ACCURATE AND EARLY DETECTION OF ALZHEIMER’S DISEASE USING A GENE EXPRESSION SIGNATURE IN BLOOD

Anders Lönneborg1, Birgitte Booij1, Marianne Jensen1, Ken Bårdsen1, Nina Hagen1, Lena Kristiansen1, Torbjørn Lindahl1, Solve Sæbø2, Guri Feten2, Bengt Winblad3, Khalid Iqbal4, Praveen Sharma1

1) DiaGenic ASA, Grenseveien 92, NO-0661 Oslo, Norway; 2) Dept. of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, P.O.Box 5003, NO-1432 As, Norway; 3) Dept. of Neurobiology, Div. for Aging Research Center, Karolinska Institutet, Novum, Pico 5, SE-141 57 Huddinge, Sweden; 4) New York State Inst. for Basic Res. in Devel. Disabilities, 1050 Forest Hill Road, Staten Island, NY 10314, U.S.A.

Introduction
With potential disease modifying drugs for Alzheimer’s soon to come we need accurate and easy-to-use early identification techniques for Alzheimer’s so that we can initiate treatment earlier. Today the procedure for diagnosis of Alzheimer’s Disease (AD) involves clinical interviews with the patient and very often also a close relative, and tests to determine any cognitive impairment often completed with structural imaging and sometimes also with functional imaging and measurements of neurophysiological function. With all these interviews and 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 an easily available clinical sample, is much desired.

The potential use of blood-based gene expression profiling in diagnosis of brain disorders has been described1-4, and further studies also demonstrate a significant degree of covariability in gene expression between brain tissue and peripheral blood5. We have previously presented results from a pilot study suggesting that a blood-based gene expression based test can be developed for AD6. We have more recently confirmed our initial findings in a large-scale study using AB Human Whole Genome microarrays. 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 Parkinson disease, another neuro-degenerative disease. In an initial validation with a subset of identified informative gene probes on the quantitative real-time PCR (QRT-PCR) based TaqMan Low Density Array (LDA), Applied Biosystems) showed that high diagnostic accuracy was retained 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 customised arrays on two different platforms that employ alternate technologies for gene expression profiling. One is based on microarray technology (CodeLink™ Customized Multisassay Bioarrays), while the other relies on QRT-PCR technology (TaqMan® LDA) (figure 1). A new cohort of AD and non-AD samples were collected and employed in the studies presented herein.

Materials and methods
Whole blood was collected from 119 individuals in PAXgene™ Blood RNA tubes at different hospitals in Norway. In total, 53 patients diagnosed with AD (based on the NINCDS-ADRDA criteria for dementia syndromes). 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 (figure 2) and quality assessed by NanoDrop spectrophotometer and Agilent 2100 Bioanalyzer.

Table 1. Demographic information of patient and control samples. AD and age-matched control samples were the same for TaqMan and CodeLink. 28 young control samples were different. Sample info for Taqman is shown. Study | Sensitivity | Specificity | Accuracy | LR+ | 95% confidence interval
--- | --- | --- | --- | --- | ---
AB1700 Whole Genome Array | 85% | 88% | 87% (95%) | 7.3 | 0.7 | 5.0
TaqMan® LDA | 74% | 84% | 79% (95%) | 4.6 | 0.7 | 3.3
CodeLink™ BioArray | 82% | 87% | 85% (95%) | 6.5 | 0.7 | 3.8

The biological processes in which the genes represented within the 96-assay format TaqMan® LDA are involved, are shown in Figure 4. The largest categories represent protein metabolism and modification, nucleoside, nucleotide and nucleic acid metabolism. These were also the largest categories represented in our previous reported AB1700 WGA study before the enrichment of the most informative genes were made.

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® LDA, assays with either missing values or having an average CT > 30 (160 assays in total) were removed prior to data analysis. Partial Least Squares Regression (PLSR) and Leave one out cross-validation (LOOCV) were used for model building and to estimate the prediction accuracy. A novel approach based on Jackkniifing and combining sets of significant genes was used to select a smaller number of genes which could fit within a 96-assay format. This approach was repeated for each segment of the cross validation loop. 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 estimation of prediction accuracy. The samples that were excluded highly influenced the models and were regarded as outliers.

Results and Discussion
We have developed a signature within a 96-assay format on TaqMan® LDA. As is shown in Table 2 the gene expression signature developed correctly predicted the class of 89/112 samples (accuracy 79%), including 36/49 AD samples (sensitivity, 74%) and 53/63 Non-Alzheimer controls (specificity, 84%). The ROC curve and AUC of this signature is presented in Figure 3.

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 55 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 2. The AUC in these studies ranged between 0.89 to 0.93 thus demonstrating improved diagnostic value compared to existing clinical diagnostic practices.

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

References
7. Silence RNA interference technologies for cancer diagnosis and therapeutics.

Figure 3. ROC curves for AB1700 WGA, TaqMan® LDA and Codelink™ BioArray data. The respective AUC’s are shown.

Figure 4. Biological processes involved in the Alzheimer’s specific gene expression signature for a 96-assay format on TaqMan® LDA.