European Journal of Pharmacology 759 (2015) 101–117

Contents lists available at ScienceDirect

European Journal of Pharmacology

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

Mouse models rarely mimic the transcriptome of human neurodegenerative diseases: A systematic bioinformatics-based critique of preclinical models

Terry C. Burns*, Matthew D. Li, Swapnil Mehta, Ahmed J. Awad, Alexander A. Morgan

Department of Neurosurgery, Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, 300 Pasteur Drive, Stanford, CA 94305, USA

ARTICLE INFO

Article history:
Received 10 February 2015
Received in revised form
12 March 2015
Accepted 12 March 2015
Available online 23 March 2015

Keywords:
Huntington's disease
Parkinson's disease
Alzheimer's disease
Amyotrophic lateral sclerosis
Neurodegeneration
Bioinformatics
Gene set enrichment

ABSTRACT

Translational research for neurodegenerative disease depends intimately upon animal models. Unfortunately, promising therapies developed using mouse models mostly fail in clinical trials, highlighting uncertainty about how well mouse models mimic human neurodegenerative disease at the molecular level. We compared the transcriptional signature of neurodegeneration in mouse models of Alzheimer's disease (AD), Parkinson's disease (PD), Huntington's disease (HD) and amyotrophic lateral sclerosis (ALS) to human disease. In contrast to aging, which demonstrated a conserved transcriptome between humans and mice, only 3 of 19 animal models showed significant enrichment for gene sets comprising the most dysregulated up- and down-regulated human genes. Spearman's correlation analysis revealed even healthy human aging to be more closely related to human neurodegeneration than any mouse model of AD, PD, ALS or HD. Remarkably, mouse models frequently upregulated stress response genes that were consistently downregulated in human diseases. Among potential alternate models of neurodegeneration, mouse prion disease outperformed all other disease-specific models. Even among the best available animal models, conserved differences between mouse and human transcriptomes were found across multiple animal model versus human disease comparisons, surprisingly, even including aging. Relative to mouse models, mouse disease signatures demonstrated consistent trends toward preserved mitochondrial function protein catabolism, DNA repair responses, and chromatin maintenance. These findings suggest a more complex and multifactorial pathophysiology in human neurodegeneration than is captured through standard animal models, and suggest that even among conserved physiological processes such as aging, mice are less prone to exhibit neurodegeneration-like changes. This work may help explain the poor track record of mouse-based translational therapies for neurodegeneration and provides a path forward to critically evaluate and improve animal models of human disease.

1. Introduction

Animal models provide a critical platform upon which translational efforts for treating human neurodegenerative diseases are built. While there is no substitute for studying true human biology, animal models provide opportunities for experimentation that are often impossible in human patients. Transgenic animals carrying human mutations provide an opportunity to understand mechanisms underlying human disease pathogenesis. Moreover, animal models routinely serve as gatekeepers to putative therapies being considered for clinical trials (Burns and Verfaillie, this issue). Unprecedented progress has been made in the past two decades based in part on animal models of Huntington's disease (HD), amyotrophic lateral sclerosis (ALS), Parkinson's disease (PD) and Alzheimer's disease (AD). With this new understanding, hundreds of pharmaceutical agents that have shown promise in preclinical animal models of neurodegenerative disease have progressed to clinical trials. Unfortunately, almost none have proven effective in humans.

This stark reality has prompted a thoughtful re-evaluation of the role of mouse models of neurodegeneration and neuroinflammation (Cavanaugh et al., 2014; Doody et al., 2014; Gladstone et al., 2002; O'Collins et al., 2006; Panza et al., 2014; Scott et al., 2008). Comparison of transcriptome data between experiments is made challenging by the wide variety of methodologies employed for comprehensive transcriptome analysis. Even identical protocols routinely yield incomparable results between experiments due to batch effects. Comparisons between species are made even more challenging by a lack of clear homologs for many genes, species-specific differences in the function of certain genes and timing of

* Corresponding author. Tel.: +1 612 812 7223.
E-mail addresses: tcburns@stanford.edu, burn0280@gmail.com (T.C. Burns).

http://dx.doi.org/10.1016/j.ejphar.2015.03.021
0014-2999/

cellular responses to stimuli, as well as a lack of standardized methodological approaches to inter-species comparisons. Nevertheless, continually improved annotations and the development of robust bioinformatics techniques now permit meaningful comparisons of transcriptional responses between species.

We recently took advantage of the rapidly expanding inventory of transcriptional profiles for human neurodegenerative diseases to perform a meta-analysis focusing on AD, PD, HD, and ALS (Li et al., 2014). This work identified a common neurodegenerative disease module that is shared across human neurodegenerative diseases. Given that transcriptome data additionally exist for several animal models of neurodegenerative diseases, we sought here to address the following questions:

1) Can transcriptional responses between species be meaningfully compared, using the relatively conserved aging process as a positive control?
2) Are the transcriptional signatures of human neurodegenerative diseases appropriately reflected in animal models?
3) Do variables such as disease stage or brain region analyzed substantially confound the results of comparative analyses?
4) Could alternate mouse models exist that more closely mimic human neurodegeneration than standard models of AD, PD, ALS and HD?
5) Do reproducible differences exist between human neurodegeneration and mouse models across multiple diseases?

2. Materials and methods

2.1. Human neurodegenerative disease meta-signatures

We previously performed a meta-analysis of human neurodegenerative diseases (Li et al., 2014). We used the individual ranked gene lists for each disease generated in the course of our prior analysis as a baseline against which to compare the transcriptome of the corresponding animal models. The total ranked gene lists for each disease are provided in Supplemental Table S1. In the "discovery" portion of our prior meta-analysis, the effect sizes for each individual AD, PD, HD and ALS data set were then combined to determine the pooled effect size for human neurodegeneration using the random effects inverse-variance technique. The resultant ranked gene lists are provided in Supplemental Table S1.

2.2. Human aging signatures

We previously identified 3 human aging brain data sets each including at least 30 patients: E-GEOD 30272, E-GEOD 11882, and E-GEOD 1572 (Li et al., 2014). For each gene in the neurodegeneration data sets, the Kendall tau coefficient between the log2 transformed gene signal intensity and age was determined using the "Kendall" R package. The resulting ranked gene lists are provided in Supplemental Table S1.

2.3. Animal models of neurodegeneration

We searched the public data repository ArrayExpress (November 2014) for gene expression microarray data sets from mouse models of neurodegenerative disease using search terms "neurodegeneration," "Alzheimer," "Parkinson," "Huntington," and "amyotrophic lateral sclerosis." Additionally, mouse models identified from review of relevant literature were included as well as select studies of mouse brain aging. Available processed data sets were included if they met the following criteria: (1) samples were from mouse CNS tissue samples, including any region of forebrain for AD and HD, midbrain or forebrain samples for PD, and spinal cord samples for ALS; and (2) the microarray platform had accessible probe-to-gene mapping annotations available for use through the GSEA Java application v2.0.14. Data sets with <3 mice per group were excluded if another sufficient data set was available for that disease/model at a comparable or later time point. When multiple time points were available, we included only the latest available time point, with the exception of E-GEOD-31372, E-GEOD-4390, and E-MEXP-453, for which each of the 2–3 available time points were evaluated in parallel. All samples included were from fresh whole tissue, with the exception of E-MEXP-453, a study of motor neurons purified by single-cell laser capture. When data for multiple brain regions were available, we selected one or two brain regions most relevant to the disease pathophysiology. The ArrayExpress identifiers (e.g. E-GEOD, E-MEXP) for each study evaluated are included in tables where relevant throughout the manuscript. Table 1 provides the complete list of data sets utilized for this study. These were compared to the human neurodegenerative disease meta-analysis ranked gene lists, as provided in Supplemental Table S1.

2.4. Enrichment for human disease signatures in animal models

Gene set enrichment analysis is a well-established technique for comparing genomic responses between independent samples, diseases and species (Yu et al., 2011). We used Gene Set Enrichment Analysis (GSEA) (Subramanian et al., 2005) to evaluate for the enrichment of (1) human neurodegenerative disease signatures in human neurodegenerative disease ranked gene lists; (2) human aging signatures in mouse aging samples; and (3) human neurodegenerative disease signatures in mouse models of neurodegeneration. Additionally, to evaluate for alterations in enrichment for functional gene modules and cell-type signatures, sets of genes found in the literature to be co-regulated in healthy (Oldham et al., 2008) and AD brain samples (Zhang et al., 2013) were included, along with gene sets representing the most highly differentially expressed genes in specific prospectively isolated CNS cell types as well as gliosis and microglial activation. GSEA was performed using default settings, including 1000 permutations based on gene set. The Pre-rank tool was used for GSEA based upon ranked gene lists. Analysis was based upon probe set IDs from the processed public data sets, or gene symbol for ranked gene lists. False discovery rates (FDR) <0.05 were considered significant. Gene sets employed for our analysis are provided in Supplemental Table S2.

2.5. Spearman's rank correlation

Spearman's rank correlation analysis of the mouse and human gene expression datasets was performed using R/Bioconductor. For every ranked gene list, genes with duplicate entries were omitted (i.e., probe sets mapping to multiple genes). Using the 1315 genes that were in common across all 51 lists, we calculated Spearman's rank correlation for each pairwise comparison between ranked gene lists. Heat maps featuring hierarchical clustering (Euclidean distance) of Spearman's rank correlations were generated using the pheatmap R package.

2.6. Comparison of mouse and human neurodegenerative diseases signatures

GSEA is based in part upon the generation of a ranked gene list with the most highly upregulated genes at the top and the most highly downregulated genes at the bottom of the ranked gene list. The order of genes in the list thereby summarizes the transcriptome for given condition. In order to determine which genes were differentially regulated in response to human disease versus mouse models of disease, we compared the percentile rank position of each gene in the ranked gene lists for mouse models and human diseases.