

ARTICLE

# Genome-wide association of major depression: description of samples for the GAIN Major Depressive Disorder Study: NTR and NESDA biobank projects

Dorret I Boomsma<sup>\*1</sup>, Gonneke Willemsen<sup>1</sup>, Patrick F Sullivan<sup>2</sup>, Peter Heutink<sup>3</sup>, Piet Meijer<sup>4</sup>, David Sondervan<sup>3</sup>, Cornelis Klufft<sup>4</sup>, Guus Smit<sup>5</sup>, Willem A Nolen<sup>6</sup>, Frans G Zitman<sup>7</sup>, Johannes H Smit<sup>8</sup>, Witte J Hoogendijk<sup>8</sup>, Richard van Dyck<sup>8</sup>, Eco JC de Geus<sup>1</sup> and Brenda WJH Penninx<sup>6,7,8</sup>

<sup>1</sup>Netherlands Twin Register, Department of Biological Psychology, VU University Amsterdam, Amsterdam, Netherlands; <sup>2</sup>Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; <sup>3</sup>Medical Genomics Section, Department of Clinical Genetics, VU Medical Center, Amsterdam, Netherlands; <sup>4</sup>Gaubius Laboratory, TNO-PG, Leiden, Netherlands; <sup>5</sup>Department of Molecular and Cellular Neurobiology, Faculty of Earth and Life Sciences, VU University Amsterdam, Amsterdam, Netherlands; <sup>6</sup>Department of Psychiatry, University Medical Center Groningen, University of Groningen, Groningen, Netherlands; <sup>7</sup>Department of Psychiatry, Leiden University Medical Center, Leiden, Netherlands; <sup>8</sup>Department of Psychiatry, EMGO Institute, Institute of Neuroscience, VU Medical Center, Amsterdam, Netherlands

To identify the genomic regions that confer risk and protection for major depressive disorder (MDD) in humans, large-scale studies are needed. Such studies should collect multiple phenotypes, DNA, and ideally, biological material that allows gene expression analysis, transcriptomic, proteomic, and metabolomic studies. In this paper, we briefly review linkage studies of MDD and then describe the large-scale nationwide biological sample collection in Dutch twin families from the Netherlands Twin Register (NTR) and in participants in the Netherlands Study of Depression and Anxiety (NESDA). Within these studies, 1862 participants with a diagnosis of MDD and 1857 controls at low liability for MDD have been selected for genome-wide genotyping by the US Foundation for the National Institutes of Health Genetic Association Information Network. Stage 1 genome-wide association results are scheduled to be accessible before the end of 2007. Genome-wide association results are open-access and can be viewed at the dbGAP web portal (<http://www.ncbi.nlm.nih.gov>). Approved users can download the genotype and phenotype data, which have been made available as of 9 October 2007.

*European Journal of Human Genetics* (2008) 16, 335–342; doi:10.1038/sj.ejhg.5201979; published online 16 January 2008

**Keywords:** major depression; GAIN; biological sample collection; genome-wide association analysis; linkage review

\*Correspondence: Professor DI Boomsma, Department of Biological Psychology, VU University, Van der Boerhorststraat 1, Amsterdam 1081 BT, The Netherlands.

Tel: +31 20 598 8787; Fax: +31 20 598 8832;

E-mail: dorret@psy.vu.nl

Received 26 June 2007; revised 13 November 2007; accepted 15 November 2007; published online 16 January 2008

## Introduction

The defining features of major depressive disorder (MDD) are marked and persistent depressed mood, associated with physical and cognitive signs and symptoms (anhedonia, insomnia or hypersomnia, appetite and weight changes, psychomotor agitation or retardation, fatigue or loss of energy, excessive guilt or worthlessness, poor

concentration or indecisiveness, and recurrent thoughts of death or suicide). MDD is distinct from 'normal' sadness by its persistence (ie,  $\geq 2$  weeks), the presence of accompanying signs and symptoms, and its associated impairment. The definition of MDD excludes other disorders that can have substantial depressive symptoms. These include bipolar disorder, schizophrenia, schizoaffective disorder, and depressive symptoms resulting from another disorder (eg, medical condition like hypothyroidism).<sup>1</sup>

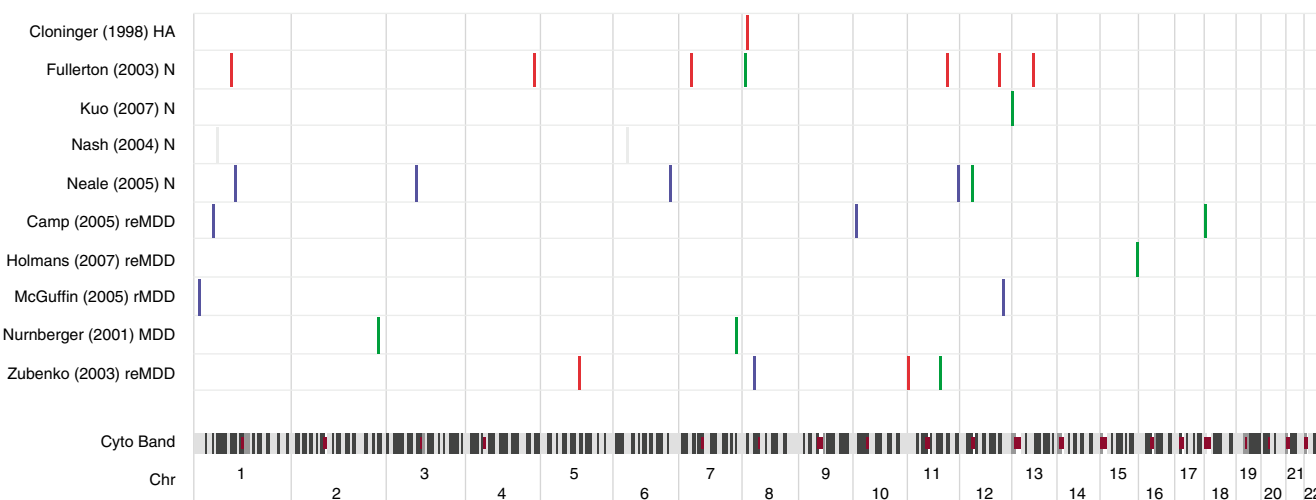
Combining available estimates, the lifetime prevalence of MDD is  $\sim 15\%$  and is twofold higher in women than in men.<sup>2</sup> The course of MDD is typified by recurrence of illness. In a meta-analysis, 76% had  $\geq 1$  recurrence over a 10-year follow-up.<sup>3</sup> MDD is associated with large morbidity (greater than chronic medical conditions such as diabetes and arthritis),<sup>4–6</sup> excess mortality from suicide and other causes,<sup>7–10</sup> and substantial direct and indirect costs ( $> \$43$  billion/year in the US).<sup>11</sup> The WHO projected MDD to be the second leading cause of disability worldwide by 2020.<sup>12</sup> Thus, MDD is a first-rank public health problem.

It has been firmly established that variation in MDD liability is in part genetic.<sup>13</sup> Family studies find a significantly higher lifetime prevalence of MDD in biological relatives of MDD probands (pooled odds ratio = 2.84, 95% CI = 2.31–3.49). Twin studies have established that the familial component reflects genetic vulnerability rather than shared environmental risk. The meta-analytic heritability estimate is 37% (95% CI = 31–42%) with a minimal contribution of environmental effects common to siblings (0, 95% CI = 0–5%), and substantial individual-specific environmental effects/measurement error (63, 95% CI = 58–67%).

Despite the evidence for heritability of MDD, as with most complex traits, identification of vulnerability genes has not been yet very successful. Results from studies of first-stage genome-wide linkage scans for MDD or related personality traits are summarized in Figure 1. One study was of MDD,<sup>14</sup> two of recurrent MDD,<sup>15,16</sup> and three of recurrent, early-onset MDD.<sup>17–19</sup> Five studies investigated quantitative traits related to MDD, that is harm avoidance<sup>20</sup> and neuroticism.<sup>21–24</sup> Two studies were secondary analyses of cohorts selected for alcoholism<sup>14</sup> or nicotine dependence.<sup>23</sup>

Results from linkage studies of MDD and associated personality traits are summarized in Figure 1. Regions on chromosomes 1, 4, 5, 7, 8, 11, 12, and 13 show a significant linkage signal (LOD  $> 3$ ) in at least one study. However, taken together these studies do not indicate a large degree of replication. This lack of success so far, may be due to the relatively small sample sizes of most linkage studies. Likewise, the generally small sizes, the different definitions of the phenotype, and the limited number of polymorphisms that could be included in association studies of candidate genes have not lead to consistent results. This is even true for candidate genes that have been assessed in large sized studies.<sup>25–27</sup>

However, since 2000 there have been exceptional advances in our knowledge of the human genome along with a precipitous drop in the cost of genotyping. These advances directly led to genome-wide association studies whereby large case–control samples are individually genotyped for 500 000 or more single nucleotide polymorphisms. Examples where genome-wide association studies has provided new genetic insights include the



**Figure 1** First-stage genome-wide linkage scans for MDD or related personality traits (neuroticism and harm-avoidance). The x-axis shows the human genome from 1ptel to 22qtel. Each row shows the findings for that study. Studies of bipolar disorder, candidate region linkage studies, fine-mapping results, and statistical analyses of epistasis or studies that included covariate models were excluded. The width of the coloured bars indicates the genomic physical distance/region implicated by a particular study. LOD/equivalents are plotted. Key: red LOD  $\geq 3$ , green LOD  $\geq 2$ , blue LOD  $\geq 1.5$ , grey otherwise.

*CFH* gene and age-related macular degeneration,<sup>28</sup> the *FTO* gene and body mass index in children and adults,<sup>29</sup> the *TCF7L2* gene and type-2 diabetes<sup>30</sup> and the *IL23R* gene and inflammatory bowel disease.<sup>31</sup> Recently, the Wellcome Trust Case Control Consortium<sup>32</sup> identified 24 association signals for seven major diseases, including bipolar disorder, using 2000 cases and 3000 controls.

In 2006, a consortium of investigators at the VU University Amsterdam, universities in Groningen and Leiden, and UNC-Chapel Hill in the US was selected for GWA genotyping as one of the six Genetic Association Information Network (<http://www.fnih.org/GAIN>) studies.<sup>33</sup> The GAIN initiative is a component of the public-private partnership of the Foundation for the National Institutes of Health Inc., and funding is via Pfizer, Affymetrix, and Abbott Laboratories. GWA genotyping for 600 000 single nucleotide polymorphisms in MDD cases and population-based controls has been conducted by Perlegen Sciences. Stage 1 genome-wide association results are scheduled to be available before the end of 2007 and will be available to the scientific community via the NCBI dbGAP web portal (<http://www.ncbi.nlm.nih.gov>). Approved users can download genotype and phenotype data with the restriction that use of genotype and phenotype data is restricted to psychiatric health and related somatic conditions.

In this paper, we describe the approach and logistics of the two large-scale projects in the Netherlands collecting biological samples for genetic studies, which together constitute the GAIN-MDD study. We give an overview of the phenotyping for MDD and the selection of case and control subjects for GWA genotyping in the GAIN-MDD study.

## Subjects

Subjects eligible for inclusion in the GWA study of MDD come from two large longitudinal projects: 1702 depressed cases come from the Netherlands Study of Depression and Anxiety (NESDA, [www.nesda.nl](http://www.nesda.nl)), and 1700 controls come from the Netherlands Twin Registry (NTR, [www.tweelingenregister.org](http://www.tweelingenregister.org)). In addition, 160 cases from the NTR and 157 controls from the NESDA study were included in the original selection, and both parents of 33 controls to form 33 trios and 21 duplicate samples.

### MDD cases

Depressed subjects (cases) are mainly derived from NESDA, a longitudinal cohort study designed to be representative of those with depression and anxiety disorders in different health care settings and in different stages of the developmental history of the disorders. Details on objectives, recruitment, and methods of NESDA have been described elsewhere.<sup>34</sup> In short, recruitment of participants took place from September 2004 through February 2007. Cases

from mental health care organizations were recruited from seven outpatient regional facilities in the Netherlands (total catchment area = 1 175 000 persons). When new patients were diagnosed with a depressive and/or anxiety disorder during a standardized intake assessment, they were asked to participate in NESDA. For recruitment of respondents in primary care, a three-stage screening procedure was used. A random sample of patients who consulted a GP in the last 4 months, irrespective of the reason for consultation, filled out a screening questionnaire.<sup>35</sup> Those who screened positive were interviewed by phone with the short-form of the Composite International Diagnostic Interview (CIDI).<sup>36</sup> Those with current MDD or an anxiety disorder were asked to participate in NESDA and were invited for a baseline assessment, which included a full CIDI interview. A community sample of NESDA cases was derived from the Netherlands Mental Health Survey and Incidence Study, a community-based study examining the prevalence of psychiatric disorders in the Netherlands.<sup>37</sup> A multistage, stratified, random sampling procedure recruited 7076 respondents from different households in 90 Dutch municipalities. Those who were diagnosed with lifetime MDD or anxiety disorder during one of the CIDI interviews in 1996, 1997 or 1999, were approached for participation in NESDA. Finally, the NESDA sample includes a subgroup previously recruited for the Adolescents at Risk for Anxiety and Depression Study,<sup>38</sup> a prospective cohort study examining the development and course of MDD and anxiety disorders among the offspring (18–25 years) of parents with depressive or anxiety disorders. Potential cases in the NTR were identified based on a multivariate composite score that combined survey data on depression, anxiety, and neuroticism.<sup>39</sup> From this group of potential cases, we selected unrelated subjects, who received a CIDI interview by telephone.

Regardless of recruitment setting, similar inclusion, and exclusion criteria were used to select MDD cases. Inclusion criteria were a lifetime diagnosis of DSM-IV MDD as diagnosed through the CIDI psychiatric interview (section E, version 2.1), an age between 18 and 77 years and self-reported western European ancestry. Persons who were not fluent in Dutch and those with a primary diagnosis of a psychotic disorder, obsessive compulsive disorder, bipolar disorder, or severe substance use dependence were excluded. The 1862 cases included in GAIN, were recruited through mental health care organizations (785), primary care practitioners (603), and community samples (218 Netherlands Mental Health Survey and Incidence Study, 96 Adolescents at Risk for Anxiety and Depression Study and 160 NTR).

### Control subjects

Control subjects are mainly derived from the NTR, which has collected longitudinal data by mailed surveys every 2–3 years since 1991. Data collection has occurred in 1991,

1993, 1995, 1997, 2000, 2002/3, and 2004/5 and takes place in twins and their family members. There are nearly 22 000 participants from 5546 families (some families are linked and come from larger pedigrees). The majority of families were recruited when the twins were young adults through City Council registration systems in 1990–91 and in 1992–93. After 1993 an effort was made to recruit adult and older twins through a variety of approaches. Details on recruitment, response rates, response bias, and demographic characteristics of the sample have been described previously.<sup>40</sup> The sex distribution is 40% men and 60% women for twins, 45 and 55% for the siblings of twins, 47 and 53% for parents, and 64 and 36% for spouses. Longitudinal phenotyping by survey includes assessment of depressive symptoms (multiple instruments), anxiety, neuroticism, and other personality measures. Additionally, subjects are asked about life events, lifestyle, education, health, and religious background.

From the group of NTR participants for whom both survey data and biological samples were available, we selected controls for the MDD cases. Controls never scored high ( $>0.65$ ) on a general factor score for anxious depression. The factor score is a combined measure of neuroticism, anxiety, and depressive symptoms assessed via longitudinal questionnaires.<sup>39</sup> It has a mean of 0 and a SD of 0.7. Subjects never reported a history of MDD in any survey or at the blood sampling visit (either as a complaint for which treatment was sought from a specialist, reported medication use, or via the CIDI). Controls and their parents were born in the Netherlands or western Europe. If there were multiple eligible controls in a family, we first matched on sex and age, and used the highest number of completed questionnaires as an additional criterion. Only biologically unrelated subjects were selected.

The NESDA controls (133 from general practice, 24 from Adolescents at Risk for Anxiety and Depression Study) came from a larger healthy control group. Controls did not have a lifetime diagnosis of MDD or an anxiety disorder as assessed by the CIDI, and all controls reported low depressive symptoms at baseline ( $<16$  on K-10 (29),  $<4$  on Inventory of Depressive Symptoms<sup>41</sup>).

### IRB approvals and informed consent

The NESDA and NTR studies were approved by the Central Ethics Committee on Research involving human subjects of the VU University Medical Center, Amsterdam, an Institutional Review Board certified by the US Office of Human Research Protections (IRB number IRB-2991 under Federal wide Assurance-3703; IRB/institute codes, NESDA 03-183; NTR 03-180). All subjects provided written informed consent. As part of the GAIN application process, consent forms were re-reviewed. For NESDA, only 22 respondents refused informed consent for genetic research

(1.7% of all respondents approached). NTR participants were approached with a broad consent form for the entire NTR biobank project.

### Phenotyping

#### MDD, depressive symptoms, and other psychopathology indicators

**NESDA** The Composite International Diagnostic Interview, section E, version 2.1<sup>36</sup> was used to diagnose MDD. The interview also provides information on age of onset, number of episodes of MDD, and specific symptoms of depression. Information on lifetime comorbid panic disorder with or without agoraphobia, generalized anxiety disorder, social phobia, alcohol use, and dependence was also collected with the CIDI. Depression and anxiety severity indicators include the Inventory of Depressive Symptoms-self-report (IDS-SR<sup>41</sup>), the Fear Questionnaire<sup>42</sup> and the Beck Anxiety Inventory.<sup>43</sup> Neuroticism, an endophenotype for MDD, was assessed with the NEO.<sup>44</sup> The family tree inventory was used to examine depression in first degree relatives.

**NTR** Phenotyping for depression in the NTR survey studies took place with the Beck depression inventory (BDI<sup>45</sup>) in 1993 and 1997 and with the depression scale from the YASR<sup>46</sup> in 1991, 1995, 1997, and 2000. Neuroticism was assessed with the ABV<sup>47</sup> in five out of six surveys (not in 1995) and with the NEO<sup>44</sup> in 2004. Anxiety (STAI<sup>48</sup>) was assessed at five out of seven surveys (not in 1995 and 2004). NTR cases and a subsample of controls underwent the CIDI protocol by phone; either as part of an earlier study<sup>49</sup> on the heritability of major depression and anxiety disorders or as part of the selection procedure for GAIN.

### Biobank procedures and DNA isolation

Before the start of the NESDA and NTR biological sample collection, processing and storage protocols were harmonized. Blood sampling for the NESDA participants took place during the baseline visit and was done between 0830–0930 h at one of the seven field sites, all at walking distance of laboratories facilities for immediate processing.

For NTR, biological samples were taken at the respondents' home between 0700 and 1000 h. Starting in 2004, adult participants registered with the NTR were invited by letter to participate in a project in which blood and morning urine samples would be collected. Eligible participants were  $\geq 18$  years, had returned at least one survey or came from a family in which at least one person completed a questionnaire. The letter was followed by a telephone call to schedule an appointment for a home visit. With fertile women, an appointment was made for the third to fifth day of the menstrual cycle when possible. For women taking oral contraceptives, an appointment

was made for their pill-free week. In twin individuals, we also took buccal swabs for DNA isolation<sup>50</sup> as blood group chimerism in twins is not rare.<sup>51</sup> Tubes were stored, as appropriate, in melting ice (0–2°C) during transport, at room temperature, or in an insulated box with a constant temperature of 37°C. Effects of transportation and storage on blood quality, RNA with and without challenge, lymphocytes and other parameters were examined in a pilot project<sup>52</sup> and appear to be negligible.

All venous blood samples were taken after overnight fasting using a safety-lock butterfly needle. For NESDA participants, a total of 10 blood tubes were drawn in the following order: 3 × 7 ml and 1 × 2 ml EDTA, 2 × 7 ml and 1 × 2 ml Heparin, 1 × 4.5 ml CTAD, 1 × 2 ml NaF, and 1 × 6 ml ACD. For the NTR participants, a total of seven blood tubes were drawn in the following order: 2 × 9 ml EDTA, 2 × 9 ml Heparin, 1 × 4.5 ml CTAD, 1 × 7 ml serum, and 1 × 2 ml EDTA. For DNA isolation blood collected in the EDTA anticoagulant tubes was used. For DNA isolation of NESDA samples the Qiagen FlexiGene® DNA AGF3000 Kit for large volumes of fresh whole human blood was used on an AutoGenFlex 3000 workstation (Autogen, Holliston, MA, USA). The GENTRA Puregene® DNA isolation kit (manual) was used on frozen whole blood samples of the NTR. DNA concentrations were determined using the PicoGreen® dsDNA Quantitation kit from Molecular Probes. All procedures were performed according to the manufacturer's protocols.

### GAIN sample

For the case and the control groups we selected unrelated subjects. Baseline characteristics of cases and controls, and of an additional NTR comparison group, are summarized in Tables 1 and 2. Tables 1 and 2 describe the data as deposited on the dbGaP website ([www.ncbi.nlm.nih.gov/dbgap](http://www.ncbi.nlm.nih.gov/dbgap), 9 October 2007). After excluding subjects with genotyping

or sample problems, data from 1821 MDD cases and 1822 controls are available for analyses. The actual numbers for genome-wide association analyses may turn out to be slightly smaller, as extra analyses of the genotype data may lead to removal of additional subjects. For both MDD cases and for controls there is information available on the website about demographics, personality traits (eg, neuroticism), and lifestyle (eg, smoking and alcohol use). The phenotype information for NESDA respondents is at present mainly from the baseline interview data. The baseline interview took 4 h on average and consisted of the CIDI, a series of self-report questionnaires, a medical interview, and a fasting blood draw. For NTR respondents, a total of up to seven surveys were mailed to the participating twin families over a period of 15 years. NTR respondents who took part in the biobank protocol were

**Table 2** Characteristics of NTR comparison sample and GAIN controls

	NTR comparison sample N = 4111	GAIN controls N = 1822 <sup>a</sup>
Age (years)	43.6 ± 15.3	45.1 years ± 14.1
Female, %	59.3	61.8
Educational level (% low/middle/high)	9.2/61.7/29.2	5.7/56.2/38.1
With partner, %	82.9	87.0
Smoking (current), %	37.9	20.4
Alcohol use (last year), %	79.4	80.0
NEO neuroticism	31.7 ± 6.3	28.2 ± 5.5
ABV neuroticism	49.9 ± 25.3	36.7 ± 16.0
ABV somatic complaints	17.9 ± 5.6	15.9 ± 3.8
STAI trait anxiety	34.1 ± 8.8	30.1 ± 5.6
YASR anxious depression	5.0 ± 4.2	3.22 ± 2.7
Beck depression	2.1 ± 2.8	1.11 ± 1.4

For continuous traits, the mean score and SD are listed.

<sup>a</sup>158 controls are from NESDA.

**Table 1** Characteristics of MDD cases by setting and total MDD group

	Community N = 446 <sup>a</sup>	MDD cases by setting		Total MDD cases N = 1821
		Primary care N = 772	Mental health care N = 603	
Age (years)	41.2 ± 13.9	45.9 ± 11.7	37.9 ± 11.4	42.1 ± 12.7
Female, %	72.0	72.3	64.7	69.7
Educational level (% low/middle/high)	7.6/62.7/29.6	7.0/54.5/38.5	8.1/66.2/25.8	7.5/60.4/32.1
With partner, %	72.5	69.8	63.8	68.5
Smoking (current), %	40.1	39.2	47.3	42.1
Alcohol use (last year), %	63.7	67.5	66.7	66.3
NEO neuroticism score	35.8 ± 8.1	38.6 ± 8.0	42.7 ± 6.6	39.3 ± 8.0
Family history of MDD, % yes	89.2	85.0	83.6	85.4
Depression; age of onset	26.8 ± 12.5	29.1 ± 12.7	25.8 ± 11.6	27.5 ± 12.3
IDS depression <sup>b</sup>	17.2 ± 11.3	23.8 ± 12.1	33.0 ± 12.7	25.8 ± 13.5
BAI anxiety <sup>b</sup>	8.2 ± 7.5	13.2 ± 9.4	18.9 ± 11.5	14.3 ± 10.6

For continuous traits, the mean score and SD are listed.

<sup>a</sup>Community cases (446): 214 from NEMESIS, 95 from ARIADNE, 137 from NTR.

<sup>b</sup>Symptom scores are assessed at study baseline, which does not always equal the period of MDD episode.

visited at home for an additional medical interview, a fasting blood draw, and collection of urine samples.

Table 1 presents the phenotype data for MDD cases, as a function of the recruitment setting, and for the total group. Patients from mental care are on average somewhat younger, have less often attained a higher educational level, less often have a spouse/partner and tend to smoke more heavily. Their profiles on the neuroticism and depression scales are more unfavourable. From comparing the total MDD group to the GAIN control group in Table 2, it may be seen that the control group is slightly older, better educated, and more often with partner/spouse. With respect to lifestyle, MDD cases smoke more often (42 vs 20%) but are less often current drinkers (66 vs 80%).

Table 2 shows characteristics of GAIN controls. In addition, Table 2 provides descriptive statistics, based on the survey data, for a large group of NTR participants who are not in GAIN. The subjects in the comparison group were selected as follows: the subject was not included in GAIN and was not a family member of a GAIN participant; was 18 years or older and survey data on personality traits were available. Only unrelated persons were included in the comparison group and subjects from the comparison group were matched on age and sex with the GAIN control group. Table 2 shows that the GAIN controls are 'hyper-controls': they are better educated, are more often married, and smoke less often. Their scores on personality traits related to MDD and on depression indices from the survey data show a somewhat more favourable profile than the NTR comparison sample.

## Discussion

Progress in unravelling the genetic architecture of multifactorial traits such as MDD has been slow. The reasons have been noted numerous times and include inadequacies in study design, inappropriate sample-recruitment strategy, and inadequately powered samples.<sup>53</sup> We have tried to carefully select patients and controls who are at low liability of developing MDD. Given the sample size of our study we have 80% power to detect a QTL with genotype relative risk of 1.33 if the minor allele frequency is at least 0.40 (under the assumption of a log-additive model with a population prevalence for MDD of 15%<sup>54</sup>). At 80% power, the genotype relative risk increases to 1.38 and 1.58 if the minor allele frequency is 0.25 and 0.10, respectively. The majority of genetic studies performed to date have had insufficient power to uncover genetic variants of small effect.

The genetic basis of liability to MDD and related disorders has been established through meta-analysis of data from twin and family studies.<sup>13</sup> The meta-analytic heritability estimate was 37%. We found in the Dutch population a comparable estimate of heritability, based on

twice the correlation in dizygotic twin and sibling pairs, of 36% for major depression.<sup>49</sup> Furthermore, no qualitative or quantitative sex differences were observed in genetic architecture, which is a promising starting point for genetic linkage and association analyses, as it allows pooling of data from men and women.

The GAIN MDD study will generate GWA data on MDD cases and controls at low risk of MDD derived from two large biobank enterprises in the Netherlands. All cases were diagnosed as having DSM-IV MDD using the CIDI psychiatric interview. Cases are a highly representative sample of those affected with this disorder in the Netherlands through ascertainment in a variety of health care settings. The control sample consisted of so-called 'hyper-controls' who were selected from the lower end of the relevant risk distribution.<sup>32</sup> This should increase the power of the study, compared to the use of unselected control samples.

One of the major strengths of the NTR and NESDA studies is that participants are followed longitudinally and that additional phenotyping and new biological sample collection are both possible. This means that genome-wide genotype data can be used in future studies of, for example, treatment outcomes or of MDD trajectories. Also, to follow up on the independently replicated genes, detailed information is available on two major theoretical correlates of MDD, that is, hypothalamic-pituitary-adrenal axis (HPA-axis) and autonomic nervous system functioning. Additional association to such correlated biological phenotypes would give construct validity to the identified risk genotypes and point to biological mechanisms. A genotype that is associated with both depression and increased sympathetic nervous system activity, for instance, would suggest that the biological action of the gene may involve the noradrenergic synapse and its signalling cascade.

NESDA and NTR subjects have provided information on major life events, and will, in follow-up studies, also be queried about the occurrence of these events. Longitudinal data, including the course of MDD, will be available from NESDA follow-up assessments at 1, 2, 4, 6, and 8 years after baseline. For NTR participants, up to seven surveys are available and the eighth survey is scheduled for 2008. These data can be used for Stage 2 studies, in which gene-environment interactions are assessed. Further plans for future use of the GAIN data include data pooling meta-analytic studies of MDD and other psychiatric disorders and the detailed analyses of genotypic and phenotypic data in monozygotic twins who are MDD discordant.

## Acknowledgements

*We would like to acknowledge support from NWO: genetic basis of anxiety and depression (904-61-090); resolving cause and effect in the association between exercise and well-being (904-61-193); twin-family database for behavior genomics studies (480-04-004); twin research focusing on behavior (400-05-717), Center for Medical*

Systems Biology (NWO Genomics); *Spinozapremie* (SPI 56-464-14192); Centre for Neurogenomics and Cognitive Research (CNCR-VU); *genomewide analyses of European twin and population cohorts* (EU/QLRT-2001-01254); *genome scan for neuroticism* (NIMH R01 MH059160); *Geestkracht program of ZonMW* (10-000-1002); *matching funds from universities and mental health care institutes involved in NESDA* (GGZ Buitenamstel-Geestgronden, Rivierduinen, University Medical Center Groningen, GGZ Lentis, GGZ Friesland, GGZ Drenthe). Major funding for this project is from the Genetic Association Information Network of the Foundation for the US National Institutes of Health, a public-private partnership between the NIH and Pfizer Inc., Affymetrix Inc. and Abbott Laboratories.

## References

- 1 American Psychiatric Association: *Diagnostic and Statistical Manual of Mental Disorders*. 4th edn. Washington, DC: American Psychiatric Association, 1994.
- 2 Weissman MM, Bland R, Joyce PR, Newman S, Wells JE, Wittchen H-U: Sex differences in rates of depression: cross-national perspectives. *J Affect Disord* 1993; **29**: 77–84.
- 3 Piccinelli M, Wilkinson G: Outcome of depression in psychiatric settings. *Br J Psychiatry* 1994; **164**: 297–304.
- 4 Wells KB, Stewart A, Hays RD *et al*: The functioning and well-being of depressed patients: results from the Medical Outcomes Study. *JAMA* 1989; **262**: 914–919.
- 5 Broadhead WE, Blazer DG, George LK, Tse CK: Depression, disability days, and days lost from work in a prospective epidemiologic survey. *JAMA* 1990; **264**: 2524–2528.
- 6 Judd LL, Paulus MP, Wells KB, Rapaport MN: Socioeconomic burden of subsyndromal depressive symptoms and major depression in a sample of the general population. *Am J Psychiatry* 1996; **153**: 1411–1417.
- 7 Tsuang MT, Woolson RF: Excess mortality in schizophrenia and affective disorders. *Arch Gen Psychiatry* 1978; **35**: 1181–1185.
- 8 Berglund M, Nilsson K: Mortality in severe depression: a prospective study including 103 suicides. *Acta Psychiatr Scand* 1987; **76**: 372–380.
- 9 Black DW, Winokur G, Nasrallah A: Is death from natural causes still excessive in psychiatric patients? *J Nerv Ment Dis* 1987; **175**: 674–680.
- 10 Zilber N, Schufman N, Lerner Y: Mortality among psychiatric patients – the groups at risk. *Acta Psychiatr Scand* 1989; **79**: 248–256.
- 11 Greenberg PE, Stiglin LE, Finkelstein SN, Berndt ER: The economic burden of depression in 1990. *J Clin Psychiatry* 1993; **54**: 405–418.
- 12 Murray CJL, Lopez AD: Evidence-based health policy: lessons from the Global Burden of Disease Study. *Science* 1996; **274**: 740–743.
- 13 Sullivan PF, Neale MC, Kendler KS: Genetic epidemiology of major depression: review and meta-analysis. *Am J Psychiatry* 2000; **157**: 1552–1562.
- 14 Nurnberger Jr JI, Foroud T, Flury L *et al*: Evidence for a locus on chromosome 1 that influences vulnerability to alcoholism and affective disorder. *Am J Psychiatry* 2001; **158**: 718–724.
- 15 McGuffin P, Knight J, Breen G *et al*: Whole genome linkage scan of recurrent depressive disorder from the depression network study. *Hum Mol Genet* 2005; **14**: 3337–3345.
- 16 Holmans P, Weissman MM, Zubenko GS *et al*: Genetics of recurrent early-onset major depression (GenRED): final genome scan report. *Am J Psychiatry* 2007; **164**: 248–258.
- 17 Zubenko GS, Maher B, Hughes III HB *et al*: Genome-wide linkage survey for genetic loci that influence the development of depressive disorders in families with recurrent, early-onset, major depression. *Am J Med Genet B Neuropsychiatr Genet* 2003; **123**: 1–18.
- 18 Holmans P, Zubenko GS, Crowe RR *et al*: Genomewide significant linkage to recurrent, early-onset major depressive disorder on chromosome 15q. *Am J Hum Genet* 2004; **74**: 1154–1167.
- 19 Camp NJ, Lowry MR, Richards RL *et al*: Genome-wide linkage analyses of extended Utah pedigrees identifies loci that influence recurrent, early-onset major depression and anxiety disorders. *Am J Med Genet B Neuropsychiatr Genet* 2005; **135**: 85–93.
- 20 Cloninger CR, Van Eerdewegh P, Goate A *et al*: Anxiety proneness linked to epistatic loci in genome scan of human personality traits. *Am J Med Genet* 1998; **81**: 313–317.
- 21 Fullerton J, Cubin M, Tiwari H *et al*: Linkage analysis of extremely discordant and concordant sibling pairs identifies quantitative-trait loci that influence variation in the human personality trait neuroticism. *Am J Hum Genet* 2003; **72**: 879–890.
- 22 Nash MW, Huezio-Diaz P, Williamson RJ *et al*: Genome-wide linkage analysis of a composite index of neuroticism and mood-related scales in extreme selected sibships. *Hum Mol Genet* 2004; **13**: 2173–2182.
- 23 Neale BM, Sullivan PF, Kendler KS: A genome scan of neuroticism in nicotine dependent smokers. *Am J Med Genet B Neuropsychiatr Genet* 2004; **132B**: 65–69.
- 24 Kuo PH, Neale MC, Riley BP *et al*: A genome-wide linkage analysis for the personality trait neuroticism in the Irish affected sib-pair study of alcohol dependence. *Am J Med Genet B Neuropsychiatr Genet* 2007; **144**: 463–468.
- 25 Willis-Owen SA, Turri MG, Munafo MR *et al*: The serotonin transporter length polymorphism, neuroticism, and depression: a comprehensive assessment of association. *Biol Psychiatry* 2005; **58**: 451–456.
- 26 Munafo MR, Clark T, Flint J: Does measurement instrument moderate the association between the serotonin transporter gene and anxiety-related personality traits? A meta-analysis. *Mol Psychiatry* 2005; **10**: 415–419.
- 27 Middeldorp CM, de Geus EJ, Beem AL *et al*: Family based association analyses between the serotonin transporter gene polymorphism (5-HTTLPR) and neuroticism, anxiety and depression. *Behav Genet* 2007; **37**: 294–301.
- 28 Klein RJ, Zeiss C, Chew EY *et al*: Complement factor H polymorphism in age-related macular degeneration. *Science* 2005; **308**: 385–389.
- 29 Frayling TM, Timpson NJ, Weedon MN *et al*: A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* 2007; **316**: 889–894.
- 30 Grant SF, Thorleifsson G, Reynisdottir I *et al*: Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet* 2006; **38**: 320–323.
- 31 Duerr RH, Taylor KD, Brant SR *et al*: A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science* 2006; **314**: 1461–1463.
- 32 Wellcome Trust Case Control Consortium: Genome-wide association study of 14000 cases of seven common diseases and 3000 shared controls. *Nature* 2007; **447**: 661–678.
- 33 The GAIN Collaborative Research Group: New models of collaboration in genome-wide association studies: the Genetic Association Information Network. *Nat Genet* 2007; **39**: 1045–1051.
- 34 Penninx BW, Beekman AT, Smit JH *et al*: The Netherlands Study of Depression and Anxiety (NESDA): rationales, objectives and methods. *Int J Meth Psychiatr Res*, provisionally accepted. (in press).
- 35 Kessler RC, Barker PR, Colpe LJ *et al*: Screening for serious mental illness in the general population. *Arch Gen Psychiatry* 2003; **60**: 184–189.
- 36 World Health Organization: *Composite International Diagnostic Interview, Core Version 2.1: Interviewer's Manual*. Sydney, Australia: World Health Organization, 1997.
- 37 Bijl RV, van Zessen G, Ravelli A, de Rijk C, Langendoen Y: The Netherlands Mental Health Survey and Incidence Study (NEMESIS): objectives and design. *Soc Psychiatry Psychiatr Epidemiol* 1998; **33**: 581–586.

- 38 Landman-Peeters KM, Hartman CA, van der Pompe G, den Boer JA, Minderaa RB, Ormel J: Gender differences in the relation between social support, problems in parent-offspring communication, and depression and anxiety. *Soc Sci Med* 2005; **60**: 2549–2559.
- 39 Boomsma DI, Beem AL, Berg M van den *et al*: Netherlands twin family study of anxious depression (NETSAD). *Twin Res* 2000; **3**: 323–334.
- 40 Boomsma DI, de Geus EJC, Vink JM *et al*: Netherlands Twin Register: from twins to twin families. *Twin Res Hum Genet* 2006; **9**: 849–857.
- 41 Rush AJ, Gullion CM, Basco MR, Jarrett RB, Trivedi MH: The inventory of depressive symptomatology (IDS): psychometric properties. *Psychol Med* 1996; **26**: 477–486.
- 42 Marks IM, Matthews A: Brief standard self-rating for phobic patients. *Behavior Res Therapy* 1979; **17**: 263–267.
- 43 Beck AT, Epstein N, Brown G, Steer RA: An inventory for measuring clinical anxiety: psychometric properties. *J Consult Clin Psychol* 1988; **56**: 893–897.
- 44 Costa Jr PT, McCrae RR: Domains and facets: hierarchical personality assessment using the revised NEO personality inventory. *J Pers Assess* 1995; **64**: 21–50.
- 45 Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J: An inventory measuring depression. *Arch Gen Psychiatry* 1961; **4**: 53–63.
- 46 Achenbach TM: *The Young Adult Self Report*. Burlington, VT: University of Vermont, Dept of Psychiatry, 1990.
- 47 Wilde GJS: *Neurotische labiliteit gemeten volgens de vragenlijst-methode (The questionnaire method as a means of measuring neurotic instability)*. Amsterdam: Van Rossen, 1970.
- 48 Spielberger CD, Gorsuch RL, Lushene RE: *STAI Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press, 1970.
- 49 Middeldorp CM, Birley AJ, Cath DC *et al*: Familial clustering of major depression and anxiety disorders in Australian and Dutch twins and siblings. *Twin Res Hum Genet* 2005; **8**: 609–615.
- 50 Min JL, Lakenberg N, Bakker-Verweij M *et al*: High microsatellite and SNP genotyping success rates established in a large number of genomic DNA samples extracted from mouth swabs and genotypes. *Twin Res Hum Genet* 2006; **9**: 501–506.
- 51 Van Dijk BA, Boomsma DI, Man AJM de: Blood group chimerism in human multiple births is not rare. *Am J Med Genet* 1996; **61**: 264–268.
- 52 Hoekstra C, Meijer P, Kluff C *et al*: Genetics of dizygotic twinning. A feasibility study for a biobank. *Twin Res* 2004; **7**: 556–563.
- 53 Hattersley AT, McCarthy MI: What makes a good genetic association study? *Lancet* 2005; **366**: 1315–1323.
- 54 Gauderman WJ, Morrison JM: QUANTO 1.1: a computer program for power and sample size calculations for genetic-epidemiology studies, 2006, <http://hydra.usc.edu/gxe>.