein kinase Mζ. *Journal of Neurochemistry*, *128*(5), 686–700. doi:10.1111/jnc.12480

Volk, H. E., McDermott, K. B., Roediger, H. L., & Todd, R. D. (2006). Genetic influences on free and cued recall in long-term memory tasks. *Twin Res Hum Genet*, *9*(5), 623–631. doi:10.1375/183242706778553462

Vukojevic, V., Kolassa, I.-T., Fastenrath, M., Gschwind, L., Spalek, K., Milnik, A., … de Quervain, D. J.-F. (2014). Epigenetic modification of the glucocorticoid receptor gene is linked to traumatic memory and post-traumatic stress disorder risk in genocide survivors. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *34*(31), 10274–84. doi:10.1523/JNEUROSCI.1526-14.2014

Wilson, R. S., Barral, S., Lee, J. H., Leurgans, S. E., Foroud, T. M., Sweet, R. A., … Bennett, D. A. (2011). Heritability of different forms of memory in the Late Onset Alzheimer's Disease Family Study. *J Alzheimers Dis*, *23*(2), 249–255. doi:10.3233/JAD-2010-101515

Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: a tool for genome-wide complex trait analysis. *American Journal of Human Genetics*, *88*(1), 76–82. doi:10.1016/j.ajhg.2010.11.011

7 Declaration by candidate

I declare herewith that I have independently carried out the PhD – thesis entitled „Why complex human phenotypes need complex data analytics - insights from the fields of molecular and cognitive neuroscience“. This thesis consists of original research articles that have been written in cooperation with the enlisted co-authors and have been published in peer-reviewed scientific journals or are in preparation for publication / submitted for publication. Only allowed resources were used and all references used were cited accordingly.

Date: _____

Signature: _____