Diagnosis of Alzheimer’s Disease (AD) involves clinical interviews with the patient and preferably also a close relative to assess for functional decline and change in behavior. Tests are performed to determine any cognitive impairment (MMSE, clock drawing, OLT). These tests are often complemented with structural imaging (MR, CT), and sometimes also with functional imaging (PET, SPECT) and measurements of neurophysiological function (EEG, qEEG). However, with all these tests it is still difficult to make an accurate diagnosis especially at an early stage of the disease. An accurate, convenient and efficient test to detect AD, especially a test based on peripheral blood as the clinical sample, is much desired.

The potential use of blood-based gene expression profiling in diagnosis of brain disorders has been described [1-4] and further studies also demonstrate a significant degree of covariability in gene expression between brain tissue and peripheral blood cells [5]. We have previously presented results from a pilot study suggesting that a blood-based gene expression base test potentially can be developed for AD [Abstract number: 162960, 2005 12th IPA, Stockholm, Sweden]. We have more recently confirmed our initial findings in a large-scale study using AB Human Whole Genome microarrays [Abstract number: 03-01-01, 2006, 10th ICAD Conference, Madrid, Spain].

A novel blood-based gene expression signature was identified that could discriminate AD patients and age-matched controls with high accuracy, sensitivity, and specificity. The signature also accurately distinguished AD patients from those having another neuro-degenerative disease, Parkinson disease. Results from an initial validation of a subset of identified informative gene probes on the quantitative real-time PCR (qRT-PCR) based TaqMan® Low Density Arrays® (LDA, Applied Biosystems) showed that high diagnostic accuracy was retained on TaqMan® LDA and the platform potentially can be used for diagnostic purposes.

We have now used the informative gene probes identified in previous studies to design and develop customized arrays on two different platforms that employ alternate technologies for gene expression profiling. One is based on microarray technology (CodeLink™ Customized MultiAssay BioArrays), while the other relies on qRT-PCR technology (TaqMan® Low density arrays). A new cohort of AD and non-AD samples were employed in the study presented here.

Materials and methods

Whole blood was collected from 119 individuals in PAXgene™ Blood RNA tubes at different hospitals and institutions in Norway. These included 53 patients diagnosed with AD (based on the ICD-10 criteria for dementia syndrome), 58 age-matched controls and 8 control samples from young individuals to rule out any possibility of AD, see Table 1. The same samples were used both in the TaqMan® LDA and in the CodeLink® study.

Total RNA was extracted from blood samples using PAXgene™ Blood RNA kit and quality assessed by NanoDrop spectrophotometer and Agilent 2100 Bioanalyzer. cDNA was prepared using the high-capacity cDNA archive kit from Applied Biosystems and gene expression analysis was conducted on ABI Prism 7900HT Fast System using TaqMan® LDA containing 384 selected arrays, including endogenous controls. For customized microarrays, 30-mer probes were designed for the 1498 gene probes identified as informative in previous studies and immobilized on CodeLink™ Bioarrays with 16 Multissay format. In addition, different positive and negative controls were also immobilized.

The generated expression data from the two customized arrays were normalized to take account of the differences in the probe intensities resulting from conditions such as differences in hybridization and labeling steps. In case of TaqMan, assays with either missing values or having an average Ct > 30 (160 assays in total) were removed prior to data analysis. Partial Least Square Regression (PLSR) and Leave one out cross-validation (LOOCV) were used for model building and to estimate the prediction accuracy. A novel approach combining double and triple cross-validation (CV) routine was used to select a smaller number of genes, which could fit within a 96-assay format and the prediction accuracy of these genes was estimated. However, during gene selection, data from only 112 samples were used for modeling. The samples that were excluded had a high influence on the model.

For CodeLink, experiments were run in batches and the generated data were batch-adjusted prior to data analysis. PLSR was used for model building and prediction accuracy determined by Leave one out cross-validation. Among the 119 samples that were included in the study, only 114 samples were used for model building and estimation of prediction accuracy. The samples that were excluded highly influenced the models and were regarded as outliers.

Table 1. Demographic information of patient and control samples. AD and age-matched control samples were the same for TaqMan and CodeLink. 2/8 young control samples were different; Sample info for TaqMan is shown.

<table>
<thead>
<tr>
<th></th>
<th>Age (years)</th>
<th>MMSE score</th>
<th>Gender distribution</th>
<th>Female (%)</th>
<th>Male (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alzheimer’s disease (N=53)</td>
<td>77,4</td>
<td>7.5</td>
<td>55</td>
<td>88</td>
<td>21,1</td>
</tr>
<tr>
<td>Age- matched controls (N=58)</td>
<td>75,8</td>
<td>7.6</td>
<td>59</td>
<td>93</td>
<td>29,2</td>
</tr>
<tr>
<td>Young controls (N=8)</td>
<td>22,6</td>
<td>2.4</td>
<td>20</td>
<td>28</td>
<td>-</td>
</tr>
</tbody>
</table>

Results and Discussion

The prediction accuracy of the developed signature using all the 214 assays on TaqMan is shown in Table 2. The gene expression signature developed correctly predicted the class of 96/119 samples (accuracy 81.7%), including 39/53 AD samples (sensitivity, 74%) and 57/66 Non-Alzheimer controls (specificity, 86%). From the 214 data, we identified a limited number of assays that fit within a 96-assay format including assays used for normalization and quality control purposes and developed an Alzheimer’s specific gene expression signature. The estimated accuracy for this signature was not significantly different than for the 214 assays. The ROC curve and AUC of this signature is presented in Figure 1.

In case of CodeLink customized arrays, the Alzheimer’s specific gene expression signature was developed using complete pre-processed data. As shown in Table 2, the developed gene expression signature correctly predicted the class of 97 samples (accuracy, 85%), including 42 AD samples (sensitivity, 82%) and 49 non-Alzheimer’s controls (specificity, 87%).

The prediction results reported here are comparable to those previously presented using AB1700 Human Whole Genome Microarrays. The comparative ROC and AUC obtained in these studies is presented in Figure 1. The AUC in these studies ranged between 0.89 to 0.93 thus demonstrating improved diagnostic value compared to existing clinical diagnostic practice.

The biological processes in which the genes represented within the 96-assay format for TaqMan® LDA are involved, are shown in Figure 2. The largest categories represent protein metabolism and modification, nucleoside, nucleotide and nucleic acid metabolism and as yet unclassified biological processes. These were also the largest categories represented in our previously reported AB1700 WGA study before the enrichment of the most informative genes were made.

Table 2. Prediction results from AB1700 Human whole genome array, TaqMan® LDA and CodeLink™ BioArray. LR+: positive likelihood ratio. (%): 95% Confidence interval.

<table>
<thead>
<tr>
<th>Study</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
<th>LR+</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB1700 Whole Genome Array</td>
<td>85%</td>
<td>88%</td>
<td>87% (5%)</td>
<td>7.3</td>
</tr>
<tr>
<td>TaqMan® Low Density 214-assay Array</td>
<td>74%</td>
<td>86%</td>
<td>81% (7%)</td>
<td>5.3</td>
</tr>
<tr>
<td>CodeLink™ BioArray</td>
<td>82%</td>
<td>87%</td>
<td>85% (7%)</td>
<td>6.5</td>
</tr>
</tbody>
</table>

Conclusions

- Results presented show that the informative genes identified using whole genome microarrays can be used to generate a highly accurate Alzheimer’s specific gene expression signature in peripheral blood on different platforms.
- An Alzheimer’s specific gene expression signature has been developed for a 96-assay format on TaqMan® LDA and on CodeLink® BioArrays with accuracies of diagnostic value in a clinical setting.
- By repeated multi-centre studies we have documented that we can develop a robust, platform independent gene expression signature for Alzheimer’s disease thus enabling a convenient diagnostic tool to be developed.
- The main biological processes involved in the Alzheimer’s specific gene expression signature in blood are protein metabolism and modification, nucleoside, nucleotide and nucleic acid metabolism and as yet unclassified processes.

References


Competing interests

All authors of DiaGenic have competing interests. Authors from other institution have no competing interest.