A Gene Expression Pattern in Blood for the Early Detection of Alzheimer’s Disease

Birgitte Boonstra Booija,1, Torbjørn Lindahlb, Peter Wetterbergb, Nina Voss Skaaneb, Solve Sæbøc, Guri Fiterc, Phil D. Ryea,2, Lena Iren Kristiansena, Nina Hagena, Marianne Jensena, Ken Bårdsena, Bengt Winbladd, Praveen Sharmaa and Anders Lönneborga

a DiaGenic ASA, Grenseveien, Oslo, Norway
b Memory Clinic, Department of Geriatrics, Medical Division, Ullevål University Hospital, Oslo, Norway
c Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, Ås, Norway
d Department of Neurobiology, Care Sciences and Society, Karolinska Institutet Alzheimer Disease Research Center, Huddinge, Sweden

Accepted 8 september 2010

Abstract: A whole genome screen was performed using oligonucleotide microarray analysis on blood from a large clinical cohort of Alzheimer’s disease (AD) patients and control subjects as clinical sample. Blood samples for total RNA extraction were collected in PAXgene tubes, and gene expression analysis performed on the AB1700 Whole Genome Survey Microarrays. When comparing the gene expression of 94 AD patients and 94 cognitive healthy controls, a Jackknife gene selection based method and Partial Least Square Regression (PLSR) was used to develop a disease classifier algorithm, which gives a test score indicating the presence (positive) or absence (negative) of AD. This algorithm, based on 1239 probes, was validated in an independent test set of 63 subjects comprising 31 AD patients, 25 age-matched cognitively healthy controls, and 7 young controls. This algorithm correctly predicted the class of 55/63 (accuracy 87%), including 26/31 AD samples (sensitivity 84%) and 29/32 controls (specificity 91%). The positive likelihood ratio was 8.9 and the area under the receiver operating characteristic curve (ROC AUC) was 0.94. Furthermore, the algorithm also discriminated AD from Parkinson’s disease in 24/27 patients (accuracy 89%). We have identified and validated a gene expression signature in blood that classifies AD patients and cognitively healthy controls with high accuracy and show that alterations specific for AD can be detected distant from the primary site of the disease.

Keywords: Alzheimer’s disease, blood, biomarker, diagnostic test, microarray, RNA

Supplementary data available online: http://www.j-alz.com/issues/23/vol23-1.html#supplementarydata04

INTRODUCTION

Alzheimer’s disease (AD) affects over 34 million worldwide causing an estimated societal cost of 422 billion USD in 2009 [1]. Taken together with the fact that currently available treatments offer only symptomatic benefit, there is a strong driving force for the development of disease modifying therapeutics. Diagnosing AD at an early stage is crucial both in a clinical trial setting to evaluate drug efficacy and for implementing optimum patient management strategies. In clinical practice AD diagnosis is mainly based on clinical criteria [2], and although research criteria guidelines [3] recommend additional support from biomarkers such as PET imaging or cerebral spinal
fluid biomarkers, high cost and low patient compliance make introducing these diagnostic tools in a wide clinical setting challenging. To make an accurate diagnosis at an early stage of the disease therefore remains difficult and a convenient blood based test would clearly be a valuable asset in the diagnostic workup of patients presenting at clinics with memory complaints.

The potential use of blood-based gene expression profiling in diagnosis of brain disorders has been described [4–12] and further studies also demonstrate a significant degree of covariability in gene expression between brain tissue and peripheral blood cells [13, 14]. Gene expression studies for detection of AD have been described [15, 16], but these studies were performed on relatively small sample sizes and no models for AD prediction were developed. In a pilot study [8], also with few samples and using cDNA clones, a model for AD prediction could be built indicating the potential for the development of a blood-based gene expression test for AD.

The aim of this study was to employ a large sample size and a commercially available platform in order to identify a gene expression signature in peripheral blood with diagnostic value for the detection of AD.

MATERIALS AND METHODS

Subjects

All patients and controls for the study were recruited from different health institutions in the Oslo area of Norway in 2004–2005. Blood sample collection was approved by the Regional Ethical Committee of Norway (Ref. No. 195-04074, dated 22 April 2004) and written informed consent was obtained from each subject included in the study. Patients were recruited at their first visit to the memory clinics at Ullevål university hospital, Aker hospital, Diakonhjemmet hospital and Lovisenberg hospital. Cognitively healthy age and gender matched control subjects were recruited at eight different sites including elderly centers in the area. As additional control subjects young individuals (<30 years of age) were recruited at the Østfold University College, Fredrikstad, and patients diagnosed with Parkinson’s disease (PD) were recruited from the “Norwegian Parkinson’s Association”. Blood sampling was performed either at the clinical site or at Furst Medical Laboratory according to a standard procedure. Demographic and clinical information was obtained for all patients and controls included in the study (for study flow diagram, see Supplementary Figure 1; available online: http://www.j-alz.com/issues/23/vol23-1.html#supplementarydata04). Diagnosis of AD was according to ICD-10. A few subjects with mild cognitive impairment (MCI) according to Petersen criteria [17] were included and clinical follow-up data for these was later obtained. Patients diagnosed with other forms of dementia (including vascular dementia (VaD), frontotemporal dementia (FTD), Lewy bodies dementia (DLB)) were excluded from the study.

Blood sampling

Venous blood samples (2.5 mL) were drawn into PAXgene™ tubes (Becton Dickenson, Qiagen Inc., Valencia, CA) according to the manufacturer’s instructions and inverted 8–10 times. Tubes were incubated at room temperature (18–25°C) overnight prior to freezing at ≤−20°C. Tubes were transported on dry ice to the DiaGenic laboratory in Oslo within 2 weeks and stored at ≤−70°C until processed further. All samples were processed within 6 months of the blood draw.

RNA extraction

Total RNA was extracted from blood samples using PAXgene™ Blood RNA kit (Qiagen Inc., Valencia, CA) according to the manufacturer’s instructions. Total RNA was stored at −70°C or below. Prior to labeling, RNA was concentrated using GeneChip Blood RNA Concentration Kit (Affymetrix, Santa Clara, CA) and assessed for quality and quantity using either the NanoDrop ND-1000 spectrophotometer (NanoDrop, Wilmington, DE) or GeneQuant Pro spectral photometer (Amersham Bioscience, Freiburg, Germany) and the Agilent 2100 BioAnalyzer (Agilent, Santa Clara, CA).

Microarray experimental design

All samples were organized in batches of 12 due to microarray experimental steps. Each batch consisted of two technical replicates used for quality control assessment, one PD patient and nine samples that were balanced in AD samples (positives) and healthy controls (negatives). One of the healthy control samples in each batch was a young control, and the remaining controls were age matched with the AD samples. Alternating batches had either one more positive than negative sample, and a total of 28 batches were randomly selected to be part of the test set and the remaining 21 batches were used for the training set.
Ten MCI subject samples together with two technical replicates were run in an additional batch. RNA samples were shipped on dry ice to the IMGM laboratories (Martinsried, Germany) for microarray analysis. The study was blinded for the operators.

**Microarray procedure**

Microarray analyses were performed using the ABI700 Whole Genome Survey microarrays, which contained 32,878 oligonucleotide probes (60-mers) representing 29,098 genes, and processed according to the manufacturer’s protocol (Applied Biosystems, CA). The NanoAmp RT-IVT Labeling Kit (Applied Biosystems, CA) was used to prepare Digoxigenin (DIG)-labeled cRNA. Briefly, 2 μg total RNA was used to synthesize cDNA using T7-oligo(dT) primers. Following second strand synthesis, the cDNA was purified and subjected to an *in vitro* transcription (IVT) labeling using T7 RNA polymerase and DIG-UTP (Roche Applied Science, Indianapolis). The labeled and fragmented cRNA was hybridized onto the array for 16 h at 55°C. Following hybridization, slides were manually washed and prepared according to the manufacturers’ recommendation with detection using the Anti-Digoxigenin-AP Fab fragments (Roche Applied Science, Indianapolis) and the Chemiluminescence Detection Kit (Applied Biosystems, CA). The hybridization signal was recorded by the ABI700 reader and processed by ABI700 Expression Array System Analyzer software (Applied Biosystems, CA).

**Microarray data quality control**

To ensure high quality, the following quality criteria were applied on the microarray data: 1) homogenous background signal based on visual inspection; 2) the average signal on each microarray to be above 5000; 3) the average signal/noise (S/N) ratio over 32,878 probes to be above three; 4) the median background signal to be below 600 on each microarray; 5) the present call rates to be near 50%; and 6) the probe signals of two technical duplicates to have a correlation coefficient ($r$) of at least 0.98. Additional evaluations were whether 7) technical duplicates have the present call rates at least 90%, and 8) technical duplicates have at least 85% identical probes among the probes called present.

**Microarray data pre-processing**

According to the manufacturers’ instructions, only gene probes with S/N > 3 and binary encoded flag value <8192 were defined as valid measurements. Gene probes that had invalid measurements in more than 10% of all training set arrays were removed entirely from the data set resulting in a set of 11,013 gene probes remaining for further data analysis. Any remaining missing probes were estimated using k-nearest neighbor (k-NN) imputation using $k=10$ [18]. The data from the 11,013 gene probes were log2 transformed and normalized by subtracting each gene probe value for a sample with that sample’s mean value (mean normalization). Finally, the data were adjusted for batch effects, using ANOVA correction for each individual gene probe [19].

**Model building and variable selection**

Partial Least Square Regression (PLSR) with Jackknife variable selection was used for model building [20, 21]. Software package R was used to perform analysis [22]. PLSR is similar to Principal Component Regression (PCR) [23] but while principal components maximize X-variance, the PLSR components maximize X–Y covariance [24]. The gene expression data served as predictors for predicting a dummy-coded response vector. The response vector was given the value −1 or 1 for each sample depending on it being a healthy control or an AD patient, respectively. A new gene expression sample was classified as AD if the predicted value was larger than zero and as healthy control otherwise. Jack-knife feature selection was used to select significant gene probes using a $p$-value of 0.05. A double cross validation approach minimizing data over-fit [25] was then used to estimate the efficacy of the variable selection method and the following performance data was calculated using the clinical diagnosis as reference: accuracy, specificity, sensitivity, AUC from the ROC curve and a positive likelihood ratio.

As PLSR in general handles multiple and potentially highly correlated variables very well, it was assumed that it would be favorable to include more gene probes in the final model rather than to minimize the number of variables before test set validation and thus risk losing biological information. Therefore significant probes were collected from all segments of leave-one-out cross-validation (LOO-CV) analysis. Any gene probe that was found significant at least once was included in a final calibration set from all the train-
ing set samples and a PLSR model was fitted to these data.

**Test set validation of the disease classifier**

When the classification algorithm was finalized, the data from the 7 test set batches were made available for analysis. The test set samples were all from unique donors who had not been included in the gene selection or model building. In addition 28 PD patients were included in the microarray analysis spread throughout the training and test set batches. The PD samples run in the training set batches contributed to the overall signal quality control but did not contribute to the gene selection or model building. For each sample a test score (prediction) value was assigned. The predictions were compared to the clinical diagnosis and the accuracy, specificity, sensitivity, AUC from the ROC curve and a positive likelihood ratio was calculated.

**RESULTS AND DISCUSSION**

**Subjects included**

A total of 320 subjects were recruited to the study (see Supplementary Figure 1 for study flow diagram). These comprised 126 AD patients, 98 age-matched cognitively healthy controls, 28 young controls, 28 PD patients, and 10 subjects with MCI. Thirty (30) patients with other dementias or uncertain diagnosis were excluded from the study.

The gene selection and model building was performed on a balanced set of samples from AD patients \( (n = 94) \) and controls \( (n = 94) \). The controls included 73 age-matched cognitively healthy controls and 21 young controls. These 188 samples comprised the training set and the demographic distribution of these subjects is shown in Table 1. The test set comprised 63 samples from unique donors who had not been included in the gene selection or model building. These included 31 AD, 25 age-matched controls, and 7 young controls, with the demographics of the test set also given in Table 1. In addition, 28 PD patients (Table 1) were included in the microarray analysis and were run spread throughout the training and test set batches.

**Sample and data quality**

The average RNA Integrity Number (RIN) value \( (\pm SD) \) on the RNA samples before labeling was 7.6 \( \pm \) 0.7 covering a range from 5.2 to 9.2. Investigations revealed that the RIN value did not correlate with neither the classification error nor the numerical prediction errors of the classification model; with \( p \)-value 0.06062 for a LOOCV approach and 0.7191 for batch-wise CV. A two-way variance analysis model based on interplay between class and classification and with RIN value as response confirmed these findings.

The labeling of one batch (#23) failed, and was therefore repeated. The PD sample in this batch failed the re-labeling and was therefore excluded from further analysis. Also one AD sample (batch #8) failed the labeling and was excluded from analysis. Four (4) of the 28 batches run did not meet the data quality criteria as they had unacceptable average signal, S/N ratio and median background signal. The hybridizations of these batches were therefore repeated, whereupon all batches passed the quality criteria. Additional quality evaluations showed that three (3) batches performed less well with respect to present call rates in technical duplicates (one (1) batch) and common present probes in technical duplicates (all three (3) batches), but overall, the data Table 1

<table>
<thead>
<tr>
<th>Subjects included</th>
<th>n</th>
<th>Age (years)</th>
<th>MMSE score</th>
<th>Gender distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training set</td>
<td></td>
<td>Average SD</td>
<td>Min Max</td>
<td>Average SD</td>
</tr>
<tr>
<td>Alzheimer’s disease</td>
<td>94*</td>
<td>76.9 8.2 52 95</td>
<td>22.2 4.7 3 30</td>
<td>56 44</td>
</tr>
<tr>
<td>Age-matched controls</td>
<td>73</td>
<td>77.9 7.4 46 90</td>
<td>28.8 1.0 27 30</td>
<td>86 14</td>
</tr>
<tr>
<td>Young controls</td>
<td>21</td>
<td>22.0 2.7 19 29</td>
<td>– – – –</td>
<td>65 35</td>
</tr>
<tr>
<td>Test set</td>
<td></td>
<td>Average SD</td>
<td>Min Max</td>
<td>Average SD</td>
</tr>
<tr>
<td>Alzheimer’s disease</td>
<td>31</td>
<td>78.1 7.5 59 91</td>
<td>22.1 5.0 6 29</td>
<td>65 35</td>
</tr>
<tr>
<td>Age-matched controls</td>
<td>25</td>
<td>78.2 5.6 61 85</td>
<td>28.9 1.0 27 30</td>
<td>80 20</td>
</tr>
<tr>
<td>Young controls</td>
<td>7</td>
<td>23.1 3.0 19 27</td>
<td>– – – –</td>
<td>71 29</td>
</tr>
<tr>
<td>Parkinson’s disease</td>
<td>27*</td>
<td>66.9 8.5 51 81</td>
<td>– – – –</td>
<td>38 62</td>
</tr>
<tr>
<td>MCI</td>
<td>10</td>
<td>77.2 6.6 67 90</td>
<td>26.4* 2.1 23 29</td>
<td>60 40</td>
</tr>
</tbody>
</table>

*One AD and one PD sample failed in the labeling reaction and are not included in the statistics.

*One MCI without MMSE score.
was of good quality. The microarray quality data for all accepted batches can be seen in the Supplementary Figures 2A–G.

Construction and characterization of the disease classifier

Using PLSR with LOO-CV and performing Jackknife feature selection as described on all training set samples (n = 188), a set of 1239 gene probes were found to be significant in at least one of the CV segments with p < 0.05. Six (6) latent components were necessary in the final PLS model to achieve a cross validated minimum in the error rate (see Supplementary Figure 3). This was also the number of components most typically used in the cross validation segments used to select the final set of probes (not shown).

For the training set on which the model was built, there was no significant difference between the AD and age-matched controls with regard to age (p = 0.40). However, there was a significant difference in gender distribution (p = 0.000022), with a bias towards female controls compared to the AD subjects. It was therefore of interest to estimate the effect on the classification error with gender (and age) as a co-variant. This analysis was done using Least Squares-Partial Least Squares (LS-PLS) regression [26], and was found to only have a small positive effect when using only 1 to 4 components to classify the cases and controls (see Supplementary Figure 3). The model giving the optimum results utilizes 6 components, and at this level there was no positive contribution observed by including the gender and/or age as co-variants, and these co-variants were therefore not built into the final model.

The efficacy of the variable selection and classifier estimation method, as tested using a double cross validation approach, is presented in Table 2. Of the 94 AD samples in the training set, 80 were predicted correctly, while 83 of the 94 healthy samples were assigned to the correct class, showing an overall accuracy of 87%.

Test set validation of the disease classifier

The independent test set samples were only revealed when the disease classifier based on the 1239 gene probes had been finalized. The efficacy results in the test set validation are given in Table 2. Of the 63 AD and control samples included in the test set, 55 were correctly classified, with the distribution of the test scores for the patients and controls shown in Fig. 1A. Twenty-six (26) of 31 AD were correctly classified (sensitivity 84%), as were 22/25 age-matched controls and 7/7 young controls giving a specificity of 91% (±10%). The PLR was 8.9 on the test set samples, and the ROC (Fig. 1B) showed an AUC of 0.94.

Table 2

<table>
<thead>
<tr>
<th>Efficacy</th>
<th>Training Set</th>
<th>Test set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>163/188</td>
<td>87 (5%)</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>80/94</td>
<td>85 (7%)</td>
</tr>
<tr>
<td>Specificity</td>
<td>83/94</td>
<td>88 (6%)</td>
</tr>
<tr>
<td>Age-matched controls</td>
<td>63/73</td>
<td>22/25</td>
</tr>
<tr>
<td>Young controls</td>
<td>20/21</td>
<td>7/7</td>
</tr>
<tr>
<td>Parkinson’s Disease*</td>
<td>24/27</td>
<td>89 (12%)</td>
</tr>
<tr>
<td>MCI†</td>
<td>7/10</td>
<td>70 (34%)</td>
</tr>
<tr>
<td>PLR</td>
<td>7.3</td>
<td>8.9</td>
</tr>
<tr>
<td>AUC</td>
<td>0.93</td>
<td>0.94</td>
</tr>
</tbody>
</table>

*not included in overall efficacy calculations.
PLR = positive likelihood ratio. AUC = area under the ROC curve.

The distribution of the test scores for the patients and controls in the training set is given in Fig. 1A. When plotting the sensitivity versus 1-specificity in a receiver operating characteristics (ROC) curve (Fig. 1B), there was a good separation of the two groups with an area under curve (AUC) of 0.93.

Approximately 50% of all AD patients included in the study were on either acetylcholinesterase inhibitors (donepezil, galantamine, or rivastigmine) or the NMDA antagonist memantine. There was no apparent effect on the use of these drugs on the classification accuracy, with 31/37 (84%) of AD patients on donepezil correctly classified. The corresponding figures for the other AD drugs were 16/18 (89%) on galantamine, 4/5 (80%) on rivastigmine and 2/3 (67%) on memantine being correctly classified.

Neither did the most common co-morbidities interfere with the classification. For the AD patients and age-matched controls included in the study, the following were corre-
Fig. 1. Efficacy of the identified disease classifier. A) Classification scores (AD > 0 > controls). Healthy control subjects are indicated as C-y (young controls) and C-a (age-matched controls). Alzheimer’s disease patients as AD, Parkinson’s disease patients as PD, and subjects with mild cognitive impairment as MCI. Training set (in black): among the 94 AD patients in the training set, 80 were correctly classified, while 83 of the 94 healthy samples were assigned to the correct class. Validation samples (in red): among the 31 AD patients, 26 were correctly classified, while 29 of the 32 healthy samples were assigned to the correct class. Twenty-four of 27 PD patients were correctly classified as healthy, whereas out of 10 MCI subjects the disease classifier predicted the final diagnosis correctly for 6/8 that were available for follow-up. A score > 0 classifies a subject as having AD while a score < 0 classifies a subject as non-AD. Training set samples are shown in black and validation samples in red.

B) ROC curves from training set (in black) and test set (in red). In black: prediction of the 188 training set samples based on double cross-validation results on the 1239 probe list. Training set gave a classification accuracy of 87% and an area under curve (AUC) of 0.93 reflecting a good separation of the two groups. In red: ROC curve based on 63 test set samples, where the classification accuracy is 87% and the AUC 0.94.
directly classified: 9/11 (82%) with cancer, 58/66 (88%) with hypertension, 9/11 (82%) with diabetes, 10/10 (100%) with rheumatoid arthritis, 18/21 (86%) with heart failure, 22/25 (88%) with coronary disease, 12/14 (86%) with previous strokes, 18/19 (95%) with depression, 3/4 (75%) with known chronic alcoholism, 22/26 (85%) with allergies, 17/20 (85%) with acute illness and 13/16 (81%) with elevated homocysteine levels (>15 μM). Although the subject numbers for some of the co-morbidities are small, the results are comparable to the overall performance and exemplify the robustness of the identified gene expression signature for detecting AD.

**Performance of the disease classifier towards other neurological disease**

Of the 27 PD patients included in the study and where acceptable microarray data quality was obtained, 24 (89%) were correctly predicted as non-AD by the disease classifier. This is in agreement with the overall specificity as obtained in the test set validation (Table 2), and indicates that the constructed classification algorithm is specific for picking up the changes in gene expression occurring in blood of AD patients.

**Performance of the disease classifier in predicting MCI to AD converters**

Of the few (10) MCI subjects included in the study, one patient (classified as AD with the disease classifier) deceased prior to clarification of the diagnosis (although it was mentioned that cognitive status was a complicating factor). For another subject (classified as cognitively healthy with the disease classifier) the disease status is still MCI after 4.25 years, but the subject has had no clinical progression in this period. Of the remaining 8 MCI cases included, 5 of these converted to AD, 2 to other dementias (1 semantic and 1 vascular) and one reverted to cognitive healthy status with an average follow-up time of 24 months. When tested as MCI subjects, the disease classifier correctly predicted the outcome of 6/8 of these, including 4/5 AD conversions and 2/2 other dementias were predicted as non-AD. Although the numbers are low this indicates the potential predictive power of the identified gene expression signature.

**Classification score according to AD grade**

Subjects included in the study were retrospectively graded according to disease severity using available clinical data and primarily MMSE grade into the following categories: Healthy controls (0), Very mild AD (1), Mild AD (2), Moderate AD (3) and Severe AD (4). In addition, the 10 MCI patients were allocated grade 0.5. The results are presented in Fig. 2. It can be seen that the disease classifier is able to detect the different AD stages with similar level of accuracy. The MCI subjects fall in between the healthy controls and grade 1 reflecting the mixture of AD and healthy/other status these may develop into. This indicates that when the disease pathology first is present, a systemic response to the disease sets in which remains stable over the course of the disease. This implies that the identified gene expression signature could be of value since diagnosing AD at an early stage is clinically the most challenging. More subjects are needed to determine if the gene expression signature also can be of diagnostic value at the MCI stage.

**Biological significance of the 1239 informative gene probes identified**

In this study 1239 informative gene probes were identified. The panther website (http://www.pantherdb.org/) was used to search for the biological significance of the genes represented by these 1239 probes. Compared with what would be expected if there was a random distribution of biological functions among the genes identified, five processes were significantly overrepresented ($p < 0.05$) in the signature (Fig. 3A): protein folding, chromatin packaging, and remodeling, reverse transcription, and stress response. There were also certain biological processes significantly under-represented ($p < 0.05$), including skeletal development, cell adhesion-mediated signaling, developmental processes, steroid metabolism, and interestingly, also neuronal activities. Genes that encode histones, various chaperones, transferases, annexins, TGF-beta receptors, and DNA-binding proteins are all significantly overrepresented (Fig. 3B). Approximately 70% of the informative probes hybridize to genes with a known molecular function or an ascribed biological process. The remaining approximately 30% have unknown molecular function and the biological process in which they are involved is not yet known.

Of the 1239 identified gene probes, at the time of writing only one represents a gene which is among the top 20 list of genes identified from genome wide association studies described in the AlzGene database [27]. This is sortilin-related receptor SORL1, a neuronal apolipoprotein E receptor which previously is shown to have significantly reduced expression both in brain tissue and in lymphocytes of AD patients [28]. The genes in the AlzGene list are ranked in accordance...
with their link to an increased risk of developing AD and such an increased genetic risk is not the same as an altered expression of the gene. First of all the level of gene expression does not necessarily correlate well with the level of the corresponding protein. In addition, a genetic risk factor, usually a minor allele difference, can exert an effect on most of the steps from transcription of the gene to the functionality of the final gene product, the protein. Any alteration in any of these steps may increase the risk of developing the disease.

The statistical approach presented in this study identifies gene products with diagnostic value based on gene expression rather than genetic risk factors for developing the disease. This is clearly shown when only one of the genes in the top 20 AlzGene list were found among the identified genes despite the fact that most of the top ranking AlzGene genes were represented on the AB1700 microarray platform.

Using a cut-off of 1.2 fold-change, which is frequently applied in gene expression studies, 99 gene probes were found to be upregulated in AD versus healthy controls. Similarly, 131 gene probes were found to be down-regulated in AD versus healthy controls, using a cut-off of 0.8 fold-change. The number of genes found to be differentially expressed (using cut-off of >1.2 and <0.8 fold change) is typically what is found in other gene expression studies studying response to disease in blood [29]. In PD 128 gene probes were upregulated and 127 down-regulated compared to healthy controls. There was an overlap of 59 upregulated and 19 down-regulated gene probes in AD and PD.

We have, using a large sample size, identified a gene expression signature consisting of 1239 gene probes that can classify AD patients from cognitively healthy controls with a high accuracy ($87 \pm 8\%$). The signature further shows high specificity towards another neurodegenerative disease, namely PD. The efficacy results reported here are in line with the consensus criteria for AD biomarkers which states that a biomarker for AD should have a sensitivity and specificity of $>80\%$ [30]. The fact that there is no apparent interference from common co-morbidities as well as existing AD medications, indicates the robustness of the identified gene expression signature in picking up the AD specific gene expression pattern.

A subset of the informative probes identified in this study was transferred to a real-time PCR platform in order to validate that the genes maintained their informative value. In such an initial gene validation study 52
Fig. 3. Biological processes (A) and molecular functions (B) over- and under-represented in the gene expression signature.
of the same patient samples (24 AD, 21 age-matched, and 7 young controls) were run on micro fluidic cards comprising a total of 178 commercially accessible gene assays. Hundred and fifty-five of these assays retained their informative value with a PLS model build on these able to distinguish AD and controls with an accuracy of 85% (±69%) as determined with a leave-one out cross validation (data not shown).

The gene expression signature identified in this whole-genome screen study is being further refined and developed into a clinical test for use on a real-time PCR platform and will offer clinicians a blood based tool to supplement the existing diagnostic workup for early diagnosis of AD.

ACKNOWLEDGMENTS

The following hospitals contributed with clinical samples: Ullevål University hospital (Oslo), Rikshospitalet (Oslo), Lovisenberg Diakonale Hospital (Oslo), Diakonhjemmet Hospital (Oslo) and Aker Hospital (Oslo). We gratefully acknowledge the contribution of clinical expertise from Knut Engedal (Oslo). We thank Fürst Medisinske Laboratorium (Oslo) for blood samples: Ullevål University hospital (Oslo), Rikshospitalet (Oslo), Lovisenberg Diakonale Hospital (Oslo), Diakonhjemmet Hospital (Oslo), and Aker Hospital (Oslo). We gratefully acknowledge the contribution of clinical expertise from Knut Engedal (Oslo). We thank Fürst Medisinske Laboratorium (Oslo) for blood sampling and IMGM laboratories (Martinsried, Germany) for performing the microarray work. Funding was partly provided by The Norwegian Research Council Functional Genomics program. Authors’ disclosures available online (http://www.j-alz.com/disclosures/view.php?id=614).

REFERENCES


