Genes & Diseases (2016) 3, 228–237

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: http://ees.elsevier.com/gendis/default.asp

FULL LENGTH ARTICLE

Computational identification and analysis of neurodegenerative disease associated protein kinases in hominid genomes

Saranya Jayapalan, Devika Subramanian, Jeyakumar Natarajan*

Data Mining and Text Mining Lab, Department of Bioinformatics, Bharathiar University, Coimbatore, India

Received 12 February 2016; accepted 15 April 2016
Available online 23 April 2016

KEYWORDS
Hominids;
Interaction analysis;
Neurodegenerative diseases;
Phosphorylation sites;
Protein kinases

Abstract Protein kinases play an important role in the incidence of neurodegenerative diseases. However their incidence in non-human primates is found to be very low. Small differences among the genomes might influence the disease susceptibilities. The present study deals with finding the genetic differences of protein kinases in humans and their three closest evolutionary partners chimpanzee, gorilla and orangutan for three neurodegenerative diseases namely, Alzheimer's, Parkinson's and Huntington's diseases. In total 47 human protein kinases associated with three neurodegenerative diseases and their orthologs from other three non-human primates were identified and analyzed for any possible susceptibility factors in humans. Multiple sequence alignment and pairwise sequence alignment revealed that, 18 human protein kinases including DYRK1A, RPS6KB1, and GRK6 contained significant indels and substitutions. Further phosphorylation site analysis revealed that eight kinases including MARK2 and LTK contained sites of phosphorylation exclusive to human genomes which could be particular candidates in determining disease susceptibility between human and non-human primates. Final pathway analysis of these eight kinases and their targets revealed that these kinases could have long range consequences in important signaling pathways which are associated with neurodegenerative diseases.

* Corresponding author. Department of Bioinformatics, Bharathiar University, Coimbatore, India.
E-mail addresses: saranya.jgs@gmail.com (S. Jayapalan), devikaparvathy@gmail.com (D. Subramanian), n.jeyakumar@yahoo.co.in (J. Natarajan).
Peer review under responsibility of Chongqing Medical University.

http://dx.doi.org/10.1016/j.gendis.2016.04.004

List of abbreviations

PK	protein kinase
HGNC	Human Genome Nomenclature Committee
GO	gene ontology
OMIM	online mendelian inheritance in man
MSA	multiple sequence alignment
HPK	human protein kinase
NHP	non-human primates

Introduction

Neurodegeneration is referred to as the progressive loss of structure and function of neurons. Neurodegenerative diseases constitute one of the major challenges of modern medicine, including Alzheimer's disease, Parkinson's disease, Huntington's disease, Pick's disease etc.[1,2] Alzheimer's disease is characterized by loss of neurons and synapses in the cerebral cortex region and subcortical regions. This loss results in gross atrophy of the affected regions, including degeneration in the temporal lobe and parietal lobe, and parts of the frontal cortex and cingulate gyrus.[3–6] Parkinson's disease is a degenerative disorder of the central nervous system, the mechanism by which the brain cells in Parkinson's are lost may consist of an abnormal accumulation of the protein alpha-synuclein bound to ubiquitin in the damaged cells.[7–9] The causes of Huntington's disease are astrogliosis and loss of medium spiny neurons. The areas affected are mainly in the striatum, but also the frontal and temporal cortices.[10]

Many neurodegenerative diseases are caused by genetic mutations, most of which are completely located in unrelated genes. Identification of disease causing genes is one of the major challenges in the human genome studies. The use of linkage analysis and cloning techniques has led to the detection of genes involved in abundant Mendelian genetic disorders.[1,11]

Protein kinases are the most common protein domains implicated in neurodegenerative diseases.[12] Kinome is a division of genome that consists of protein kinase genes. Protein kinases are major regulatory enzymes that participate in the process of protein phosphorylation. It is an essential process in many cellular and signal transduction processes.[13] Protein kinases (PK's) act as key regulators of cell functions by catalyzing the addition of a negatively charged phosphate group to proteins. Because of its role in every aspect of regulation and signal transduction, they provide new targets for drug development.[14,15]

Human (*Homo sapiens*) share maximum sequence similarity with their closest ancestral species chimpanzee (*Pan troglodytes*), gorilla (*Gorilla gorilla*) and orangutan (*Pongo abelii*) genomes.[16] Comparing the human genome to the genomes of other great apes can provide a window into the molecular changes that causes this difference in disease susceptibility between human and non-human primates. The sequence divergence between the human, and chimpanzee has been a subject of numerous studies.[17–19] Human brain aging equally exhibits similarities and differences with great apes. The amyloid-β protein deposits are seen in hippocampus and neocortex of aged chimpanzees, gorillas, and orangutans in the form of diffuse plaques and vascular lesions similar to humans.[20,21] But gene expression changes in the neocortex during aging in human makes them more susceptible to age related neurodegenerative diseases.[22] Even though it is widely considered that they are exclusive to the human species, various studies have shown the existence of some neurodegenerative diseases in senile non-human primates. They are found to be present with some features including neuropathological changes and cognitive-behavioral symptoms, but with lower levels of incidence.[22,23]

The discovery of gene mutations dominates our understanding of neurodegenerative diseases. However they have a more complex etiology influenced by lifestyle and environmental factors in addition to a number of uncharacterized variants in a high number of genes. An important pathological hallmark of ND diseases is usually the accumulation and aggregation of misfolded proteins as with amyloid-β (Aβ) in Alzheimer's disease (AD),[24,25] α-synuclein in Parkinson's disease (PD)[26] and huntingtin protein in Huntington's disease (HD).[27] Recently biological pathways and networks have become focal points to examine the genetic architecture of neurodegeneration. Vijay K Ramanan and Andrew J Saykin[28] reviewed the major pathways implicated in neurodegeneration.

The present study compares the key protein kinase sequences of humans, chimpanzee, gorilla and orangutan and tries to analyze the kinase level evolutionary relationships among the four lineages. The study of disease susceptibility genes is an emerging task in the field of genetic research that could lead to fruitful information.

Materials and methods

Hominid protein kinase data sets

In our recent publication,[29] we reported in detail about the collection of hominid protein kinase data sets. The same is briefly outlined below. Human protein kinases were directly taken from our in-built database Homo-Kinase (http://www.biominingbu.org/homokinase).[30] For the other three species — Chimpanzee, Gorilla and Orangutan, the proteomes were retrieved from Uniprot data source (http://www.uniprot.org/) and their functional annotation were checked for the presence of the following three GO terms i) ATP binding, ii) kinase activity, and iii) protein phosphorylation. Protein with above three GO annotation terms were classified as true protein kinases and used for further analysis. In total, 499 human, 478 chimpanzee, 468 gorilla and 470 orangutan protein kinases were used.

Identification and classification of neurodegenerative disease associated protein kinases

From the above human protein kinase data set, protein kinase genes associated with the three target diseases (Alzheimer's, Parkinson's and Huntington's) were identified and grouped using OMIM (www.ncbi.nlm.nih.gov/Omim)[31]