Strategies for the Identification of Novel Brain Specific Genes Affected in Alzheimer Disease


ABSTRACT: The pathological changes that occur in Alzheimer disease (AD) brain lead to a large loss of various classes of neurons and the production of novel proteinaceous elements such as neuritic plaques and neurofibrillary tangles. For the neuronal loss to occur and these elements to arise, there must be a disturbance in the expression or regulation of genes that code for proteins required for normal cell maintenance, or perhaps even for the expression of genes unique to AD. We describe the construction of a cDNA library from the human substantia innominata and strategies for isolating genes that are expressed differentially between brain regions and which may be affected by AD. Some of the results obtained using these strategies and a preliminary description of a novel brain specific mRNA of 15.5kb, whose expression is increased in AD affected temporal cortex, are presented.

RÉSUMÉ: Génétique moléculaire de la maladie d’Alzheimer familiale Les changements pathologiques qui surviennent dans le cerveau de patients atteints de la maladie d’Alzheimer (MA) provoquent une perte considérable de neurones de différentes classes et une production d’éléments protéiques nouveaux, dont les plaques nécritiques et les amas neurofibrillaires. Pour que cette perte neuronale puisse survenir et que ces éléments apparaissent, il doit exister une perturbation de l’expression ou de la régulation des gènes qui codent pour les protéines nécessaires au maintien de la cellule normale, ou peut-être même à l’expression de gènes uniques à la MA. Nous décrivons la construction d’une librairie d’ADN complémentaire à partir de la substance innominée humaine et les stratégies pour isoler les gènes qui sont exprimés de façon différentielle entre les régions du cerveau, gènes qui pourraient être atteints dans la MA. Nous présentons quelques uns des résultats obtenus en utilisant ces stratégies ainsi qu’une description préliminaire d’un nouveau ARN messager de 15.5 kb, spécifique au cerveau, dont l’expression est augmentée dans le cortex temporal de patients atteints de la MA.

Alzheimer disease (AD) is a neurodegenerative disease of the elderly that is characterized by a severe loss of cognitive function. To date, little is known about the cause of the neurodegeneration that occurs in AD. Hallmark pathological features of AD are the loss of neurons together with the formation of neurofibrillary tangles (NFT) and amyloid containing neuritic plaques (NP). The application of basic scientific methodologies to AD research has concentrated on elucidating the nature and etiology of these abnormal proteinaceous structures. The plaques and tangles are most abundant in the temporal cortex, parietal cortex and hippocampus and occur to a lesser extent in the frontal cortex and basal forebrain. There is now substantial evidence indicating selective vulnerability of certain neuronal groups. For example, the magnocellular cholinergic neurons of the basal forebrain are particularly vulnerable in AD. This is reflected by a decrease in the specific activity of neocortical choline acetyltransferase (ChAT), the enzyme responsible for the synthesis of acetylcholine. Although a number of other neuronal systems are selectively affected, e.g. noradrenergic neurons of the locus coeruleus, serotonergic neurons of the raphe complex, neurons of the central and medial nuclei of the amygdala, hippocampal pyramidal neurons, and neocortical glutamatergic pyramidal neurons, neocortical ChAT levels remain one of the most reliable diagnostic criteria for AD. The molecular events that lead to the deposition of amyloid and the formation of neuritic plaques, the degeneration of neurons into insoluble paired helical filament containing tangles, and the selective loss of certain neuronal populations are the main identified areas of AD research. It can be assumed that these molecular events are accompanied by an alteration in the normal expression of the genes that are necessary for maintaining normal cell function at one of the levels of gene regulation, i.e. RNA transcription, RNA stability, RNA translation or post-translational modifications of proteins. We, and other research groups, are using this
assumption to apply recombinant DNA methodology to the investigation of gene expression in AD affected brains. There are two methods of approach: a) the analysis of levels of expression of identified genes, and b) the identification and analysis of unidentified genes which show differential expression between normal and AD affected brains. In this paper, we provide some review information on these approaches and some of our experimental findings using the second approach.

Examination of the Expression of Identified Genes in AD Brains

Measurements of the expression of identified genes in various regions of AD affected brains have been the object of a number of investigations, in attempts to identify key genetic alterations in AD. The majority of these studies have been with various probes derived from the sequence of the amyloid precursor protein (APP) gene. It has been suggested that abnormal expression of this gene may be the cause of β-amylloid protein deposition in senile plaques. Several groups have used synthetic oligonucleotide probes, designed from the amino acid sequence obtained by Glenner and Wong from extracted vascular amyloid, to isolate cDNA copies of the gene for the amyloid precursor protein. It was found that this gene was widely expressed in a number of tissues besides brain and was highly homologous to genes present in species that do not develop amyloid containing plaques with age. It was subsequently shown that alternate forms of the originally identified APP gene existed that contained 56 and 75 extra amino acids. The extra 56 amino acids showed about 50% homology to the Kunitz type serine protease inhibitor. The third form has an extra 19 amino acids added to this domain. The three forms of APP mRNA are referred to as APP695, APP751 and APP770. The APP695 form is more highly expressed in nervous tissue, particularly neurons, while the insert containing forms occur widely in neural and non-neural tissues. With the complete cDNA probe, Northern blot hybridization analyses showed a decrease in the expression of APP in AD affected temporal or frontal cortex compared to tissue from non-demented cases. With probes specific for the APP695, APP751 and APP770 forms, Northern blot hybridizations of RNAs from cortex of AD and normal cases revealed an increase in AD in the APP770 transcript with no change in the expression of APP695 and APP751. However, in another study, a 65% decrease in the expression of the APP695 transcript was found in AD cortex compared to cortex from neurologically normal cases with no change being found in the levels of the APP751 transcript. Quantitative in situ hybridization studies have shown an apparent increase in APP gene expression in magnocellular neurons of the nucleus basalis of Meynert in AD due to increased expression of the APP695 transcript. An alteration in the expression of the APP gene in the paracubiculum neurons of the hippocampus in AD has also been observed. The relevance of APP gene expression to the pathology of AD is still unclear, as there is no clear difference in the expression in brain regions or neuronal cell types that are not affected in AD (e.g. hypothalamus) compared to those that are affected in AD (e.g. temporal cortex). Furthermore, DNA sequence analysis has shown that there is no difference between the APP gene of neurologically normal and AD affected individuals. This would seem to indicate that the AD brain amyloid protein is not a mutated form of the normal one, and that the deposition in plaques may occur due to abnormal processing of the normal APP. The initial interest in the APP gene was heightened by its localization to chromosome 21. The gene was thought at first to be in the vicinity of a DNA probe that identified a genetic defect for some families with AD. Subsequent work has shown that recombination events occur between the FAD locus detected by this probe and the APP gene sufficient to establish that they are separated by more than eight million base pairs. There is no evidence of gene duplication of the APP gene in the majority of AD affected individuals.

The genes of some of the proteins found to be associated with neurofibrillary tangles (NFT) have been identified (tau, microtubule associated protein 2 (MAP 2), neurofilament complex components, ubiquitin). The expression of the microtubule associated protein tau gene was shown not to be different in AD affected tissue. Significant decreases in the expression of one of the genes of the neurofilament complex, the neurofilament L gene, were observed in AD. It is unclear whether this decrease results from overall loss of neurons expressing this gene or is occurring as an earlier event that may be contributing to the degeneration of affected neurons. Increases in levels of glial fibrillary acidic protein (GFAP) mRNA, a specific marker for astrocytes, have also been observed in AD affected brains, but this is probably a result of reactive astrogliosis.

Indications that significant data may be obtained on pathological mechanisms of AD by the study of neuronal growth and regulatory factors have come from studies of expression of genes of nerve growth factor (NGF) and its receptor. It was shown that there was no significant difference in levels of NGF mRNA expression in cortex of normal and AD cases. Nerve growth factor appears to have a role in the development and survival of central cholinergic neurons. Treatment with this agent can prevent the degeneration of lesioned cholinergic neurons in rats. NGF appears to function as a retrograde trophic messenger for basal forebrain cholinergic neurons. Similarly, it was shown that there was no difference in the levels of expression of NGF receptor mRNA in the human basal forebrain between normal and AD cases. As the loss of magnocellular neurons from this brain region is well documented, it would appear that the surviving neurons are expressing increased levels of NGF receptor mRNA. The degeneration of these neurons may be occurring as a result of the interruption of the retrograde supply of NGF.

Analysis of the Expression of Unidentified Genes in AD

Although there is obviously great value in analysing the expression of key marker genes whose protein products have been characterized, these genes represent only a small fraction of those expressed specifically in human brain. The genes studied as possibly being related to neuronal degeneration have been those of high abundance neuronal structural elements. The techniques of differential/subtractive hybridization offer the possibility of directly identifying genes whose expression is increased or decreased in AD, even though the proteins these genes are coding for remain unknown. This methodology has been widely applied to the study of gene expression in developing tissues, especially the brain, and in characterizing genes expressed in mutant strains of animals. These techniques have allowed the identification of anonymous genes whose expres-
sion is increased or decreased between two populations of RNA. However, several factors must be considered in applying these techniques to neurodegenerative diseases. The AD brain has a large neuronal loss with associated hypertrophy of astrocytes. The use of differential/subtractive screening techniques to compare genes expressed in AD affected brain tissue with those expressed in normal brain will primarily identify abundant astrocyte genes. A considerable amount of secondary screening of differentially expressed genes, along with careful interpretation, may be required to identify those genes whose induction is not secondary to the degenerative process. For example, subtractive hybridization of scrapie infected brain cDNA from normal hamster brain identified three such genes: GFAP, metallothionein II and the B chain of α-crystallin, all of which are considered to be nonspecific markers of neurodegeneration. However, the approach of differential hybridization has been successfully applied by May and colleagues to isolate a neuronal gene whose expression is increased in AD affected hippocampus.

Analysis of Gene Expression in Human Substantia Innominata

The aim of the project we have undertaken is to investigate the molecular biology of the neurons present within the substantia innominata (SI), an area highly enriched in magnocellular cholinergic neurons, and the effect of AD on genes expressed in this region. For these studies using postmortem human brain tissue, total cellular RNA was prepared by the method of Chirgwin. This method has been used to prepare relatively intact RNA from frozen tissue, as assessed by its limited extent of degradation, and its ability to be an efficient substrate for reverse transcription and in vitro translation. The extent of degradation in any sample is judged by the relative intensities of the 28S and 18S ribosomal bands, when samples are separated by formaldehyde agarose gel electrophoresis and stained with ethidium bromide, and by the intactness of hybridization signals on Northern blots using probes for the APP and glutamic acid decarboxylase genes. These probes detect major mRNA bands of 3.5 and 3.7 kb, respectively. In our experience, the intactness of these signals gives a more reliable estimation of degradation than does the use of an actin probe (which detects an mRNA of 1.8 kb). Relative yields of total cellular RNA from hippocampus, temporal cortex and SI of AD and neurologically normal cases are shown in Figure 1. There was no significant difference between yields of RNA from temporal cortex and SI of AD and normal cases. However, a significant decrease in RNA yield from hippocampus of AD was observed.

cDNA Library Construction and Screening Methodology

Using mRNA from the SI of two neurologically normal cases (ages 79 and 80 years), a cDNA library was constructed. Poly A+ RNA was prepared by batch absorption to and elution from oligo dT cellulose. Double stranded cDNA was prepared by the method of Gubler and Hoffman. After the addition of EcoRI adaptors, the cDNA was ligated to EcoRI cleaved plasmid vector pT7T318U. This was used to transform E.Coli (strain DH5α) to ampicillin resistance using the technique of electroporation. A library of about one million independent clones was produced. Fifty thousand primary clones were plated out onto nitrocellulose membranes for screening, while the rest were amplified and frozen.

As the amount of RNA available from the human SI is limited (1-2 µg poly A+/case), large scale differential hybridization or subtractive hybridization procedures were not feasible. In order to identify regionally expressed genes, a negative screening protocol was adopted that would allow the removal of genes expressed in a brain region that did not contain the cells of interest. This procedure was similar to that used to identify cholinergic specific genes of the Torpedo ray. The library was first screened by colony hybridization using a cDNA probe made from cerebellar RNA. The rationale is that the cerebellum is only marginally affected by the degenerative changes of AD, and furthermore contains no cell bodies of cholinergic neurons. In addition, it was assumed that this procedure would detect (and thus exclude) any abundant astrocytic genes, whose expression is likely to be increased during gliotic changes in AD (eg GFAP). This screening procedure (summarized in Figure 2) should allow the identification of neuronal genes whose expression is increased in, or restricted to, the SI, although differential hybridization will not be able to detect genes of less than 0.06% abundance. In addition, due to the higher content of white matter in the SI compared to the cerebellum, oligodendrocytic genes will also be detected by this screening method. Clones that did not react with cDNA probes of cerebellar RNA (8% of initial numbers) were replated in a gridded array on membranes and rescreened with higher cerebellum probe input, to increase the sensitivity of the assay. This removed a further 75% of the clones, leaving 2% of the initial number. Plasmid DNAs were prepared from these clones and equivalent amounts were immobilized onto replicate membranes. These membranes were hybridized with cDNA probes prepared from normal SI RNA, AD SI RNA, cerebellum RNA, liver RNA, placenta RNA or corpus callosum RNA. It was apparent from the slot hybridizations that there were no moderate abundance genes detectable that were enriched in the SI (Figure 3). Fourteen clones were selected that gave differential signals between SI and cerebellum and no detectable signal with liver or corpus callosum RNA. Subsequent Northern blot hybridizations were performed on these to determine brain specificity and regional expression (Table 1, part A). The limited sensitivity of cDNA probes for
hybridization detection of clones of low abundance genes was evident as all, except one of the selected clones that detected transcripts by Northern hybridization, were found to be detectable in cerebellum RNA and 8/11 of these were detectable in liver RNA. Clone pSI3a-64 only produced a hybridization signal with SI RNA. This signal was of very low intensity and formed a smear, rather than identifying a distinct band. Such a result is consistent with the presence of repetitive DNA sequences within the cDNA. This clone is being investigated further. The low abundance of regional specific clones is in line with the results of Travis et al. They screened a monkey cortex library by subtractive hybridization and identified 5 clones out

### Table 1: Detection of RNAs by Selected Clones

<table>
<thead>
<tr>
<th>Clone Number</th>
<th>Insert size (kbp)</th>
<th>Cbm</th>
<th>Signal detection</th>
<th>Placenta</th>
<th>SI</th>
<th>Transcript</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pSI3a-9</td>
<td>1.9</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>nd</td>
<td>2100 (WM+)</td>
</tr>
<tr>
<td>pSI3a-24</td>
<td>1.7</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>1320</td>
</tr>
<tr>
<td>pSI3a-30</td>
<td>1.2</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>++</td>
<td>1900</td>
</tr>
<tr>
<td>pSI3a-36</td>
<td>0.8</td>
<td>++</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>4100</td>
</tr>
<tr>
<td>pSI3a-50</td>
<td>1.3</td>
<td>++</td>
<td>+</td>
<td>+</td>
<td>nd</td>
<td>1600</td>
</tr>
<tr>
<td>pSI3a-60</td>
<td>1.8</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>nd</td>
<td>smear</td>
</tr>
<tr>
<td>pSI3a-64</td>
<td>1.6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>nd</td>
<td>no band</td>
</tr>
<tr>
<td>pSI3a-99</td>
<td>2.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>nd</td>
<td>no band</td>
</tr>
<tr>
<td>pSI3a-156</td>
<td>1.9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>nd</td>
<td>no band</td>
</tr>
<tr>
<td>pSI3a-164</td>
<td>0.6</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>nd</td>
<td>950,2150</td>
</tr>
<tr>
<td>pSI3a-182</td>
<td>0.8</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>no band</td>
</tr>
<tr>
<td>pSI3a-190</td>
<td>2.0</td>
<td>+/-</td>
<td>+</td>
<td>+</td>
<td>nd</td>
<td>2240</td>
</tr>
<tr>
<td>pSI3a-197</td>
<td>1.1</td>
<td>+</td>
<td>+</td>
<td>++</td>
<td>+</td>
<td>1600 (WM+)</td>
</tr>
<tr>
<td>pSI3a-203</td>
<td>0.9</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td>+</td>
<td>4300 (WM+)</td>
</tr>
<tr>
<td>B.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pSI3ac-2</td>
<td>0.8</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>1500 (WM+)</td>
</tr>
<tr>
<td>pSI3ac-7</td>
<td>1.4</td>
<td>++</td>
<td>+</td>
<td>nd</td>
<td>++</td>
<td>4800</td>
</tr>
<tr>
<td>pSI3ac-16</td>
<td>0.6</td>
<td>++</td>
<td>+</td>
<td>+</td>
<td>++</td>
<td>5 bands</td>
</tr>
<tr>
<td>pSI3ac-18</td>
<td>1.0</td>
<td>+</td>
<td>+/-</td>
<td>nd</td>
<td>+</td>
<td>3300 (WM+)</td>
</tr>
<tr>
<td>pSI3ac-22</td>
<td>0.8</td>
<td>++</td>
<td>++</td>
<td>nd</td>
<td>++</td>
<td>1400 (WM+)</td>
</tr>
<tr>
<td>pSI3ac-24</td>
<td>3.0</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>15500</td>
</tr>
</tbody>
</table>

- not detectable, +/- weak signal, + signal, ++ stronger signal, nd not determined, cbm cerebellum, SI substantia innominata, WM white matter

Transcript size estimated (bases RNA) relative to migration of B.R.L. RNA ladder.
and the cDNA inserts separated from the vector by gel electrophoresis. The DNA samples were transferred to hybridization membranes and the replicate filters were hybridized with cDNA probes made from normal SI RNA, AD SI RNA, cerebellum RNA, liver RNA, corpus callosum RNA and placenta RNA. Plasmid DNAs were isolated from selected clones which gave a differential hybridization signal by this procedure. These were digested with Eco R1 restriction endonuclease and the cDNA inserts separated from the vector by gel electrophoresis. The DNA samples were transferred to hybridization membranes and the replicate filters were hybridized with cDNA probes made from normal SI RNA, AD SI RNA, cerebellum RNA, liver RNA, corpus callosum RNA and placenta RNA. This modified screening procedure allowed the selection of clones that, although expressed in cerebellum, were differentially expressed in the SI. A summary of Northern hybridization data of clones selected by this alternate method is given in Table 1, part B. One clone (pSI3ac-24) containing a 3 kb insert, which was identified by the above described screening procedure, was selected for further characterization. The detailed characterization of this clone will be published elsewhere (manuscript in preparation, Boyes et al). This clone identified a mRNA of 15.5 kb that was detectable by Northern hybridization at different abundances in various brain regions and was not detectable in placenta or liver RNA samples. Attempts were made to quantify the relative levels of this gene in RNA extracted from the SI and temporal cortex of a number of AD and neurologically normal control cases, by Northern hybridization analyses. However, the intactness of the 15.5 kb band varied from case to case. In addition, because of its large size, a certain amount of degradation was always observed. As this degradation was inconsistent, quantification by Northern hybridization and gel autoradiograph densitometry appeared to be unreliable and a nuclease protection assay was used instead. This assay showed that this mRNA was detectable at a 30% higher level in AD temporal cortex RNA compared to normal cortex RNA, but at the same level in AD SI.
RNA compared to normal SI RNA. In addition, this gene was present in 4-fold greater amount in normal SI compared to normal cerebellum.

**CONCLUSION**

There is evidence now available for alterations of normal gene expression in AD affected brains. How these changes may lead to the neuronal degeneration and formation of amyloid containing neuritic plaques and neurofibrillary tangles can still only be the subject of speculation and a considerable amount of further investigation. Strategies for the identification of genes, from an SI cDNA library, that are differentially expressed in different brain regions and whose expression is altered in AD affected brains, have been developed. These strategies have led to the identification of anonymous genes that have been characterized on the basis of their expression in different brain regions. There is now available a range of scientific techniques that can be used to identify the proteins these genes code for and in which type of cells they occur. It is now possible to obtain the DNA sequence, identify the proteins these genes code for, and prepare synthetic peptide segments of these proteins. These peptides can be used to produce antibodies, which can be used to identify and isolate the proteins from brain tissue. With these techniques, new approaches for research in AD are now open.

**ACKNOWLEDGEMENTS**

This work was supported by grants from the Alzheimer Disease and Related Disorders Association, Medical Research Council of Canada and Alzheimer Society of British Columbia. DGW is a research fellow of the Alzheimer Society of British Columbia.

**REFERENCES**