Strategies for the study of neuropsychiatric disorders using endophenotypes in developing countries: a potential databank from China

Raymond C. K. Chan¹*, Irving I. Gottesman², Xiaojia Ge³ and Pak C. Sham⁴

¹ Neuropsychology and Applied Cognitive Neuroscience Laboratory, Key Laboratory of Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing, China
² Developmental Neuroimaging Laboratory, Departments of Psychiatry and Psychology, University of Minnesota, Minneapolis, MN, USA
³ Institute of Child Development, University of Minnesota, Minneapolis, MN, USA
⁴ Department of Psychiatry and State Key Laboratory for Brain and Cognitive Sciences, The University of Hong Kong, Hong Kong Special Administrative Region, China

Edited by:
Judith M. Ford, Yale University School of Medicine, USA

Reviewed by:
William Iacono, University of Minnesota, USA
Gregory Light, University of California at San Diego, USA

*Correspondence:
Raymond C. K. Chan, Institute of Psychology, Chinese Academy of Sciences, 4A Datun Road, Beijing 100101, China.
e-mail: rckchan@psych.ac.cn, rckchan2003@yahoo.com.hk

INTRODUCTION
The identification of genes with large effects (odds ratios > 1.5) that contribute to susceptibilities to neuropsychiatric disorders such as schizophrenia and bipolar disorder has progressed to genome-wide association studies (GWAS) of large samples. Despite this technological advance, the numbers of confirmed loci have remained remarkably small, explaining little of the substantial heritabilities of these disorders (International Schizophrenia Consortium et al., 2009; Shi et al., 2009). This mirrors the state of affairs for type 2 diabetes (Zeggini et al., 2008; Panoutsopoulou and Zeggini, 2009) and other complex diseases or traits (Manolio et al., 2009). One problem of the current approaches is the complete reliance on clinical diagnosis as a sufficient phenotype to pursue, which does not allow us to use subclinical phenotypic information or endophenotypic information (Gottesman and Gould, 2003) which may be informative for identifying carriers of vulnerable genotypes. Psychiatric diagnoses may represent the joint effects of multiple neurocognitive and psychosocial processes, which in turn are partially determined by genetic polymorphisms in susceptibility genes. A promising research direction is the identification of neurobiological or neurobehavioral characteristics that underlie neuropsychiatric disorders, and to find genetic polymorphisms that determine susceptibility to the disorders through their effects on these characteristics.

The concept of endophenotype for psychopathology was first introduced by Gottesman and Shields (1972, 1973) and has received significant interest after a re-conceptualization in 2003 (Gottesman and Gould, 2003). Related terminologies to present this construct, such as target features (Tsuang et al., 1991; Faraone et al., 1999), phenomes (Mahner and Kary, 1997; Freimer and Sabatti, 2003; Sabb et al., 2008), and intermediate phenotype (Meyer-Lindenberg and Weinberger, 2006; Insel and Cuthbert, 2009) facilitate the same basic strategy for optimal reductionism of diagnostic phenotypes. An endophenotype can be defined as an internal construct that cannot be observed with unaided eyes but can fill the gap between clinical symptoms and syndromes on the one end and putative genes on the other of a causal chain (Gottesman and Gould, 2003; Hasler et al., 2006). Potential endophenotypes should fulfill several criteria such as association with the illness in the population, substantial heritability, state-independence (with exceptions), familial association, co-segregation, and reliability and validity of measurement (Chen and Faraone, 2000; Cornblatt and Malhotra, 2001; Gottesman and Gould, 2003; Chan and Gottesman, 2008). The endophenotype strategy (Ritsner, 2009) has greatly increased statistical power as it legitimizes the use of unaffected close relatives as logical extensions of the study of probands only.

With the completion of the human genome project, the main hurdle for biomedical science in this post-genomic era is how to characterize the manifold human “endophenotypes” from the molecular level to the mind level, and most importantly, detecting “unharmonious minds” in psychiatric disorders (Freimer and Sabatti, 2003; Meyer-Lindenberg and Weinberger, 2006; Glahn et al., 2007; Sabb et al., 2008). The development of cognitive neurosciences, neuropsychology and imaging genetics – a strategy for mapping neural structure and activity as a function of genotype in humans – has encouraged a conceptual transformation by showing that the greater power of endophenotypes lies in using genetic risk variants in the clinically unaffected relatives of cases as tools for the discovery of the mediating neural mechanisms that bridge the gap from DNA...
sequence to pathological behavior. The use of electrophysiology, neurochemistry, and neuropsychology can also be powerful tools to index intermediate neurobiological processes that are influenced by genetic variation and epigenetic processes (Petronis, 2010).

**TRENDS IN DEVELOPED COUNTRIES**

The completion of genome sequencing does not mean we can grasp a more holistic picture about the etiologies of different neuropsychiatric disorders as the study of these disorders is complicated by the polymorphisms of the diseases, the impact of environment, including $G \times E$ interactions (Ge et al., 1996; Plomin et al., 2008) and variable phenotypic manifestations related to developmental stages and epigenetic (Ptak and Petronis, 2008; Kaminsky et al., 2009; Savitz and Drevets, 2009). We are still unable to specify precisely the phenotypes (the observed manifestation of the genotypes such as symptoms observed in psychiatric patients) in those individuals whose genomes we investigate. We continue to depend mainly on standard psychiatric disease diagnoses, which are both incomplete and imprecise, as representations of human phenotypes.

Most of the findings in this field so far are reported from Western countries with several sites doing relevant work. The International human phenotype project (HPP) (e.g., Freimer and Sabatti, 2003) and the International Consortium for Brain Mapping (ICBM) (Mazziotta et al., 2001) are international networks that merged worldwide data to characterize phenotypes. In particular, the HPP is aimed at providing all levels of phenomic information in understanding the diagnoses of human diseases, whereas the ICBM is focused on brain structure and morphological information in human brain. These two projects are too broad and general so that it is not specifically designed for neuropsychiatric disorders. Direct and time-consuming assessments by experts with new tools will probably be essential for collecting most human phenotypic data, especially at this stage of their infancy.

The following consortia or networks are more specifically established for seven neuropsychiatric disorders. The first is a strong network developed at University of California at San Diego with collaborating universities, known as the Consortium on the Genetics of Schizophrenia (COGS) (Greenwood et al., 2007; Braff et al., 2008). It is the pioneering project initiated by the NIMH and aimed at establishing a multisite database to understand the genetic basis of candidate endophenotypes for schizophrenia. The COGS focuses on neurophysiological (e.g., prepulse inhibition, oculomotor antisaccade) and neurocognitive endophenotypes from the U Penn Computerized Neurocognitive Battery, Verbal working memory (LNS) and verbal memory (CVLT) and continuous performances test (Braff et al., 2007; Calkins et al., 2007). They are also performing candidate gene analyses, developing novel statistical genetics approaches, and other methods development. The COGS database is also a resource for other emerging strategies (epigenetics and CNV-based studies). Moreover, they have built up their network for about 7–8 years and have collected several hundred family units (schizophrenic patients, non-psychotic probands, parents) and parallel healthy controls from the community. Moreover, there is an independent database similar to the COGS network by Raquel Gur (Gur et al., 2007). It is more focused in nature by reaching consensus on the most promising traits and concentrates on these, to achieve larger sample sizes and more detailed information across multi-generational families.

A related organization is the MATRICS Psychometric Study initiated by the NIMH and contracted out to the University of California at Los Angeles. Although their original aim was to refine a set of standard psychometric tests or tools suitable for schizophrenia research, their refinement of the tools paves a crucial impact on the selection of potential cognitive endophenotypes for schizophrenia. However, this type of initiative was limited to laboratory-based level of measurement and is relatively difficult to articulate to real life scenarios and functional outcome directly. Most recently, The CNTRICS initiative (R13MH078710) is now being conducted to specifically develop new cognitive neuroscience/translational tasks for drug development and discovery. For example, tests that were deemed well-established and “mature” (mismatching negativity and prepulse inhibition) are not part of CNTRICS. This project attempts to draw consensus about the phenotypes or endophenotypes that are most valuable for schizophrenia treatment research.

Another consortium is specifically examining the cognitive phenotype of psychiatric diseases, namely cognitive phenomics, located in the University of California at Los Angeles (Sabb et al., 2008). Their main aim is specifically designed to traverse multiple disorders, and multiple species to gain traction on phenotypes important to mental disorder and their treatment. Such an approach has two more additional merits for studying neuropsychiatric disorders endophenotypes. First, it bridges the gap between microscopic (e.g., synaptic and molecular abnormalities) and macroscopic (e.g., clinical syndromes observed in schizophrenia and bipolar disorders) knowledge. Second, unlike the COGS, the cognitive phenomics database targets a wider scope of neuropsychiatric disorders in addition to schizophrenia spectrum disorders. Therefore, such an approach may stimulate further similar projects or databases that enable broadly collaborative knowledge-building and translational research. Several other related centers or programs have been set up in European countries and Australia providing comprehensive repositories for phenotype data that can be used on a genome-wide scale for animal and human study (e.g., Mouse Phenome Database; Australian Phenomics Centre, c.f. Freimer and Sabatti (2003)).

One additional issue is that new databases are rapidly being developed to represent gene expression and proteomic knowledge on psychiatric disorders. For example, SchizophreniaGene (SzGene) has been developed by Lars Bertram and colleagues formerly at Harvard Medical School and Massachusetts General Hospital (now in Berlin) to collect and synthesize systematically the genetic data published in peer-reviewed scientific journals. For variants (or polymorphisms) with published phenotypes in at least four independent case-control studies, the Bertram team has also systematically meta-analyzed the available data in an effort to tease out the most promising schizophrenia candidate genes. SzGene currently lists 43 top genes, changing monthly, which include candidate genes involved in dopaminergic neurotransmission (e.g., DRD1, DRD2, DRD4, COMT), genes discovered by fine-mapping of linkage hot-spots (e.g., NRG1, DAOA, DISC1), as well as genes identified by GWAS (e.g., ZFN804A, NOTCH4, RELN, NRGN). A promising newcomer is now released from the Laboratory of Patrick F Sullivan at the University of North

---

1. www.jax.org/phenome
2. www.apf.edu.au

---
Carolina, Chapel Hill, known as the Sullivan Lab Evidence Project (SLEP) (Konneker et al., 2008) and is bound to attract attention and trials. SLEP is a searchable archive of findings from psychiatry genetics that is freely available on the web. Through this searchable archive, researchers can access and retrieve data concerning genome-wide linkage, genome-wide association, and microarray studies for a wide range of neuropsychiatric disorders such as ADHD, autism, bipolar disorder, eating disorders, major depression, alcohol and nicotine dependence, and schizophrenia.

THE SITUATION AND FUTURE ROADMAP IN CHINA

There is no systematic and comprehensive study of endophenotypes for neuropsychiatric disorders in China. There are several sites doing specialized work scattered in different places across the country. The Bio-X Life Science Research Centre of Shanghai Jiao Tong University has been conducting a wide range of genotyping in schizophrenia spectrum disorders. The Chinese National Human Genome Centre of Beijing and the Institute of Bioinformatics of Tsinghua University in Beijing have established a database for schizophrenia candidate genes focusing on variations. The State Key Laboratory of Brain and Cognitive Science of the Chinese Academy of Sciences in Beijing has mainly focused on the underlying cognitive processing of healthy people, using imaging technologies. The Key Laboratory of Mental Health of the Institute of Psychology of the Chinese Academy of Sciences takes an active role in studying the neurocognitive endophenotypes for neuropsychiatric disorders such as schizophrenia, bipolar disorders, and hyperactivity disorders. There are also several other institutes e.g., Chengdu, Nanjing, and Changsha that have been contributing to furthering knowledge about genetic and cognitive processes of neuropsychiatric disorders.

Although we are somewhat behind our western counterparts in this approach, we still stand to make an important and perhaps unique contribution if we can develop a systematic project in China. This program would, in principle, highlight the brain structure, functional connectivity, and neurocognitive function as well as neurological manifestations in the patients and their relatives, non-psychotic probands, and healthy controls. In so doing, we will bridge the big gap between molecular gene levels and macroscopic human mind levels.

The database should have at least two main features. First, data that are common to the western samples of databases are less subject to cultural variation in China, i.e., the supposedly universal basic cognition such as sustained attention, working memory, and inhibitory control. Second, data that are unique to the Chinese population and are more culturally relevant to the Chinese setting, i.e., the supposedly culturally specific social cognitions such as emotion perception and expression, e.g., reading text vertically, early motoric coordination from use of chopsticks. These data will be useful for data merging and cultural comparison. Given these considerations, we propose the establishment of the Consortium on Human Information and Neurocognitive Endophenotypes (CHINE) in China. The CHINE should be composed of three main parts, namely (1) neurocognitive, social cognition, and neurophysiological functions (including ERP, structural, and functional imaging); (2) behavioral genetics and genomic sequencing, and finally (3) a national/centralized twin registry data (healthy and clinical cohorts) [our co-author, the late Professor Ge has already registered more than 2,000 pairs in Beijing].

In summary, our team is optimistic about establishing a database for endophenotypes in mainland China, and invites collaborators to share in the merits and promise of a potentially useful database on phenotypes and endophenotypes for China. This initiative focuses on specific cognitive endophenotypes, e.g., neurological soft signs, which have been emerged in recent literature as important to multiple neuropsychiatric syndromes such as schizophrenia, bipolar disorders, ADHD, and dementia, but not yet included in the COGS. These neurological soft signs may not be specific to schizophrenia but also are found in other neuropsychiatric disorders such as ADHD (Casey et al., 1997; Sergeant et al., 1999; Chan et al., 2010), and major depression (Baldwin et al., 2005), bipolar disorders (Mukherjee and Shukla, 1984; Negash et al., 2004); therefore, the database may be applicable and extended to a wider range of clinical cases with greater public health impact.

CONCLUSIONS

Despite the promising and systematic work initiated by our western counterparts, there are still a number of caveats for such research programs. First, the technology needed to reliably acquire neuropsychological and neurocognitive phenotypes is complex and must be carefully adapted to large multisite-population studies. The COGS group and other subgroups may not be able to collect a huge amount of targeted families and subjects in a short period of time. The use of endophenotypes for genetic studies of the aforementioned kind requires large family and patient samples and multisite collaborations to achieve sufficient statistical power. China is a country with 1.3 billion persons, about 68.7% were in the age range of 15–59 years old (National Bureau of Statistics of China, 2001). One percent (the lifetime prevalence rate of schizophrenia) has already accounted for a large sub-population of affected participants. Such a large potential sample source complements other existing programs and initiatives.

Second, although heritability and familial patterns, as well as animal models, provide some evidence for whether or not an endophenotype reflects important genetic effects, identifying a plausible causative gene through GWAS (association studies and proteomics) is the only definitive answer to that question. The GOGS group in the United States takes two approaches to remediate such limitations: (1) determine the segregation and co-segregation of these phenotypes and families; (2) perform linkage analysis on those phenotypes that appear to show genetic transmission. The results of family studies help identify the genetic aspects of schizophrenia from physiological and cognitive perspectives, and they may also determine which of the various pathophysiological features of schizophrenia have a common genetic basis. The results of the linkage studies ultimately will be used in subsequent projects to identify candidate genes, supported by both linkage and neurobiological findings, for molecular sequencing. Moreover, the next generation of COGS (i.e., COGS-2) is a case-control design and no longer relies on families. Given the known uniqueness of genes in different ethnic groups and the interaction effects with relevant environments, the findings from the Western groups should complement and expand the findings from a non-western culture.
like the Chinese. The identification of overlapping versus distinct genetically linked features of neuropsychiatric disorders such as schizophrenia is a crucial step in the search for targets for intervention as well as validation/prioritization and continued refinements of biomarkers/endophenotype-based approaches.

In conclusion, after reviewing the pros and cons, we suggest a clear rationale for the establishment of a consortium for the study of endophenotypes in China. Such a consortium – CHINE – will be sufficiently large to function as a separate but complementary entity. In contrast to many developed countries, we strongly believe that establishment of such a geographical, cultural, and genetically homogeneous consortium can be more focused on its own cultural characteristics although it is obviously desirable that there is close communication with other consortia so as to identify those aspects of neuropsychiatric disease that are universal in the gene-to-behaviors-pathways. We hope that such a consortium can function as an important resource for the policy makers and stakeholders to use in planning for future treatment regimens and related mental healthy policy.

ACKNOWLEDGMENTS
This study was supported partially by the Key Laboratory of Mental Health, Institute of Psychology, Chinese Academy of Sciences, the Project-Oriented Hundred Talents Programme (07CX031003), the Knowledge Innovation Project of the Chinese Academy of Sciences (KSCX2-YW-R-131), National Science Foundation of China (30770723), National Outstanding Young Investigator Award (National Science Foundation of China), and a grant from National Basic Research Programme of China (973 Program) (2007CB512302) and the Lieber Prize for Outstanding Schizophrenia Research to IIG. The funding agents had no role in the decision to publish, or to prepare the manuscript.
Mukherjee, S., and Shukla, S. (1984). Neurological abnormalities in patients with bipolar disorder. *Biol. Psychiatry* 19, 337–345.

National Bureau of Statistics of China. (2001). *Bulletin of the Fifth National Census*. Retrieved May 29, 2009, from http://www.stats.gov.cn/tjgb/qgrkpcgb/t20020331_15434.htm.

Negash, A., Kebede, D., Alem, A., Melaku, Z., Deyessa, N., Shibire, T., Fekadu, A., Fekadu, D., Jacobsson, L., and Kullgren, G. (2004). Neurological soft signs in bipolar I disorder patients. *J. Affect. Disord.* 80, 221–230.

Panooutsopoulos, K., and Zeggini, E. (2009). Finding common susceptibility variants for complex disease: past, present and future. *Brief. Funct. Genomic. Proteomic.* 8, 345–352.

Petronis, A. (2010). Epigenetics as a unifying principle in the etiology of complex traits and diseases. *Nature* 465, 721–727. doi: 10.1038.

Plomin, R., Haworth, C. M. A., and Davis, O. S. P. (2008). Genetics of learning abilities and disabilities: recent developments from the U.K. and possible directions for research in China. *Acta Psychol. Sin.* 40, 1051–1061.

Ptak, C., and Petronis, A. (2008). Epigenetics and complex disease: from etiology to new therapeutics. *Annu. Rev. Pharmacol. Toxicol.* 48, 257–276.

Ritzler, M. S. (Ed.). (2009). *The Handbook of Neuropsychiatric Biomarkers*, *Endophenotypes and Genes*, Vol. 4. Frankfurt: Springer.

Sabb, F. W., Bearden, C. E., Glahn, D. C., Parker, D. S., Freimer, N., and Bilder, R. M. (2008). A collaborative knowledge base for cognitive phenomics. *Mol. Psychiatry* 13, 350–360.

Savitz, J. B., and Drevets, W. C. (2009). Imaging phenotypes of major depressive disorder: genetic correlates. *Neuroscience* 164, 300–330.

Sergeant, J. A., Van der Meere, J. J., and Oosterlaan, J. (1999). “Information processing and energetic factors in attention-deficit/hyperactivity disorder,” in *Handbook of Disruptive Behavior Disorders*, eds H. C. Quay and A. Hogan (New York: Plenum), 75–104.

Shi, J., Levinson, D. F., Duan, J., Sanders, A. R., Zheng, Y., Pe’er, I., Dudbridge, F., Holmans, P. A., Whitemore, A. S., Mowry, B. J., Ollinck, A., Amin, F., Cloninger, C. R., Silverman, J. M., Buccola, N. G., Byerley, W. F., Black, D. W., Crowe, R. R., Oksenberg, J. R., Mirle, D. B., Kendler, K. S., Freedman, R., and Gejman, P. V. (2009). Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature* 460, 753.

Tuang, M. T., Gilbertson, M. W., and Faraone, S. V. (1991). The genetics of schizophrenia: current knowledge and future directions. *Schizophr. Res.* 4, 157–171.

Zeggini, E., Scott, L. J., Saxena, R., Voight, B. F., Marchini, J. L., Hu, T., de Bakker, P. I. P., Abecasis, G. R., Almgren, P., Andersen, G., Ardlie, K., Bostrom, K. B., Bergman, R. N., Bonnycastle, L. L., Borch-Johnsen, K., Burt, N. P., Chen, H., Chines, P. S., Daly, M. J., Dooiho, P., Ding, C.-J., Donney, A. S., Duren, W. L., Elliott, K. S., Erdos, M. R., Frayling, T. M., Freathy, R. M., Gianninni, L., Grallert, H., Grarup, N., Groves, C. J., Guiducci, C., Hansen, T., Herder, C., Hitman, G. A., Hughes, T. E., Izomma, B., Jackson, A. U., Jorgensen, T., Kong, A., Kubalanza, K., Kuruvilla, F. G., Kuusisto, J., Langenberg, C., Lango, H., Lauritzen, T., Li, Y., Lindgren, C. M., Lysenko, V., Marville, A. F., Meisinger, C., Midthuell, K., Moblie, K. L., Morken, M. A., Morris, A. D., Nairus, N., Nilsson, P., Owen, K. R., Palmer, C. N. A., Payne, F., Perry, J. R. B., Pettersen, E., Platou, C., Prokopenko, I., Qi, L., Qin, L., Rayner, N. W., Rees, M., Roix, I. J., Sandbach, A., Shields, B., Sjogren, M., Steinhorsdottir, V., Stringham, H. M., Swift, A. J., Thorleifsson, G., Thorsteinsdottir, U., Timpons, N. J., Tuomi, T., Tuomilehto, J., Waller, M., Watanabe, R. M., Weedon, M. N., Willer, C. J., Illig, T., Hveem, K., Hu, F. B., Laakso, M., Stefansson, K., Pedersen, O., Wareham, N. J., Barroso, L., Hattersley, T. A., Collins, F. S., Groop, L., McCarthy, M. I., Bodnue, M., and Altshuler, D. (2008). Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat. Genet.* 40, 638–645.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 17 May 2010; accepted: 08 October 2010; published online: 28 October 2010. Citation: Chan RCK, Gottesman II, Ge and Sham PC (2010) Strategies for the study of neuropsychiatric disorders using endophenotypes in developing countries: a potential database from China. *Front. Hum. Neurosci.* 4:207. doi: 10.3389/fnhum.2010.00207

Copyright: © 2010 Chan, Gottesman, Ge and Sham. This is an open-access article subject to an exclusive license agreement between the authors and the Frontiers Research Foundation, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.