Neurobiology of Aging 34 (2013) 1825–1836

Contents lists available at SciVerse ScienceDirect

Neurobiology of Aging

journal homepage: www.elsevier.com/locate/neuaging

# Regional differences in gene expression and promoter usage in aged human brains

Luba M. Pardo [a,1], Patrizia Rizzu [a,1], Margherita Francescatto [a,b], Morana Vitezic [c,d], Gwenaël G.R. Leday [e], Javier Simon Sanchez [a], Abdullah Khamis [f], Hazuki Takahashi [c], Wilma D.J. van de Berg [g], Yulia A. Medvedeva [f], Mark A. van de Wiel [h], Carsten O. Daub [c], Piero Carninci [c], Peter Heutink [a,c,*]

[a] Section Medical Genomics, Department of Clinical Genetics, VU University Medical Center, Amsterdam, The Netherlands
[b] GABBA Program, Instituto de Ciências Biomédicas Abel Salazar, UP, Porto, Portugal
[c] RIKEN Omics Science Center, RIKEN Yokohama Institute, Yokohama, Japan
[d] Department of Cell and Molecular Biology (CMB), Karolinska Institute, Stockholm, Sweden
[e] Department of Mathematics, VU University, Amsterdam, The Netherlands
[f] Computational Bioscience Research Center, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia
[g] Department of Anatomy and Neurosciences, Section Functional Neuroanatomy, Neuroscience Campus Amsterdam, VU University Medical Center, Amsterdam, The Netherlands
[h] Department of Epidemiology and Biostatistics, VU University Medical Center, Amsterdam, The Netherlands

ARTICLE INFO

*Article history:*
Received 27 August 2012
Received in revised form 29 November 2012
Accepted 7 January 2013
Available online 19 February 2013

*Keywords:*
CAGE
Brain transcriptome
Aging

ABSTRACT

To characterize the promoterome of caudate and putamen regions (striatum), frontal and temporal cortices, and hippocampi from aged human brains, we used high-throughput cap analysis of gene expression to profile the transcription start sites and to quantify the differences in gene expression across the 5 brain regions. We also analyzed the extent to which methylation influenced the observed expression profiles. We sequenced more than 71 million cap analysis of gene expression tags corresponding to 70,202 promoter regions and 16,888 genes. More than 7000 transcripts were differentially expressed, mainly because of differential alternative promoter usage. Unexpectedly, 7% of differentially expressed genes were neurodevelopmental transcription factors. Functional pathway analysis on the differentially expressed genes revealed an overrepresentation of several signaling pathways (e.g., fibroblast growth factor and *wnt* signaling) in hippocampus and striatum. We also found that although 73% of methylation signals mapped within genes, the influence of methylation on the expression profile was small. Our study underscores alternative promoter usage as an important mechanism for determining the regional differences in gene expression at old age.



## 1. Introduction

The brain is the most complex organ of the human body, and this complexity is a major landmark of human evolution (Konopka and Geschwind, 2010). The brain can be divided into different functional and anatomic regions that are established during development and maintained throughout life. The mechanisms that regulate normal brain function and differentiation are controlled by both genetic (Johnson et al., 2009) and epigenetic factors (Miller and Sweatt, 2007), and alterations in these mechanisms can lead to neurodegenerative diseases (Abdolmaleky et al., 2005). There have been tremendous advances in our understanding of the molecular mechanisms involved in brain function, and the regional differences in these functions are beginning to be understood (Khaitovich et al., 2004; Roth et al., 2006). Less is known about the genetic mechanisms that are responsible for establishing and maintaining these differences throughout development, adulthood, and aging. Insights into these mechanisms are required to understand the differential susceptibility of distinct brain regions to neuronal insults (Double et al., 2010). For example, the genes for which mutations have been characterized in Alzheimer's disease (AD) (Joachim et al., 1989; Shen et al., 1997) and Parkinson's disease (PD) (Bandopadhyay et al., 2004) are often ubiquitously expressed whereas the observed pathology is restricted to specific brain regions and specific cell types (Double et al., 2010). Dissection of the molecular basis of this selective vulnerability will be pivotal to our understanding of disease pathogenesis and the development of specific therapies.

Much of our current insight into the molecular basis of brain function results from detailed studies of single genes or

* Corresponding author at: Department of Clinical Genetics, VU University Medical Center, van der Boechorststraat, 71081 BT Amsterdam, The Netherlands. Tel.: +31-205989962; fax: +31-2059983596.
*E-mail address:* p.heutink@vumc.nl (P. Heutink).
[1] These authors contributed equally to this work.

0197-4580/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neurobiolaging.2013.01.005

molecular mechanisms often initiated by the identification of genetic mutations (Hardy and Selkoe, 2002). However, unbiased approaches, where large numbers of genes are assessed simultaneously, are expected to be more powerful to dissect the genetic mechanisms controlling brain function. Large-scale analysis of gene expression in brain was pioneered by microarray experiments (Khaitovich et al., 2004). In recent years, high-throughput sequence-based technologies have been developed to analyze the mammalian transcriptome in more detail and at greater depth (Sandelin et al., 2007). These technologies have been decisive to uncover a complex picture of the mammalian transcriptome (Carninci et al., 2005) and to identify new mechanisms of gene regulation and control of gene expression in brain (Kang et al., 2011; Tollervey et al., 2011). Among sequence-based technologies, tag-based approaches such as cap analysis of gene expression (CAGE) have been used to comprehensively profile the transcription start sites (TSSs) and the promoter regions (Takahashi et al., 2012). CAGE is a cap-trapping–based method that profiles 5′ capped transcripts of both coding and noncoding RNA classes and has been pivotal in the discovery of alternatively regulated TSSs and novel regulatory elements (Carninci et al., 2006; Valen et al., 2009).

To understand how different promoters and control elements of genes establish and maintain region-specific expression patterns, we used CAGE in combination with massive parallel sequencing to profile TSSs of brain regions in 7 aged healthy individuals, at a genome-wide scale. We selected 5 samples of caudate nuclei, putamen, frontal and temporal cortices, and hippocampus, which are specifically vulnerable in the most prevalent neurodegenerative disorders (Double et al., 1996). First, we characterized the transcriptome of aged human brain and evaluated the extent of alternative promoter usage. Second, we quantified differences in gene expression and promoter usage across 5 brain regions. Finally, we analyzed the extent to which methylation influenced the observed expression profiles.

## 2. Methods

### 2.1. Brain specimens

The postmortem brain tissues were obtained from the Netherlands Brain Bank (Amsterdam, The Netherlands). The donors were aged subjects (age range: 70–91 years) without clinical signs of neurodegenerative or psychiatric disorders. All brains were neuropathologically evaluated by an experienced neuropathologist and classified for neurofibrillary tangles stage 0–VI (Alafuzoff et al., 2008), amyloid-beta plaques score 0–C, and Braak α-synuclein stage 0–VI using the staging protocols of Brain Net Europe and Braak (Alafuzoff et al., 2009a, 2009b; Braak et al., 2006). The dissection of the caudate, putamen, hippocampus, middle frontal gyrus (F2), and middle temporal gyrus regions was performed from snap frozen human brain sections. Tissue was stored at −80 °C until further processing. Pathologic examination of the brain specimens showed changes consistent with the age of the individuals. The age at death, cause of death, and postmortem delay until dissection are provided in Supplementary Table 1.

### 2.2. CAGE library preparation

Total RNA was extracted and purified from tissues using the Trizol tissue kit according to the instructions provided by the manufacturer (Invitrogen). RNA quality per library was assessed using the RNA integrity number with the Agilent Total RNA Nano kit (Table 1). The standard CAGE protocol (Kodzius et al., 2006) was adapted for sequencing on an Illumina platform. A thorough description of the protocol to prepare CAGE libraries and to sequence CAGE tags is presented in Takahashi et al. (2012). Briefly, complementary DNA (cDNA) was synthesized from total RNA using random primers, and this process was carried out at high temperature in the presence of trehalose and sorbitol to extend

**Table 1**
Description of the tag counts per region/sample

| Individual | Region | Batch[a] | RIN | Tag counts[b] | Unique counts[c] | Mapping rate[d] | Ribosome mapping[e] |
|---|---|---|---|---|---|---|---|
| **1** | **Caudate** | **1** | 7.6 | 1,988,794 | 935,084 | 0.856 | 0.062 |
| **1** | **Frontal** | **1** | 7 | 3,453,682 | 1,531,751 | 0.866 | 0.049 |
| **1** | **Hippocampus** | **1** | 6.5 | 2,022,640 | 979,162 | 0.811 | 0.09 |
| **1** | **Putamen** | **1** | 7.7 | 3,814,753 | 1,627,659 | 0.826 | 0.069 |
| **1** | **Temporal** | **1** | 6.3 | 4,333,255 | 1,937,270 | 0.822 | 0.07 |
| **2** | **Hippocampus** | **1** | 6.5 | 1,682,943 | 310,481 | 0.843 | 0.072 |
| **2** | **Caudate** | **2** | 7.2 | 1,663,688 | 362,468 | 0.724 | 0.088 |
| **2** | **Frontal** | **2** | 6.9 | 1,745,155 | 801,757 | 0.822 | 0.04 |
| **2** | **Putamen** | **2** | 6.5 | 1,216,441 | 274,776 | 0.702 | 0.113 |
| **2** | **Temporal** | **2** | 6.8 | 936,396 | 259,968 | 0.748 | 0.103 |
| **3** | **Frontal** | **2** | 7.1 | 2,111,277 | 505,207 | 0.779 | 0.068 |
| **3** | **Hippocampus** | **2** | 8.8 | 1,785,386 | 413,336 | 0.816 | 0.041 |
| **3** | **Temporal** | **2** | 6.8 | 1,103,935 | 255,621 | 0.84 | 0.041 |
| **4** | **Temporal** | **2** | 5.9 | 1,199,974 | 356,840 | 0.71 | 0.127 |
| **4** | **Frontal** | **2** | 6.5 | 2,035,347 | 472,327 | 0.739 | 0.107 |
| **4** | **Hippocampus** | **2** | 6.4 | 1,251,589 | 335,644 | 0.731 | 0.109 |
| **4** | **Putamen** | **2** | 6.5 | 2,541,166 | 516,842 | 0.73 | 0.121 |
| **5** | **Caudate** | **1** | 7.9 | 3,096,524 | 1,144,105 | 0.875 | 0.059 |
| **5** | **Putamen** | **1** | 6.6 | 4,029,122 | 1,541,543 | 0.834 | 0.082 |
| **6** | **Caudate** | **1** | 7.4 | 3,587,220 | 1,296,765 | 0.875 | 0.053 |
| **6** | **Putamen** | **1** | 6.3 | 2,085,385 | 795,569 | 0.868 | 0.072 |
| **7** | **Caudate** | **1** | 6.8 | 4,875,578 | 1,625,317 | 0.862 | 0.062 |
| **7** | **Frontal** | **2** | 6.2 | 2,324,932 | 407,993 | 0.731 | 0.111 |
| **7** | **Hippocampus** | **2** | 6.2 | 3,158,604 | 597,669 | 0.775 | 0.033 |
| **7** | **Temporal** | **2** | 6.2 | 1,104,711 | 241,508 | 0.699 | 0.157 |

Details on the quality control and final counts used for the analysis are presented in Supplementary data. Individual, region and batch id are presented in bold.
Key: RIN, RNA integrity number.
[a] Refers to 2 main batch effects corresponding to different period of times in which the cap analysis of gene expression libraries were prepared (Supplementary data).
[b] Refers to the total tag counts after removal of sequencing artifacts.
[c] Refers to the tag counts that map to single positions in the genome unique regions.
[d] Refers to proportion of tags that mapped to less than 10 positions.
[e] Refers to the proportion of tags that mapped to ribosomal DNA.