Technical Advance

Single-Channel Quantitative Multiplex Reverse Transcriptase-Polymerase Chain Reaction for Large Numbers of Gene Products Differentiates Nondemented from Neuropathological Alzheimer's Disease

Stavros Therianos, Min Zhu, Eunice Pyun, and Paul D. Coleman

From the Center for Aging and Developmental Biology, University of Rochester Medical Center, Rochester, New York

Effective approaches using array technologies are critical to understand the molecular bases of human diseases. The results obtained using such procedures require analysis and validation procedures that are still under development. In the context of Alzheimer's disease, in which the identification of molecular mechanisms of underlying pathologies is vital, we describe a robust assay that is the first real-time reverse transcriptase-polymerase chain reaction-based high-throughput approach that can simultaneously quantitate the expression of a large number of genes at the copy number level from a minute amount of starting material. Using this approach within the human brain, we were able to quantitate as many as 19 genes at a time with only one type of fluorescent probe. The number of genes included can be considerably increased. Examples of consistent changes in Alzheimer's disease within these 19 candidate genes included reductions in targets related to the dendritic and synaptic apparatus. These changes were specific to Alzheimer's disease when compared with Parkinson's disease cases. We also present comparison data with microarray analysis from the same brain region and the same patients. The high sensitivity and reproducibility of this technology coupled with appropriate multivariate analysis is proposed here to form a biotechnology platform that can be widely used for diagnostic purposes as well as basic research. (*Am J Pathol* 2004, 164:795–806)

Simultaneous quantitation of numerous transcripts extracted from a defined tissue sample provides fundamental information for molecular neurobiology. Within identified states of a disease, such information helps the understanding of molecular cascades underlying pathologies. In the present work, our objective was to develop and validate a technology that would allow the coincident expression profiling and analysis of a large number of genes at the copy number level and from minute quantities of starting material. To do so, we have developed the single-channel quantitative multiplex reverse transcriptase-polymerase chain reaction (scqmRT-PCR) method. This approach provides the quantitation of copy number obtained with real-time PCR. In addition, by applying a totally different PCR strategy, this new method overcomes the limit of previous multiplex methods by considerably increasing the number of genes studied.^{1,2} By combining the advantage of fluorescent dye-coupled universal primers³ and gene-specific primers, and strategically altering the composition and ratio of the primer mixture, we have been able to routinely quantitate in parallel as many as 19 genes at a time (we can now routinely quantitate 35 targets) with only one type of fluorescent probe (FAM). The use of only one fluorescent reporter avoids the high background encountered in multichannel multiplex quantitative PCR methods. The uniformity, sensitivity, and specificity of our technology is equivalent to that of single-transcript real-time PCR.⁴

The drive to develop this new technology arose from the fact that currently available methods either are limited

Supported by the National Institutes on Aging (grant RO1 to P.D.C.) and a grant from an anonymous donor.

Accepted for publication December 4, 2003.

Address reprint requests to Dr. Paul D. Coleman, Center for Aging and Developmental Biology, University of Rochester Medical Center, 601 Elmwood Ave., Rochester, NY 14610. E-mail: paul.Lcoleman@urmc.rochester.edu.

to single digit number of genes at a time, restricted by the number of different fluorescence channels available or provide only semiquantitative gene expression levels by using fluorescence intensity as an indirect index, such as cDNA microarrays or large scale oligo-arrays.⁵ The development of these latter techniques forms the backbone of functional genomics and allows the analysis of thousands of genes in a single experiment. Nevertheless, results obtained with such approaches are inherently less precise and require further validation steps.^{6,7} Thus confirming cDNA and oligoarrays analysis with alternative, reliable, and high-throughput methods is needed. Currently, such validations are generally provided either by quantitative RT-PCR, *in situ* hybridization, or Northern blots. These approaches require large amounts of starting material, which limit their use in many circumstances such as human peroperative or postmortem samples. Moreover they must be performed in a sequential manner, consuming a protracted amount of time. The scqmRT-PCR method quantifies copy numbers of numerous transcripts in parallel using significantly less starting material. The sensitivity of this approach makes it potentially applicable in single-cell transcript analysis.^{8–10}

Here we first validated scqmRT-PCR by comparison to standard quantitative RT-PCR and then applied it in the context of Alzheimer's disease (AD). Using scqmRT-PCR we compared the transcript expression level of 19 candidate genes from aged-matched controls, AD as well as Parkinson's disease (PD) cases. Then we compared these scqmRT-PCR results with the outcome from oligo-arrays on the same human patients. We also performed multivariate analysis on scqmRT-PCR results, in an attempt to better interpret the biological meaning in such a high-throughput approach. Although AD and neurobiology are our primary interest, we stress that the method described here is broadly applicable to other kinds of pathological conditions.

Materials and Methods

Human Brain Tissues

Postmortem human brain tissues from superior frontal gyrus were obtained from the brain bank at University of Rochester, Rochester, NY. All cases were characterized based on clinical and neuropathological criteria as presented in Table 1.

RNA Extraction

Total RNA from 200 mg of human brain tissue homogenates was extracted using the RNeasy Protect midi kit (Qiagen, Valencia, CA). Each RNA preparation also included DNase I and proteinase K (Qiagen) treatment according to the manufacturer's instructions. Yield of total RNA was determined by absorbance at 260 nmol/L. RNA integrity was assessed by both 260/280 nmol/L ratios (ranging from 1.98 to 2.02) and agarose gel electrophoresis.

Table 1. Gender, Age, Postmortem Delay, and Clinical Dementia Rate of the Cases Used in This Study

Cases	Age (years)	Gender (M/F)	PMD* (hour)	CDR†
Control				
C1 (99A–112)	57	M	7.5	0
C2 (97A–179)	66	M	5.33	0
C3 (97A–191)	91	M	6.1	0
C4 (97A–212)	73	F	10.2	0
C5 (97A–237)	87	F	10	0‡
AD				
AD1 (97A–224)	86	M	5	5
AD2 (98A–030)	77	F	8.55	3
AD3 (98A–077)	84	F	5.35	5
AD4 (98A–175)	84	F	9.1	3
AD5 (99A–017)	87	F	5	3
PD				
PD1 (946)	78	F	6	N/D
PD2 (2688)	73	F	6	N/D
PD3 (1593)§	79	M	5	N/D
PD4 (2792)§	80	F	5.5	N/D

*PMD, postmortem delay.

†Clinical dementia rate.

‡Retrospective CDR.

§Additional Alzheimer's disease clinical diagnosis.

Reverse Transcription

One μ g of total RNA from each sample was reverse-transcribed into cDNA in a final volume of 20 μ l containing 4 U of Omniscript reverse transcriptase (Qiagen) in the manufacturer's buffer, 0.5 mmol/L of each dNTP, 10 U RNase inhibitor (Promega, Madison, WI), and 1 μ mol/L NVd(T)'s (5'TTTTTTTTTTTTTTTTTTVN3'). The reactions took place at 37°C for 12 hours and then were stored at –20°C until further use.

Single-Channel Multiplex Quantitative PCR

Real-time PCR reactions were performed using the Amplifluor Universal Detection system (Intergen, Purchase, NY) and iCycler (Bio-Rad, Richmond, CA). PCR primers were designed using Primer3 software (available at http://www-genome.wi.mit.edu/genome_software/other/primer3.html) to specifically amplify between 177 and 237 bp for the genes of interest in the same PCR conditions and were synthesized by Invitrogen, Carlsbad, CA. For each gene of interest, an additional forward primer was ordered that contained a Z-sequence (ACTGAACCT-GACCGTACA) at the 5' end required for UniPrimer annealing. Sequences of the PCR primers are shown in Table 2. The Amplifluor Universal Detection system kit is based on sunrise primer strategy. The UniPrimer contains the same Z-sequence, labeled with a reporter (FAM, 6-carboxy-fluorescein) at 5' and a quencher dye [DAB-SYL, 4-(dimethylamine)azo benzene sulfonic acid] at 3' of Z-sequence.

For the first round of multiplex quantitative PCR, each 100- μ l PCR reaction contained 1 μ l of cDNA or plasmid, 5 U of HotStarTaq DNA polymerase (Qiagen) in the manufacturer's buffer, 0.5 mmol/L of each dNTP, and 2 μ l of primer mixture. The primer mixture was made of forward and reverse primers for all of the genes of interest, at a

Table 2. Accession Numbers, Name of Gene Fragments, Primer Sequences, and Size of Amplicons Used in This Study

Accession number	Name of gene fragment	Forward primer	Reverse primer	Size of amplicon	Number
X00351	Actin	ACTGGGACGACATGGAGAAA	CAGAGGCGTACAGGGATAGC	204	1
U36922	FKHR	GCCATGTAAGTCCCATCAGG	AAAGAAAGACACCAAGCCATTT	207	2
J05633	Integrin-5	GCTCGACAGACATCAGCACA	GTCAGGTTTCCCAGAGTGGA	200	3
L37868	Oct-3	CTGGAGAGCCATTTCCCTCAA	TGTGGTGGAGTGTCCCTACTC	200	4
M16937	HOXB7	CAGACCTACACCCGCTACCA	CTCTGCTTCAGCCCTGTCTT	213	5
L33243	PKD1	TACCGCTACACCTGGGACTT	GAGCCATTGACCTTGATGCT	200	6
M28526	PECAM-1	CTCCAGCCAACTTCACCATC	ATGACCTCAAAGTGGGCATC	208	7
X52541	EGR-1	TGCGACATCTGTGGAAGAAA	GGGATGGGTATGAGGTGGT	214	8
U72671	ICAM-5	ATGGAGCCAACTTCTCGTGT	GAGGTAGACCCTGGCCTCTG	218	9
L07044	CAMK2G	GTGCCATTGAGGTCTTCAACA	AATAAGCTGTGCCAGGGTGT	194	10
M96954	TIA-1	GCCAATGGAGCCAAAGTGTAT	CATATCCGGCTTGTTAGGA	222	11
U83410	Cul2	CTTACTCCGTGCTGTGTCCA	GCCTTATCCAACGCACATCAT	216	12
M29551	PP2CB	CCACAGGGATGTTGCCTAGT	TGCGGTGTTTCCAGAGAATTGA	221	13
AF005043	PARG	AGCTGAGCGAGATGTGGTTT	CAGGACTCGACAGCATGGTA	213	14
X07979	ITGB	AGACCTGCCTTGGTGTCTGT	CCTCGTTGTTCCCATTCATC	237	15
X65873	KIF5B	AGACCTGCCTTGGTGTCTGT	ACCTCGTTGTTCCCATTCAC	189	16
AB014556	AP180	GAGATGCCTTTGCAGCTTCT	CTGGTCTCAGGTGGCTTCAT	190	17
L37792	Syntaxin	TTTGAGCAGGTGGAGGAGAT	ATGGACTGCTCGATGCTCTT	200	18
L07807	Dynamin	CCAAGGAGAACTGCCTCATC	GCTTGTCTCCAGCACATCA	177	19

final concentration of 10 $\mu\text{mol/L}$ each. The forward primers used here did not contain the Z-sequence. The PCR program consisted of 15 minutes at 95°C to activate the polymerase, followed by 15 cycles of 20 seconds of denaturation at 95°C, 20 seconds of annealing at 60°C, and 35 seconds of elongation at 72°C. A final step of elongation at 72°C for 10 minutes was performed. This round of PCR was preamplification only and did not involve real-time PCR.

For the second round of multiplex quantitative PCR, each 50 μl real-time PCR reaction contained 1 μl of first round multiplex quantitative PCR reaction, 2.5 U of Hot-StarTaq DNA polymerase (Qiagen) in the manufacturer's buffer, 0.5 mmol/L of each dNTP, 0.02 $\mu\text{mol/L}$ forward primer, and 0.2 $\mu\text{mol/L}$ reverse primer for one gene, and 0.2 $\mu\text{mol/L}$ UniPrimer. The PCR program consisted of 15 minutes at 95°C to activate the polymerase, followed by 50 cycles of 20 seconds of denaturation at 95°C, 20 seconds of annealing at 60°C, and 35 seconds of elongation at 72°C. A final step of elongation at 72°C for 10 minutes was performed. Fluorescence intensity was measured during the annealing step of each cycle, so that unincorporated Uni-Primers were predominantly in the quenched hairpin conformation. Threshold cycle (C_T) for each reaction was analyzed using iCycler software (Bio-Rad).

All real-time PCR experiments were performed in triplicates and the average C_T for the triplicates was used in all subsequent analysis. Reactions omitting enzyme or template were used as negative controls. All reactions were resolved in 1% agarose gel to confirm the PCR specificity. The amount of transcripts was calculated by reference to respective standard curves.

Regular Quantitative PCR

Reaction mixture and conditions were the same as the second round of multiplex quantitative PCR, except that the PCR template was 1 μl of plasmid.

Cloning and Constructing Standard Curves

Regular PCR was performed using cDNA as template and the same set of primers for each gene of interest. The forward primers used did not contain Z-sequence. Each PCR product was cloned into pGEM-T Easy vector (Promega). Plasmids were quantitated by absorbance at 260 nmol/L. Eight 10-fold serial dilutions of plasmids for each gene of interest were used as templates to perform multiplex quantitative PCR individually in triplicates. Thus a standard curve was constructed for each gene of interest. A linear relationship between the threshold cycles and the log value of input plasmid DNA copy number was observed throughout the range of 10^1 to 10^8 copies.

Microarray

Double-stranded DNA was synthesized from 15 μg of total RNA by using one primer containing poly (dT) and the other primer containing T7 polymerase promoter sequence. *In vitro* transcription with the double-stranded DNA as a template in the presence of biotinylated UTP and CTP was performed using the protocol provided by Affymetrix (Santa Clara, CA). Biotinylated cRNA was purified, fragmented, and hybridized to HuGeneFL arrays following manufacturer's manual. The hybridized arrays were then washed and stained with streptavidin-phycoerythrin, and scanned with a Hewlett Packard GeneArray Scanner. Data analysis was performed using Affymetrix Genechip Expression Analysis software (version 3.1 and 5.0). Internal controls of housekeeping genes and a test chip were run before test samples.

Principle Component Analysis

Data from multiplex quantitative PCR and microarray were first transformed into Excel files, and then imported

Figure 1. Overview of scqmRT-PCR.

into S-Plus statistical software package (Insight) as data files. Principle component analysis was performed with either all or selected variables using default settings in S-Plus. The first two principle components were used to make the scatter plots. A screenplot and a loading bar graph were also generated in each analysis by the software.

Results

Experimental Design of scqmRT-PCR

The overview of the strategy used to analyze numerous transcripts in parallel is outlined in Figure 1. The first step was to design and validate a set of primers that would share the same PCR conditions. These conditions are comparable to some extent as those applicable in qualitative multiplex RT-PCR.¹¹ Common parameters included annealing temperature ($60 \pm 0.5^\circ\text{C}$), GC content ($50 \pm 5\%$), and amplicons with 180 to 200 bp. Longer PCR products tended to increase background. These features were essential to allow all of the multiplex PCR reactions to occur at maximum efficiency. Construction of a series of standard curves took place after primer design. To achieve this step, it was necessary to subclone each candidate in a plasmid vector. To do so, a regular qualitative multiplex RT-PCR was performed (data not shown). This step not only allowed subcloning the candidates of interest but also provided a quality test for the primer sets. In other words, any incompatibility among the primers such as intercomplementarity or self-complementarity could be detected at this point and relevant primers could be redesigned. The specificity of each primer pair was also tested through this step. We have found this empirical quality control to be less time consuming and more reliable than any control using primer design software. Once our targets had been subcloned, standard curves for each gene of interest were constructed (see validation below).

The third step (step 3 in Figure 1) was concerned with sample preparation. Total RNA (Rneasy Midi kit, Qiagen)

was extracted from samples following the manufacturer's instructions and reverse-transcribed into a first strand cDNA. Once the first strand cDNA had been synthesized, a first round of PCR was performed. This first round of PCR was performed with 1 μ l of the RT (step 4). In this round, all of the primers of interest were mixed together in an equimolar concentration and the final volume was 100 μ l, with the proper concentration of dNTP's, enzyme, and appropriate mix (see Material and Methods). We have limited this first round to 15 cycles to guarantee that even the most abundant messages such as β -actin were still within a linear range of amplification when starting with 1 μ g of total RNA. The PCR conditions of this first round were the same as in the previous regular multiplex RT-PCR (step 2) for subcloning and quality control. The next step (step 5) consisted of a series of single-channel real-time quantitative PCR reactions. To achieve this step, an aliquot of the first round PCR (1 μ l) was mixed on ice in a total volume of 50 μ l with each specific forward primer tagged with a universal sequence (Intergen) at 5' and the reverse primer specific for every gene of interest.¹²⁻¹⁴ Appropriate concentrations of enzyme and reagents necessary for the reaction were added to the solution. Reactions for all of the genes of interest were performed in parallel. This second round of PCR was performed in a real-time quantitative thermocycler (iCycler, Bio-Rad) and quantitation of the fluorescent emission was recorded during each cycle of PCR. The copy number for each gene of interest was then calculated based on threshold cycles using the corresponding standard curve. Here it should be emphasized that indeed up to 19 quantitations have been performed in parallel from the initial first strand cDNA (step 3). The next step (step 6) consisted of data analysis. Several analyses can be performed to achieve different goals. Examples will be presented in the next sections.

Reliability and Validation of scqmRT-PCR

As a first step toward the validation of scqmRT-PCR, we subcloned 19 candidate targets (described in Table 2) in cloning vectors (pGEM-T Easy Vector system, Promega). The specificity of our primers was further confirmed by sequencing each inserted clone of interest. Then, with plasmids containing each target gene, we performed the procedure shown in Figure 1 (steps 2 to 5). We used eight serial dilutions of equimolar concentrations of the 19 plasmids ranging from 10⁸ to 10¹ copies as a template for the first round of amplification. This allowed us to process in parallel these 19 inserts at different concentrations and let us construct 19 different standard curves from the same aliquot, each time with a different copy number in triplicate. An illustration of the uniformity of our amplifications is shown in Figure 2A. Analysis of the triplicate repeats led to standard deviations of threshold cycle number ranging from 0.05 to 0.18 cycles within the dilution series used. By using all data collected in this experiment and combining these data with appropriate selections for baseline cycles and threshold, the final result for 24 wells (ie, 8 dilutions $\times$ 3 triplicates) was a mean

Figure 2. Validation steps of scqmRT-PCR. **A:** Nineteen targets were processed in parallel from the same amount of starting material and a representative example of threshold cycles obtained during the quantitative round of PCR is shown. **B:** Standard curves related to the 19 target transcripts were constructed derived from the threshold cycles. **C:** One percent agarose gel run after the quantitation as a demonstration of amplicon specificity. Such gels are run after each experiment. **D:** Comparison of regular quantitative RT-PCR and scqmRT-PCR. The upper trend line was obtained following a regular quantitative RT-PCR protocol in which 10^4 copies of starting material were considered as an unknown copy number on the thermocycler settings. The lower standard curve represents the same experiment performed using scqmRT-PCR protocol. The same starting mRNA copy number was obtained using both techniques.

threshold cycle of 29.6 and a SD of 0.21 (SD = 0.7% of the mean). Each of the 19 background-corrected data were brought down to the PCR baseline to form standard curves as illustrated in Figure 2B, which shows a representative example of the amplification linearity that can be achieved with scqmRT-PCR. The standard curve correlation coefficients ranged between 0.999 and 0.980 throughout a range of eight orders of magnitude. This range of correlation coefficient is in accordance with what can be obtained using regular quantitative RT-PCR. After each second round of PCR, a 1% agarose gel was run to check for specificity of the amplifications (Figure 2C). This step is very important in scqmRT-PCR because one has to be sure that no nonspecific amplification or contamination arising from another set of primers has occurred. The sensitivity of scqmRT-PCR allows reproducible amplification of starting material containing 10 to 100 copies of transcript (10 copies as shown in Figure 2B). In some instances we could undoubtedly reach the threshold cycle with single copy template (data not shown). This level of sensitivity makes our procedure compatible with a large range of applications and reduces by several orders of magnitude the amount of starting material necessary for quantitation over that required for arrays or Northern blots. To further validate scqmRT-PCR we compared the results obtained after regular quantitative RT-PCR and scqmRT-PCR using several targets. Figure 2D illustrates the results we have obtained. There, we have set one dilution (10^4 copies in the example provided) as an unknown concentration in the thermocycler settings.

The first observation was that the absolute value of threshold cycles was significantly reduced using scqmRT-PCR ($P < 0.001$) when standard curves were constructed. This did not affect the calculation of the accurate copy number of starting material, because the preamplification was kept in the linear range, thus not changing the original relationship between threshold cycles and log of copy number. It should be noted that separate standard curves for regular quantitative RT-PCR were constructed. Yet, using both approaches, the same copy number of starting material was obtained. Therefore, our technology gives the same result in term of absolute quantification relative to plasmid when compared to more traditional approaches. These results demonstrating the uniformity, reproducibility, linearity, and sensitivity of scqmRT-PCR formed a technical proof of feasibility.

Application of scqmRT-PCR in the Context of AD and PD

After these validation steps, it was important to demonstrate that our approach can be applied in a biological context. To do so, we investigated the expression of 19 targets in AD (Table 2). For comparative purposes, these targets were selected based on results obtained with oligo-arrays that will be discussed below. It is important to mention that we did not use oligo-arrays as a validation of scqmRT-PCR. These results formed the backbone

of our analysis. We used regular quantitative RT-PCR to positively validate our technology (Figure 2D), as this approach is the closest to what we have developed here. After adequate proteinase K (Roche, Indianapolis, IN) and DNase I (Promega) digestions, we extracted total RNA following the manufacturers' instructions. Comparison of results obtained from total RNA *versus* mRNA showed no difference in data reproducibility (data not shown). Consequently, subsequent preparations used total RNA. We used five AD cases, three age-matched controls (described as controls in the text), and two cases whose autopsy report met the neuropathological criteria of the National Institute on Aging-Reagan Institute^{15,16} but had no clinical signs of dementia (retrospective clinical dementia rate, 0). We will describe these two cases as "preclinical" throughout the text. The AD cases satisfied both clinical and neuropathological criteria for AD.¹⁵ We have also included four PD cases, among them two presenting concomitant AD (Table 1), to test the specificity of scqmRT-PCR within different neurodegenerative diseases. Figure 3 summarizes the copy numbers per μg total RNA of 19 genes obtained from the 10 cases studied. Each measure was in triplicate (SD is plotted but is too small to be visible). PCR products for the 19 candidate targets were gel-extracted and subcloned after the second round of PCR (Figure 4F). All 19 PCR products yielded the expected sequences. Because of the inherent variability among human patients, we found it to be more informative to present the data with each individual described separately rather than grouping our results as means of controls or means of AD. The presentation of individual data provides information that may otherwise be obscured. The data show a clear separation between control and AD cases for a number of candidate messages (Figure 3A). Controls (which did not include the two preclinical cases) always had a higher mRNA copy number per μg total RNA for seven candidate genes studied: AP180, PP2CB, Dynamin, Syntaxin, PARG, CAMKG, ICAM5 (Figure 3A). A majority of these targets are related to the dendritic or synaptic apparatus, which has been proposed to be affected early in AD.^{17–19} Moreover, the expression pattern of the two preclinical cases was more similar to the AD group than to the control group (Figure 3A). Within the above seven messages, we could also observe that the AD population, which is representative of late AD stages both clinically and morphologically, was more homogenous in terms of copy numbers than compared to the control and preclinical cases, similar to what has been observed in previous studies.¹⁰ However, the greater heterogeneity in the cases representing control plus preclinical is almost entirely because of preclinical cases falling among the AD values. Thus, if the cases that were clinical controls but neuropathological AD were excluded from the more strictly defined control population, then the control cases were not more heterogeneous than AD cases. Egr-1 represents an exception to these comments because controls were more heterogeneous and overlapped with AD values (Figure 3A). Another set of genes, including HOXB7, PKD1, β -Actin, Oct-3, KIF5B, FKHR, Intergrin- β 5, and ITGB showed an heterogeneous distribution of

copy numbers and extensive overlap of both groups (Figure 3B). In the case of the endogenous endonuclease TIAL1²⁰ and the acute inflammatory response protein PECAM1²¹ the control population and the preclinical cases were more homogenous than the AD group. The hypoxia-induced mRNA regulator CUL2^{22,23} seemed to belong to this group with one noticeable exception in the control population (control case 5).

Application of Principal Component Analysis (PCA)

PCA was used to reduce the dimensionality of our data set and to extract further meaningful biological information. First, the entire set of genes was used to perform PCA and a two-dimensional plot of the first two principal components was constructed (Figure 4A). The advantage of PCA analysis is that it is inherently performed without disease group attribution or categorization. The algorithm does not take in account any previous knowledge of disease category. This analysis showed clustering of cases according to their disease status. In particular, the preclinical cases were positioned closer to the AD cluster than to the control cluster. The first two components were sufficient in this analysis to account for 75.5% of the variance among our candidates (Figure 4C). The messages that contributed heavier weights to component 1 were Dynamin, AP180, ICAM5, PP2CB, Syntaxin, and Actin. The messages that contributed heavier weights to component 2 were PKD1, KIF5B, HOXB7, Integrin5, ITGB, and FKHR. We then selected the seven genes related to the dendritic and synaptic apparatus (Figure 3A) and performed PCA on this set of candidates. Using this collection of genes we could observe a more pronounced clustering of the AD cases and again the preclinical cases were closer to AD cases than to controls (Figure 4B). Here, the first component accounted for 78.5% of the variance among our candidates (Figure 4D). Addition of the second component increased the variance accounted for to 92.3%. The fact that this set of messages allowed a clear separation between AD and control suggests that they could be used to separate AD from control groups with low probability of error. This is further strengthened by the fact that, based on the same targets related to the dendritic and synaptic apparatus, we could also separate pure age-matched PD cases from age-matched PD cases with concomitant AD (Figure 4E). Furthermore, this set of targets could be used in any postmortem situation in which the final diagnosis of AD is difficult.¹⁵ Moreover, this set of results can lead to further investigations about the connection of these genes to the cell biology of the disease. Altogether, the data presented in Figures 3 and 4 demonstrate that scqmRT-PCR can be used in a biological paradigm in which transcript populations are of major interest.

Affymetrix Samples Compared with scqmRT-PCR

We propose that another possible use of scqmRT-PCR is to validate oligo-arrays or cDNA arrays. Indeed these

Figure 3. mRNA copy number per μg total RNA comparisons from control, preclinical, and AD cases. **A:** Seven transcripts showed consistent change between controls (black, **left column**) and AD cases (**right column**). Note that the preclinical cases (gray, **left column**) matched closer to the AD group. AP180, Dynamin, Syntaxin, ICAM5, and CamK2G are related to the dendritic and the synaptic apparatus. EGR1 showed a greater heterogeneity within the control group. **B:** Eight transcripts displayed heterogeneity in their mRNA copy numbers in the three groups. **C:** The three transcripts that showed higher homogeneity within control and preclinical cases compared to AD cases.

techniques require confirmation steps and there is a lack of an economical high-throughput technique in this area. En masse identification of the mRNAs differentially ex-

pressed between controls and AD cases was achieved using Affymetrix human U95 oligonucleotide microarrays. Total RNA extracted from the same cases used for the

Figure 4. PCA performed on scqmRT-PCR results. **A:** Two-dimensional plot constructed based on the entire set of genes ($n = 19$). Cases clustered according to their disease status and preclinical cases (C, gray) were positioned closer to the AD cases (AD, black) than to the control cases (C, black). **B:** Same analysis achieved with AP180, PP2CB, Dynamin, Syntaxin, ICAM5, PARG, and CamK2G. This set of transcripts was sufficient to separate control cases from AD cases and the preclinical cases clustered closer to the AD group. **C:** Relative importance of principal components for the 19 candidate genes. Note that the first two components accounted for 75.5% of the variance among the cases. **D:** Relative importance of principal components for seven candidate genes including AP180, PP2CB, Dynamin, Syntaxin, ICAM5, PARG, and CamK2G. Here the first two components accounted for 92.1% of the variance among the cases. **E:** Two-dimensional plot including four PD cases. Two PD cases (underlined) with concomitant AD clustered close to AD cases. The same set of genes as used for **B** were used for this analysis. **F:** Representative gel run after the two rounds of PCR. Targets are in the same order as in Figure 2. The two last lanes, without amplicons, are negative controls (no RT).

rest of this study were hybridized and analyzed following published procedures.²⁴ We then compared the fold changes for each target between age-matched controls and AD cases (Figure 5A). The synaptic vesicle endocytosis AP180²⁵ was excluded because this gene is not represented on the oligonucleotide arrays. In 7 (41%) of the 17 targets we noticed a discrepancy between the two

techniques in term of a trend greater than onefold change. These seven targets were FKHR,²⁶ Integrin 5,²⁷ Oct 3,²⁸ PKD 1,²⁹ PECAM 1, EGR 1,³⁰ and KIF 5B.^{31,32} Within these seven transcripts, four of them went in opposite directions (FKHR, Integrin-5, Oct-3, and PECAM). These inconsistencies in array confirmation have been reported by others^{6,33,34} yet are underplayed in many

Figure 5. Comparison of microarrays and scqmRT-PCR. **A:** Fold changes between control and AD cases measured with either microarray data or scqmRT-PCR showed inconsistencies for several candidates including FKHR, Integrin 5, Oct 3, and PECAM 1. **B:** scqmRT-PCR two-dimensional plot of principal components constructed with 18 genes that are also present of microarrays. Note that the preclinical cases (C, gray) clustered with AD cases (AD, black). **C:** Same analysis as for **B** but based on microarray's indirect fluorescence index. Here, the two preclinical cases were not discernible from controls despite their AD histological pathology.

studies. The use of a newer statistical algorithm proposed by the manufacturer (MAS 5.0) did not change this situation. Interestingly, when PCA was used on the array data and compared with PCA on the same targets processed with scqmRT-PCR (Figure 5B), the two populations were separated as expected. However, within the PCA analysis performed on array data, the two preclinical cases failed to be separated from the controls (Figure 5C). Our data again highlight the increased sensitivity of scqmRT-PCR as well as the importance of validation steps following functional genomic approaches. It is clear that more candidates should be analyzed and compared to array data but such a high percentage of incongruity within 18 targets is notable.

Discussion

We have developed a PCR-based high-throughput methodology to gain further insight into molecular mechanisms underlying AD. By adding to proven qualitative multiplex RT-PCR methods³⁵ a totally new single-channel real-time strategy, we were able to quantify in parallel copy numbers of 19 transcripts from a fairly small amount of starting material. Using this approach we could separate an age-matched control population from both the AD and PD population with copy number precision and based on the expression pattern of a subset of functionally related candidates. Furthermore we could highlight nondemented cases with AD pathology within the heterogeneous control population. These preclinical cases

were indeed closer to the AD population than to controls at the molecular level. Finally, this approach allowed us to strengthen the precision of results obtained from oligo-arrays.

These findings suggest that the use of scqmRT-PCR could be applied routinely for both fundamental research and clinical diagnosis. Here we analyzed 19 transcripts in parallel as a proof of principle. It should be stressed that if 1 μ g of total RNA were to be used as starting material, the number of transcripts analyzed in parallel can be greatly increased based on the following facts: first we have used only 5% of the reverse-transcribed material to perform our analysis; second for each second round of PCR reaction we have used 1% of the first round products. Based on this, we estimate that with appropriate equipment, up to 500 targets could be analyzed in parallel from 1 μ g of total RNA without modifying our protocol, assuming that it is possible to generate a compatible primer mix for the first round of PCR. Alternatively, the quantity of starting material could be highly reduced to the level of single-cell transcript amplification. Indeed we perceive scqmRT-PCR as a complementary approach to other single-cell technologies such as aRNA amplification.⁹ The synergy of such techniques could give rise to a yet unachieved level of precision at the single-cell level. Here our real-time strategy was based on sunrise primers that contain a universal sequence on their 5' end.¹⁴ The method could be also adapted to a molecular beacon approach (TaqMan) without great modifications.³ Here we have used standard curves to derive absolute copy number of targets, although alternative approaches such as analysis using C(t), could also be applied. We have demonstrated that scqmRT-PCR is a uniform, reproducible, linear, and sensitive technique. This approach of real-time RT-PCR is comparable to regular quantitative PCR yet the potential number of transcripts that can be analyzed is increased by orders of magnitude. This is a breakthrough in the field of the molecular biology of AD in the sense that up to now, such an approach was limited to the study of one or a few targets at a time. With the addition of our technology in the field, the numbers of candidates analyzed in parallel can be tremendously increased. Technically, for high and moderately expressed genes, one round of PCR would be sufficient to reach a threshold cycle and lead to copy number calculation. We did not adapt this approach though for several reasons. 1) This would lead to a requirement for an increased amount of starting material (total RNA) because each reaction would have to be processed completely separately. Tissue (from autopsy, peroperative, biopsy samples, and so forth) is generally difficult to obtain and when obtained, may be in very low amounts. 2) For genes that are expressed at low levels but are carrying a lot of biological information, the first round would still be necessary. 3) Samples that have been quantified in parallel lead to more reproducible results than samples that have been quantified sequentially.

Comparisons of gene expression between normal aging and neurodegenerative diseases is frequently hampered by the fact that housekeeping genes such as GAPDH and β -actin, commonly used as reference val-

ues, are changed at the transcript level during the course of aging and disease. Altered expression of these genes in AD is consistent with the metabolic and structural changes known to occur in AD.^{36–38} The problems induced by the use of housekeeping genes as reference standards are mitigated by the quantification of absolute copy number produced by our procedure. Thus, by bringing together quantitation at the copy number level for each target and quantitation of numerous targets in parallel, we have been able to define a set of genes, primarily related to the dendrite and the synapse, that are consistently changed in AD. This is in accordance with findings that the dendritic and synaptic machinery is undoubtedly affected in AD.^{39–41} We also observe that *in situ*, the fold changes in gene products between AD and controls can be considered as rather modest when compared to what is commonly observed in animal model systems or *in vitro*. In our hands, this observation strengthens the possibility that the cellular manifestations of AD are caused by subtle changes in gene expression of numerous members within one functional group (here a class of genes related to the synaptic and the dendritic apparatus) rather than on the alteration of a single gene expression or mutation. Another set of transcripts showed a greater variability among control and disease, including the housekeeping gene β -actin. Further experiments on other classes of genes (DNA repair, molecular motors, transcription factors, cytoskeleton related proteins, and so forth) will be needed to consider whether this variability can be interpreted either as a result of the progression of the disease or as a result of individual differences in cellular mechanisms in AD. This greater variability among some gene products also highlights the possibility that numerous genes and transcripts of moderate or idiosyncratic effect are likely to play interdependent roles during the course of AD.

With our approach we detected the presence of HOXB7 in both control and AD samples. This is surprising first because this homeodomain protein has been described principally during embryonic development but even more because the expression of this particular homeodomain protein is restricted to caudal segments during embryogenesis.⁴² We propose that our results reflect the presence of macrophages within the sample homogenates rather than a genuine neuronal or glial expression.^{43,44} Both single-cell gene expression experiments and *in situ* hybridization will be necessary to further investigate this hypothesis.

Using the expression pattern of a group of genes to separate different disease subtypes has been a promising approach for clinical diagnosis.^{45–47} However, current publications were derived mainly from microarray studies, which are restricted from practical application for a number of reasons (time, costs, and so forth). Our method provides more flexibility in choosing candidate genes and allows robust separation of groups with a significantly smaller number of genes. Indeed the coupling of scqmRT-PCR with multivariate statistical analyses such as PCA is a very promising tool for the early identification of several diseases, not limited to neurodegenerative disorders. Based on our results with seven

transcripts related to the dendritic and synaptic apparatus, we suggest that this combination of molecular and multivariate statistical tools, displayed by the use of scqmRT-PCR coupled with PCA, can be used to discriminate between age-matched control and preclinical AD cases in an elegant, precise, and economical way. As an example, in this study, preclinical cases that did not meet clinical criteria for AD but did meet neuropathological criteria for AD at autopsy were separated from controls based on their gene expression. We would like to emphasize that the goal of the present study was primarily to validate the scqmRT-PCR technology and then to apply it in a neurodegenerative context to demonstrate its biological relevance. Therefore, the number of cases incorporated here does not allow any assumption about sensitivity and specificity *stricto sensu*. However, the two preclinical control cases clustering and overlapping with AD cases raise an important issue. This clustering could indeed be interpreted as a lack of both sensitivity and specificity of a diagnostic test but based on the number of cases studied, we would not and do not propose this approach, yet, as a postmortem diagnostic tool for AD. This would be even truer if only the clinical assessment was to be taken in account. To be accurate, these two cases had no clinical signs of AD but showed patent neuropathological features of the disease at autopsy in terms of plaques and tangles. On our side, using scqmRT-PCR, we find molecular signs that are going in the same direction as the neuropathology report. Our statement concerning these two cases is that they were as close as it can be to the earliest evidence of clinical signs. In other words, would it be possible to perform a premortem, and less invasive, test on living individuals, we would predict that this type of clustering of clinically silent controls to AD patients announces the entry into the disease. This will be of critical importance with respect to prognostic methods and therapeutic interventions. Furthermore, the results obtained with PD cases indicated that our approach allows a specific separation of AD cases compared to pure PD cases. Interestingly, PD cases showing associated AD clustered close to AD cases. In fact, this kind of test could be a prerequisite for any large-scale analysis in the sense that it could rapidly separate different populations of interest at a molecular level.

AD is a complex, dichotomous, and heterogeneous disease.⁴⁸ Both based on a pathobiological and a genetic linkage approach, the search for strong AD candidates now relies heavily on the use of large-scale microarrays. It is also widely recognized that although clearly essential, array approaches need independent confirmation that will distinguish among consistent, inconsistent, or likely false-positive/negative findings. scqmRT-PCR complies with the parameters of such an independent experimental technique that will allow validation of microarrays. In our study, arrays results were confirmed by scqmRT-PCR only for approximately half of the candidates. It should be pointed out that most of the candidates showed a consistent change on arrays despite the fact that the fold changes were not important. In other words, these changes were robust and reproduc-

ible within the samples studied using Affymetrix technology but not confirmed when tested with scqmRT-PCR. When these transcripts were analyzed as a whole using PCA, both technologies were able to separate controls from AD cases but arrays lacked the level of precision necessary to distinguish cases that were preclinical. Array analysis gave rise to an approximately accurate description and scqmRT-PCR increased this level to correspond more closely to postmortem diagnosis. We want to emphasize that our goal was not primarily to validate oligo-arrays with scqmRT-PCR. We performed the validation step of our technology using regular quantitative PCR with success. Our comparative results with oligo-arrays outline the limitations in terms of reproducibility that are inherent to functional genomics based on human post-mortem approaches. Clearly the use of animal models will further strengthen and validate this predictive aspect of scqmRT-PCR use. It would be interesting to compare the oligonucleotide sequences in the Affymetrix array with the sequences amplified with our primers but this information is not accessible to public use. We also think that these discrepancies highlight limitations of microarrays because of, among other reasons, differential stringency conditions among the population of targets studied. The data provided here form the canvas for further investigation into the molecular basis underlying AD. Moreover, given the ease of scqmRT-PCR, other diseases can be studied at the molecular level as well.

Acknowledgments

We thank Kelly Glajch for technical help and the Human Brain and Spinal Fluid Resource Center, VAMC, Los Angeles, CA (which is sponsored by the National Institute of Neurological Disorders and Stroke/National Institute of Mental Health, the National Multiple Sclerosis Society, the VA Greater Los Angeles Health Care System, Los Angeles, CA, and the Veterans Health Services and Research Administration, Department of Veterans Affairs), for the Parkinson's disease specimens.

References

- Chiu RW, Murphy MF, Fidler C, Zee BC, Wainscoat JS, Lo YM: Determination of RhD zygosity: comparison of a double amplification refractory mutation system approach and a multiplex real-time quantitative PCR approach. *Clin Chem* 2001, 47:667–672
- Moody A, Sellers S, Bumstead N: Measuring infectious bursal disease virus RNA in blood by multiplex real-time quantitative RT-PCR. *J Virol Methods* 2000, 85:55–64
- Bustin SA: Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J Mol Endocrinol* 2000, 25:169–193
- Freeman WM, Walker SJ, Vrana KE: Quantitative RT-PCR: pitfalls and potential. *Biotechniques* 1999, 26:112–124
- Young RA: Biomedical discovery with DNA arrays. *Cell* 2000, 102:9–15
- Rajeevan MS, Vernon SD, Taysavang N, Unger ER: Validation of array-based gene expression profiles by real-time (kinetic) RT-PCR. *J Mol Diagn* 2001, 3:26–31
- Mills JC, Roth KA, Cagan RL, Gordon JL: DNA microarrays and beyond: completing the journey from tissue to cell. *Nat Cell Biol* 2001, 3:E175–E178
- Cauli B, Porter JT, Tsuzuki K, Lambolez B, Rossier J, Quenet B, Audinat E: Classification of fusiform neocortical interneurons based on unsupervised clustering. *Proc Natl Acad Sci USA* 2000, 97:6144–6149
- Miyashiro K, Dichter M, Eberwine J: On the nature and differential distribution of mRNAs in hippocampal neurites: implications for neuronal functioning. *Proc Natl Acad Sci USA* 1994, 91:10800–10804
- Chow N, Cox C, Callahan LM, Weimer JM, Guo L, Coleman PD: Expression profiles of multiple genes in single neurons of Alzheimer's disease. *Proc Natl Acad Sci USA* 1998, 95:9620–9625
- Ruano D, Lambolez B, Rossier J, Paternain AV, Lerma J: Kainate receptor subunits expressed in single cultured hippocampal neurons: molecular and functional variants by RNA editing. *Neuron* 1995, 14:1009–1017
- Winn-Deen ES: Direct fluorescence detection of allele-specific PCR products using novel energy-transfer labeled primers. *Mol Diagn* 1998, 3:217–221
- Nuovo GJ, Hohman RJ, Nardone GA, Nazarenko IA: In situ amplification using universal energy transfer-labeled primers. *J Histochem Cytochem* 1999, 47:273–280
- Nazarenko IA, Bhatnagar SK, Hohman RJ: A closed tube format for amplification and detection of DNA based on energy transfer. *Nucleic Acids Res* 1997, 25:2516–2521
- Gearing M, Mirra SS, Hedreen JC, Sumi SM, Hansen LA, Heyman A: The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part X. Neuropathology confirmation of the clinical diagnosis of Alzheimer's disease. *Neurology* 1995, 45:461–466
- Ball M, Braak H, Coleman P, Dickson D, Duyckaerts C, Gambetti P, Hansen L, Hyman B, Jellinger K, Markesbery W, Perl D, Powers J, Price J, Trojanowski JQ, Wisniewski H, Phelps C, Khachaturian Z: Consensus recommendations for the postmortem diagnosis of Alzheimer's disease. The National Institute on Aging, and Reagan Institute Working Group on Diagnostic Criteria for the Neuropathological Assessment of Alzheimer's Disease. *Neurobiol Aging* 1997, 8(Suppl 4):S1–S2
- Minger SL, Honer WG, Esiri MM, McDonald B, Keene J, Nicoll JA, Carter J, Hope T, Francis PT: Synaptic pathology in prefrontal cortex is present only with severe dementia in Alzheimer disease. *J Neuropathol Exp Neurol* 2001, 60:929–936
- Scheff SW, Price DA, Sparks DL: Quantitative assessment of possible age-related change in synaptic numbers in the human frontal cortex. *Neurobiol Aging* 2001, 22:355–365
- Maccioni RB, Munoz JP, Barbeito L: The molecular bases of Alzheimer's disease and other neurodegenerative disorders. *Arch Med Res* 2001, 32:367–381
- Kawakami A, Tian Q, Duan X, Streuli M, Schlossman SF, Anderson P: Identification and functional characterization of a TIA-1-related nucleolysin. *Proc Natl Acad Sci USA* 1992, 89:8681–8685
- Newman PJ, Berndt MC, Gorski J, White II GC, Lyman S, Paddock C, Muller WA: PECAM-1 (CD31) cloning and relation to adhesion molecules of the immunoglobulin gene superfamily. *Science* 1990, 247:1219–1222
- Pause A, Lee S, Worrell RA, Chen DY, Burgess WH, Linehan WM, Klausner RD: The von Hippel-Lindau tumor-suppressor gene product forms a stable complex with human CUL-2, a member of the Cdc53 family of proteins. *Proc Natl Acad Sci USA* 1997, 94:2156–2161
- Pause A, Lee S, Lonergan KM, Klausner RD: The von Hippel-Lindau tumor suppressor gene is required for cell cycle exit upon serum withdrawal. *Proc Natl Acad Sci USA* 1998, 95:993–998
- Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES: Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 1999, 286:531–537
- Mao Y, Chen J, Maynard JA, Zhang B, Quirocho FA: A novel all helix fold of the AP180 amino-terminal domain for phosphoinositide binding and clathrin assembly in synaptic vesicle endocytosis. *Cell* 2001, 104:433–440
- Anderson MJ, Shelton GD, Cavenee WK, Arden KC: Embryonic expression of the tumor-associated PAX3-FKHR fusion protein interferes with the developmental functions of Pax3. *Proc Natl Acad Sci USA* 2001, 98:1589–1594
- Reynolds LE, Wyder L, Lively JC, Taverna D, Robinson SD, Huang X, Sheppard D, Hynes RO, Hodivala-Dilke KM: Enhanced pathological

- angiogenesis in mice lacking beta3 integrin or beta3 and beta5 integrins. *Nat Med* 2002, 8:27–34
28. Schreiber E, Tobler A, Malipiero U, Schaffner W, Fontana A: cDNA cloning of human N-Oct3, a nervous-system specific POU domain transcription factor binding to the octamer DNA motif. *Nucleic Acids Res* 1993, 21:253–258
29. The European Polycystic Kidney Disease Consortium. The polycystic kidney disease 1 gene encodes a 14 kb transcript and lies within a duplicated region on chromosome 16. *Cell* 1994, 78:725
30. Huang RP, Fan Y, de Belle I, Niemeyer C, Gottardis MM, Mercola D, Adamson ED: Decreased Egr-1 expression in human, mouse and rat mammary cells and tissues correlates with tumor formation. *Int J Cancer* 1997, 72:102–109
31. Niclas J, Navone F, Hom-Booher N, Vale RD: Cloning and localization of a conventional kinesin motor expressed exclusively in neurons. *Neuron* 1994, 12:1059–1072
32. Kamal A, Almenar-Queral A, LeBlanc JF, Roberts EA, Goldstein LS: Kinesin-mediated axonal transport of a membrane compartment containing beta-secretase and presenilin-1 requires APP. *Nature* 2001, 414:643–648
33. Tseng GC, Oh MK, Rohlin L, Liao JC, Wong WH: Issues in cDNA microarray analysis: quality filtering, channel normalization, models of variations and assessment of gene effects. *Nucleic Acids Res* 2001, 29:2549–2557
34. Wang X, Ghosh S, Guo SW: Quantitative quality control in microarray image processing and data acquisition. *Nucleic Acids Res* 2001, 29:E75–E75
35. Cauli B, Audinat E, Lambolez B, Angulo MC, Ropert N, Tsuzuki K, Hestrin S, Rossier J: Molecular and physiological diversity of cortical nonpyramidal cells. *J Neurosci* 1997, 17:3894–3906
36. Dickson DW: Neuropathology of Alzheimer's disease and other dementias. *Clin Geriatr Med* 2001, 17:209–228
37. Perl DP: Neuropathology of Alzheimer's disease and related disorders. *Neurol Clin* 2000, 18:847–864
38. Braak E, Griffling K, Arai K, Bohl J, Bratzke H, Braak H: Neuropathology of Alzheimer's disease: what is new since A. Alzheimer? *Eur Arch Psychiatry Clin Neurosci* 1999, 249:14–22
39. Small DH, Mok SS, Bornstein JC: Alzheimer's disease and Abeta toxicity: from top to bottom. *Nat Rev Neurosci* 2001, 2:595–598
40. DeKosky ST, Scheff SW: Synapse loss in frontal cortex biopsies in Alzheimer's disease: correlation with cognitive severity. *Ann Neurol* 1990, 27:457–464
41. Buell SJ, Coleman PD: Dendritic growth in the aged human brain and failure of growth in senile dementia. *Science* 1979, 206:854–856
42. Simeone A, Mavilio F, Acampora D, Giampaolo A, Faiella A, Zappavigna V, D'Esposito M, Pannese M, Russo G, Boncinelli E, Peschle C: Two human homeobox genes, c1 and c8: structure analysis and expression in embryonic development. *Proc Natl Acad Sci USA* 1987, 84:4914–4918
43. Lill MC, Fuller JF, Herzig R, Crooks GM, Gasson JC: The role of the homeobox gene, HOX B7, in human myelomonocytic differentiation. *Blood* 1995, 85:692–697
44. Yaron Y, McAdara JK, Lynch M, Hughes E, Gasson JC: Identification of novel functional regions important for the activity of HOXB7 in mammalian cells. *J Immunol* 2001, 166:5058–5067
45. van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH: Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002, 415:530–536
46. Pomeroy SL, Tamayo P, Gaasenbeek M, Sturla LM, Angelo M, McLaughlin ME, Kim JY, Goumnerova LC, Black PM, Lau C, Allen JC, Zagzag D, Olson JM, Curran T, Wetmore C, Biegel JA, Poggio T, Mukherjee S, Rifkin R, Califano A, Stolovitzky G, Louis DN, Mesirov JP, Lander ES, Golub TR: Prediction of central nervous system embryonal tumour outcome based on gene expression. *Nature* 2002, 415:436–442
47. Dhanasekaran SM, Barrette TR, Ghosh D, Shah R, Varambally S, Kurachi K, Pienta KJ, Rubin MA, Chinnaiyan AM: Delineation of prognostic biomarkers in prostate cancer. *Nature* 2001, 412:822–826
48. Tanzi RE, Bertram L: New frontiers in Alzheimer's disease genetics. *Neuron* 2001, 32:181–184