Available online at www.sciencedirect.com

ScienceDirect

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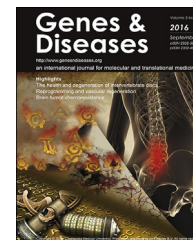
FULL LENGTH ARTICLE

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

CrossMark

Saranya Jayapalan, Devika Subramanian, Jeyakumar Natarajan*

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

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.

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 HomoKinase (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]

and Wikigenes (http://www.wikigenes.org/).[32] For this, we used MeSH terms associated with the three diseases (Alzheimer's disease for AD, Parkinson's disease for PD and Huntington's disease for HD) to extract all literature cited in OMIM. In addition, the set of human neurodegenerative PKs were also selected from the literature based on mutational analysis and/or roles in processes such as chromosomal stability, promoter methylation etc. From the total of 499 human protein kinases, 47 protein kinases were found to be associated with neurodegenerative diseases and subjected to further analysis. Similarly the orthologous protein kinases present in other three species chimpanzee, gorilla and orangutan for the above 47 human protein kinases were retrieved from their protein kinase data sets.

Sequence analysis of orthologous protein kinases

Both Multiple sequence alignment (MSA) and pairwise sequence analysis were performed between all 47 selected human protein kinases and their orthologous sequence from other three genomes to find indels and substitutions between the human PK's and other orthologs. Multiple sequence alignment (MSA) for all four genomes was performed using ClustalX 2.1[33] to find the similarities and dissimilarities between the orthologous kinases at protein sequence level. Further, any sequence characteristics specific to human kinases were also examined for all the target disease associated protein kinases.

In addition pairwise sequence alignments between human and other hominid PK's were performed using EMBOSS Needle (http://www.sanger.ac.uk/Software/EMBOSS/)[34] program to find the genetic differences between seven key protein kinases which were associated with more than one of the target diseases.

Phosphorylation site analysis

The regulation of protein kinase activity is brought about by various mechanisms in which phosphorylation also plays a very relevant role. To bring about more insights into this, the protein phosphorylation sites of a few chosen kinases with significant results in MSA were predicted using NetPhos 2.0 server[35] to identify Ser/Thr and Tyr phosphorylation sites.

Interaction analysis of selected protein kinases in humans

In order to identify additional functional relationships between the AD, PD, and HD associated kinases and to obtain a deeper understanding of the consequences of the kinases with phosphorylation sites predicted exclusively in humans, we analyzed their participation in human signaling pathways. Our aim was to examine the participation of the kinases and their targets in neurodegenerative diseases and important signaling pathways such as EGF/MAPK, neurotrophin and Notch signaling pathways. Experimentally proven targets of these kinases were obtained from STRING (version 8.3),[36] HPRD (version 5)[37] and BioGrid (version 3.0.64)[38] databases. The gene list of kinases and their targets were then analyzed for overrepresentation of disease pathways and important signaling pathways using the online tool WebGeStalt.[39] Enrichment analysis was carried out for KEGG pathways,[40] Pathway commons,[41] Wikipathways[42] and Disease association (from PharmGKB[43]) in WebGestalt.[39]

Results

A set of 499 human, 478 chimpanzee, 468 gorilla and 470 orangutan protein kinases was used for this kinase-disease association analysis. Then by literature based approaches and from OMIM[31] and Wikigenes[32] data bases we retrieved the human protein kinases associated with neurodegenerative diseases (Table 1). These neurodegenerative associated protein kinases were grouped into three categories:

(1) HPK's associated with Alzheimer's disease,
(2) HPK's associated with Parkinson's disease and
(3) HPK's associated with Huntington's disease.

A total number of 47 protein kinases associated with the three neurodegenerative diseases were identified and grouped and given in Table 1. Protein kinases associated with more than one disease are underlined. The Table 1 results revealed that 32 human protein kinases (HPKs) were found to be associated with Alzheimer's diseases, 13 HPKs with Parkinson's disease and 9 HPKs were associated with Huntington's disease. Further seven kinases (LRRK2, GSK3B, PINK1, MAPK14, FGFR3, PYK2 and R1PK2) of them being found to be associated with more than one disease. LRRK2, GSK3B, PINK1 and MAPK14 are commonly associated with both Alzheimer's disease and Parkinson's disease. FGFR3 is associated with both Alzheimer's disease and Huntington's disease whereas PYK2 and R1PK2 are associated with both Parkinson's disease and Huntington's disease. The associations have been represented in the form of a gene-disease network in Fig. 1.

For the above 47 human protein kinases, the orthologous protein kinases present in chimpanzee, gorilla and orangutan were identified and grouped using their respective protein kinase data sets. Table 2 gives the complete list of human protein kinases and their corresponding orthologs from other three species.

Table 1 Protein kinases associated with Alzheimer's, Parkinson's, and Huntington's diseases (underlined kinases have more than one disease association).

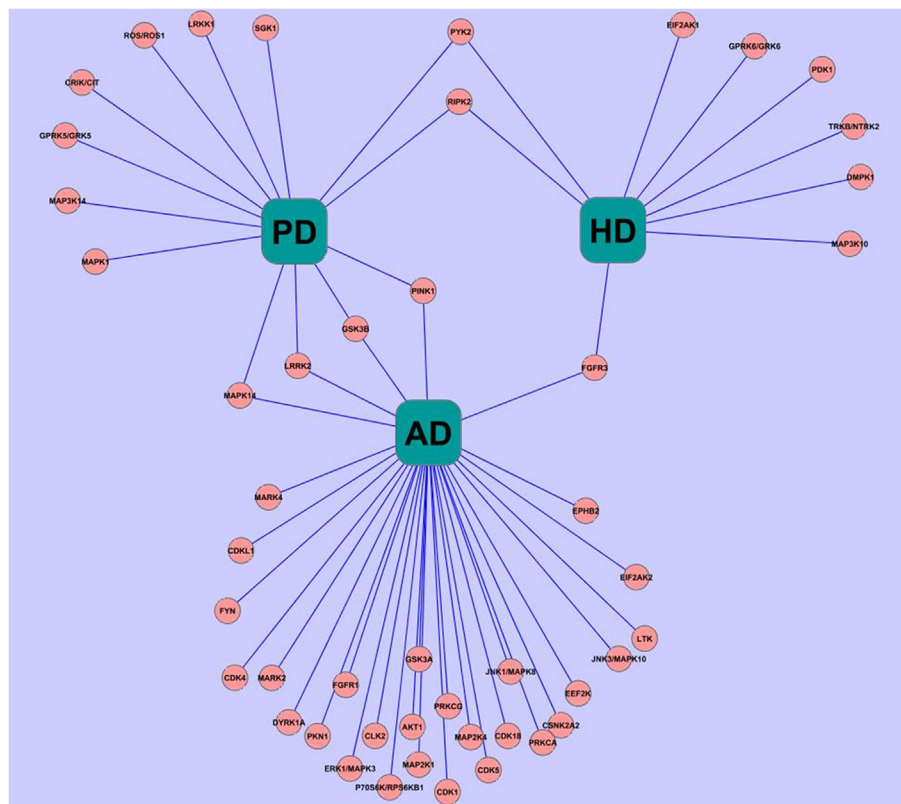| Diseases	Protein kinases
Alzheimer's	AKT1, CDK1, CDK4, CDK18, CDK5, CDKL1, CLK2, CSNK2A2, DYRK1A, EEF2K, EIF2AK2, ERK1, EPHB2, FGFR1, FGFR3, FYN, GSK3A, GSK3B, JNK1, JNK3, LRRK2, LTK, MAP2K1, MAP2K4, MAPK14, MARK2, MARK4, P70S6K, PINK1, PKN1, PRKCA, PRKCG
Parkinson's	CRIK, GPRK5, GSK3B, LRRK1, LRRK2, MAP3K14, MAPK1, MAPK14, PINK1, PYK2, R1PK2, ROS, SGK1
Huntington's	DMPK1, EIF2AK1, FGFR3, GPRK6, MAP3K10, PDK1, PYK2, RIPK2, TRKB

Fig. 1 Gene-disease network for Alzheimer's disease (AD), Parkinson's disease (PD), and Huntington's disease (HD).

Sequence analysis of orthologous protein kinases

Basic sequence analysis was performed to find the sequence identity similarities and dissimilarities between the orthologous protein kinases of human and non-human genomes. The pairwise sequence alignment of seven protein kinases namely LRRK2, GSK3B, PINK1, MAPK14, FGFR3, PYK2 and R1PK2 associated with more than one disease showed that all of them except GSK3B and MAPK14 shared greater than 95% similarity with each other (Table 3). Human GSK3B protein sequence was identical to chimpanzee sequence where as it shared 88.2% and 84.6% similarity with the gorilla and orangutan protein sequences, respectively (Table 3). Significant deviation from human sequence was observed for FGFR3 sequence of orangutan genome (70% similarity at protein level with 29.8% gaps) and MAPK14 sequence of gorilla genome (60.2% similarity at protein level with 34% gaps). Thus a preliminary analysis of key protein kinases involved in neurodegenerative diseases showed high amount of conservation within species.

ClustalX was used to build multiple sequence alignments of the selected protein kinases. Multiple sequence alignments of 47 protein kinases were carried out. For each human protein kinase under consideration, the protein sequences of the corresponding orthologous kinase from the other three genomes if present were taken for the alignment. The results revealed that some of the human kinases contained many significant indels and substitutions which are not seen in any of the other sequences. A few significant differences observed in three protein kinases DYRK1A, RPS6KB1, and GRK6 are mentioned below.

The protein kinase DYRK1A associated with Alzheimer's disease in humans contained a long insertion of 83 amino acids in human which was not present in any of the corresponding orthologs. DYRK1A sequence contained 763 amino acids in humans 754 amino acids in chimpanzee 535 amino acids in gorilla and 683 amino acids in orangutan. The orangutan DYRK1A sequence was mostly similar to the human DYRK1A except for this major deletion.

In the case of protein kinase RPS6KB1, also associated with Alzheimer's diseases there was a deletion of 22 amino acid residues in human (Fig. 2a). Terminal amino acid residues were identical in human and orangutan while they were absent in chimpanzee and gorilla. The sequences were similar at other positions of the alignment for all the four genomes.

The protein kinase GRK6 associated with Huntington's disease has 560 amino acids in human and 576 in gorilla and orangutan. The GRK6 sequence in chimpanzee consisted of 600 amino acid residues and was found to vary considerably from the other three GRK6 sequences. Among the other three gorilla and orangutan GRK6 contained a sequence segment of 17 amino acids long which was replaced by a single 'R' residue in human (Fig. 2b).

We also looked for other possible general features and variations among the protein kinases associated with each of the three diseases. One important observation was that in LTK human sequence had a low occurrence of Gly-repeats compared to the LTK's of other three organisms. The Gly repeats were either deleted in between or substituted with other residues. In addition many other deletions were also observed in LTK.

Table 2   Human neurodegenerative disease associated protein kinases and their equivalent orthologs in other hominids.

Disease	Human protein kinases	Chimpanzee	Gorilla	Orangutan
Alzheimer's	AKT1	AKT1	AKT1	—
	CDK1	CDK1	CDK1	CDK1
	CDK4	CDK4	CDK4	CDK4
	CDK5	CDK5	CDK5	CDK5
	CDK18	CDK18	CDK18	CDK18
	CDKL1	CDKL1	CDKL1	—
	CLK2	CLK2	CLK2	—
	CSNK2A2	CSNK2A2	CSNK2A2	CSNK2A2
	DYRK1A	DYRK1A	DYRK1A	DYRK1A
	EEF2K	EEF2K	EEF2K	EEF2K
	EIF2AK2	EIF2AK2	—	EIF2AK2
	EPHB2/ERK	EPHB2	EPHB2	EPHB2
	ERK1/MAPK3	ERK1	ERK1	ERK1
	FGFR1	FGFR1	FGFR1	FGFR1
	FGFR3	FGFR3	FGFR3	FGFR3
	FYN	FYN	FYN	FYN
	GSK3A	GSK3A	GSK3A	GSK3A
	GSK3B	GSK3B	GSK3B	GSK3B
	JNK1/MAPK8	JNK1	JNK1	JNK1
	JNK3/MAPK10	JNK3	JNK3	JNK3
	LRRK2	LRRK2	LRRK2	LRRK2
	LTK	LTK	LTK	LTK
	MAP2K1	MAP2K1	MAP2K1	MAP2K1
	MAP2K4	MAP2K4	MAP2K4	—
	MAPK14	MAPK14	MAPK14	MAPK14
	MARK2	—	MARK2	MARK2
	MARK4	MARK4	MARK4	MARK4
	P70S6K/RPS6KB1	P70S6K	P70S6K	P70S6K
	PINK1	PINK1	PINK1	PINK1
	PKN1	PKN1	PKN1	—
	PRKCA	PRKCA	PRKCA	—
	PRKCG	PRKCG	PRKCG	PRKCG
Parkinson's	CRIK/CIT	CRIK	CRIK	CRIK
	GPRK5	GPRK5	GPRK5	GPRK5
	GSK3B	GSK3B	GSK3B	GSK3B
	LRRK1	LRRK1	LRRK1	LRRK1
	LRRK2	LRRK2	LRRK2	LRRK2
	MAP3K14	MAP3K14	MAP3K14	MAP3K14
	MAPK1	MAPK1	—	MAPK1
	MAPK14	MAPK14	MAPK14	MAPK14
	PINK1	PINK1	PINK1	PINK1
	PYK2/PTK2B	PYK2	PYK2	PYK2
	R1PK2	R1PK2	R1PK2	R1PK2
	ROS	ROS	ROS	ROS
	SGK1	SGK1	SGK1	SGK1
Huntington's	DMPK1	DMPK1	DMPK1	DMPK1
	EIF2AK1	EIF2AK1	EIF2AK1	EIF2AK1
	FGFR3	FGFR3	FGFR3	FGFR3
	GPRK6	GPRK6	GPRK6	GPRK6
	MAP3K10	MAP3K10	MAP3K10	MAP3K10
	PDK1	PDK1	PDK1	PDK1
	PYK2	PYK2	PYK2	PYK2
	R1PK2	R1PK2	R1PK2	R1PK2
	TRKB	TRKB	TRKB	TRKB

Table 3 Protein sequence comparison results (in percentage) for the seven protein kinases associated with more than one disease. It showed high degree of conservation within species.

|	Human—Chimpanzee			Human—Gorilla			Human—Orangutan		
	Identity	Similarity	Gaps	Identity	Similarity	Gaps	Identity	Similarity	Gaps
GSK3B	100.0%	100.0%	0.0%	87.5%	88.2%	10.6%	83.6%	84.6%	12.6%
LRRK2	99.2%	99.6%	0.0%	97.3%	97.9%	1.5%	97.6%	98.7%	0.4%
PINK1	99.7%	100.0%	0.0%	99.7%	99.7%	0.0%	98.8%	99.3%	0.0%
MAPK14	100.0%	100.0%	0.0%	57.1%	60.2%	34.0%	99.7%	100.0%	0.0%
FGFR3	95.3%	96.4%	0.5%	96.7%	98.1%	0.1%	69.9%	70.0%	29.8%
PYK2	99.6%	99.8%	0.0%	99.6%	99.9%	0.0%	98.9%	99.5%	0.0%
R1PK2	99.3%	99.4%	0.2%	96.9%	97.8%	0.2%	98.9%	99.3%	0.0%

(a)

(b)
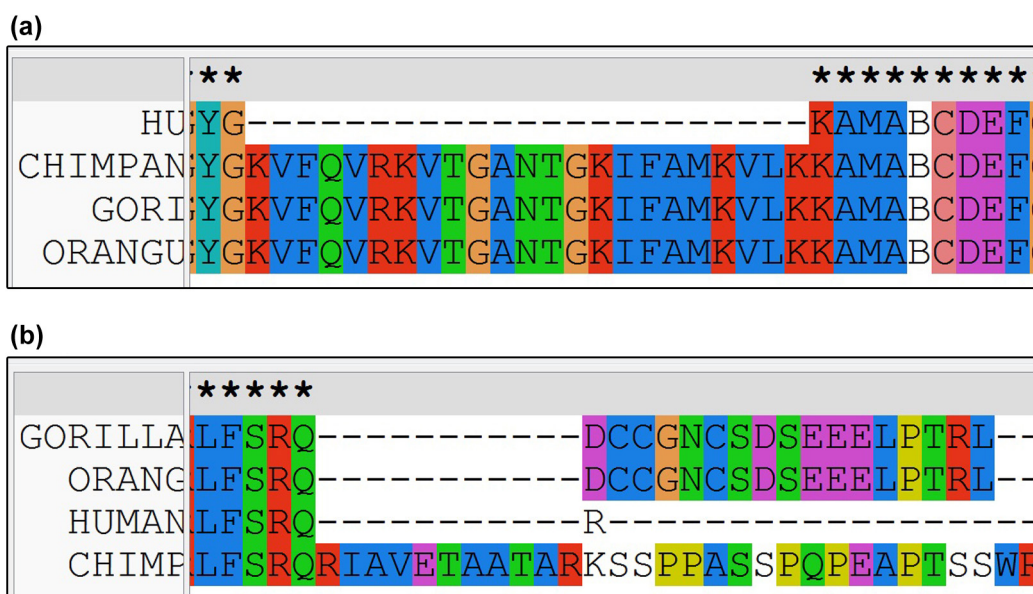
Fig. 2   a) Deletion of 22 amino acid residues in human RPS6KB1 sequence associated with Alzheimer's disease. **b)** Sequence variation in Huntington's disease associated GRK6 sequence.

Another observation was that 5 out of the 13 kinases associated with Parkinson's disease namely LRRK1, LRRK2, MAP3K14, PYK2 and ROS1 were characterized by the most number of single amino acid substitutions unlike the genes associated with other diseases.

In summary, the following 18 kinases were found to contain significant variations: CDK1, CDKL1, DYRK1A, EIF2AK2, GSK3A, MAPK8, MAPK10, LTK, MARK2, RPS6KB1, PRKCG (Alzheimer's disease associated), LRRK1, MAP3K14, PYK2, ROS1 (Parkinson's disease associated), TRKB, GRK6 and EIF2AK1 (Huntington disease associated). However a substantial proportion of differences may have no functional impact and hence it is important to ascertain the effect of each of these variations. The details are provided in Supplementary data SD1.

Phosphorylation site analysis of selected protein kinases

The varying regions in the protein kinases identified above were analyzed in depth for phosphorylations sites specific

to humans. 18 kinases found to contain significant indels/substitutions were further analyzed for phosphorylation sites in the regions seen only in human kinase sequences.

Table 4 Phosphorylation sites present in human protein kinase sequences.

S. No	Protein kinase	NetPhos score	Position and sequence segment containing the site for phosphorylation	Phosphorylation site for
01	CDKL1	0.866	273 VPIASRTE	*S*
02	DYRK1A	0.576	220 EMKYYIVHL	*Y*
03	GSK3A	0.800	458 LTPSSQALT	*S*
04	JNK3/ MAPK10	0.881	31 QVDVSYIAK	*S*
05	**LTK**	**0.995**	**251 RTQASPEKL**	***S****
06	**MARK2**	**0.981**	**648 RRNLSFRFA**	***S****
07	LRRK1	0.521	37 GAGDTGGKP	*T*
08	MAP3K14	0.592	292 ACVDSQKPL	*S*

CDKL1, DYRK1A, GSK3A, MAPK10, LTK, MARK2, LRRK1 and MAP3K14 contained specific sites of phosphorylation exclusive to human genomes only (Table 4).

The phosphorylation sites predicted in these sequences are of great research interest as they are not present in the other hominid sequences. These genes could be important candidates for further research on evolutionary changes related to neurodegenerative diseases. Phosphorylation scores closer to 1 are deemed reliable. Higher scoring sites in MARK2 and LTK are deliberated in the discussion section.

Interaction analysis of the protein kinases with exclusive phosphorylation sites

The first neighbor interactors of the above eight kinases which were predicted to contain phosphorylation sites in humans only were retrieved from various databases and literature sources such as STRING (version 8.3),[33] HPRD (version 5)[34] and BioGrid (version 3.0.64).[35] Similarly the pathways associated with each kinase were retrieved from KEGG pathway database[37] and analyzed using WebGeStalt[36] (Supplementary data SD2).

A network was constructed from the collected data to visualize the kinase—target relationships (Fig. 3). The MAP3K14 with 40 neighbors, MARK2 with 32 interactors, GSK3A with 20 interactors, MAPK10 with 19 interactors, and DYRK1A with 17 interactors were found to be the highly populated kinases. These kinases were also found to be connected to each other via a few nodes that interact with more than one kinase, namely YWHAG, MCL1, MAP2K4, AKT1, YWHAB and SFN. For the other three kinases — LTK, LRRK1 and CDKL1 had less than ten interactors and were disconnected from the other kinases.

Upon investigating the presence of the eight kinases in KEGG pathways, we found that only GSK3A, MAPK10 and MAP3K14 participated in important signaling pathways (see SD2). However, further scrutiny proved that in spite of this, six of the highly connected interactors of these nodes were part of many important signaling pathways in humans. AKT1 participated in 66 pathways, MAP2K4 in 12, YWHAB in 7, YWHAB in 6, SFN in 4 and MCL1 in 3 KEGG pathways, including important signaling pathways such as PI3K-Akt, JAK-STAT, MAPK, TNF signaling pathways etc.

The functional enrichment of all the kinases and their targets also indicated overrepresentation of many central signaling pathways of the human body. The KEGG pathways of Alzheimer's and Huntington's from KEGG were clearly overrepresented in the network with P-values of 0.0003 and 0.0034, respectively (Supplementary data SD3). Disease association analysis indicated the enrichment of neurodegenerative diseases with a P-value of 8.10e-05, Alzheimer's disease with a P-value of 0.0005 and Parkinsons's disease with P-value of 0.0343. In addition, all the pathway enrichment analyses indicate that important signaling pathways indicated in neurodegenerative pathways are significantly enriched (see SD3). Thus the insertion of phosphorylation sites in these human kinases might have long range consequences.

Discussion

Studies on experimental models are the main sources for gaining knowledge on the mechanisms of neurodegeneration. Non-human primates (NHP) are widely used as models for studies on neurodegenerative studies as rodent models lack true neurodegeneration. These non-human primates closely resemble humans in genetic
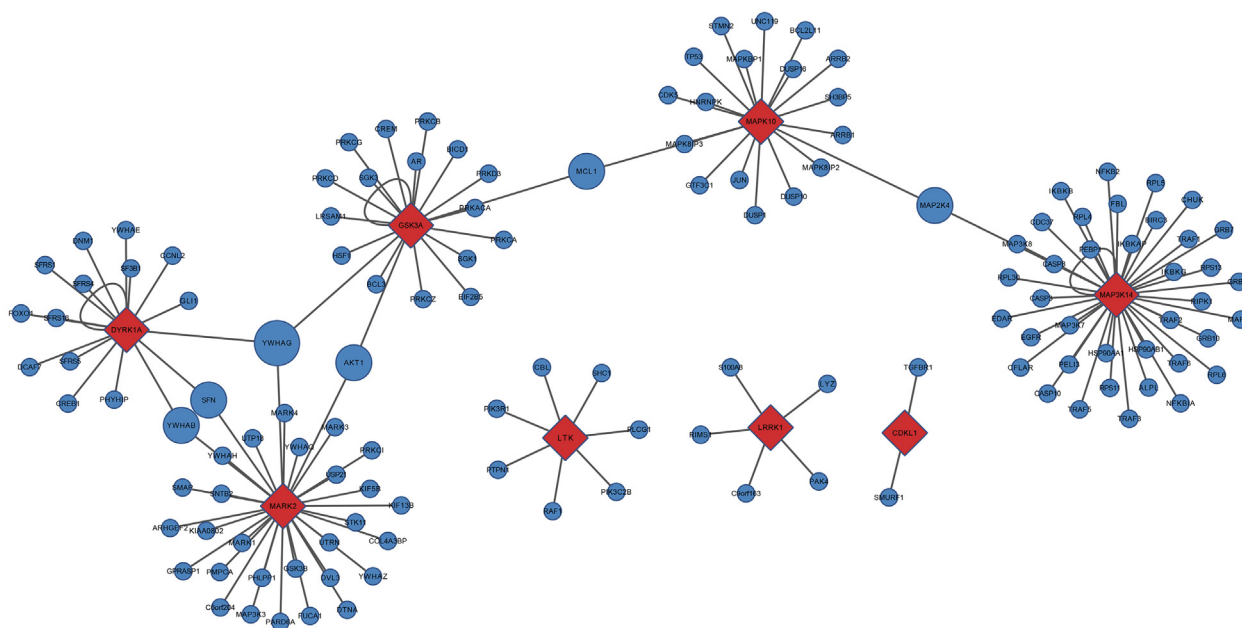

Fig. 3 The network of selected protein kinases and their targets. Red nodes: Protein kinases. Blue nodes: First neighbor interactors; Nodes that interact with more than one protein kinases are denoted with a corresponding increase in size. Links: Interactions.

neuroanatomical as well as cognito-behavioral characteristics. Accordingly the development of NHP models for neurodegenerative diseases holds greater promise for success in the discovery of diagnoses treatments and cures than approaches using other animal species.[44,45] The role of phylogenetically related NHP models in disease research including Parkinson's and Alzheimer's diseases has been reviewed in detail by Sibal LR and Samson KJ.[46] Marvanova M et al. conducted a microarray analysis and assessed the changes in global expression profiles for the validation of non-human primate as experimental models for neurological disorders.[47]

However a few differences have been observed between the nature of disease in human and non-human primates. Besides the lower incidence and less intense presentations of neurodegenerative diseases in primates they are also found to have the ability to recover spontaneously from a mild or moderate form of the induced disease.[48] We hypothesize that the preliminary differences identified in our study could be a starting point to detect the cause of higher rate of neurodegenerative diseases in humans. This provides a path to detect the possible variations between human and non-human primates that might serve as major factors — either primary causes or secondary agents of disease susceptibility.

The results of the current study on kinases are in agreement with studies that state high similarity between hominid genomes. However consequential clues regarding their variations in disease projection were also obtained. MARK2 and LTK are found to have high scoring phosphorylation sites in humans but which are absent in others.

Phosphorylation plays a very important and established role in regulating the activity of proteins including kinases. Specific sequence variations are known to contribute to their specificities and regulation.[49–51] Despite the common features of protein kinases their interactions with different substrates are very specific and are known to be regulated by different methods like secondary messenger binding

phosphorylation at catalytic kinase domain as well as outside it and by association with adapter and regulatory subunits.[52,53]

In the case of MARK2 protein kinase the human sequence was found to be 90.0%, 94.5% and 89.4% identical to the sequences in chimpanzee, gorilla and orangutan, respectively. The variations were found to be mainly present towards the beginning and end of the sequences. Around 50 residues of the gorilla and orangutan on the either edges were characterized by indels and substitutions while the remaining portions were highly similar to the human sequence. However a set of 9 residues [NLSFRFARR] was found to be inserted among the otherwise identical portions (Fig. 4a). To add to this finding, the inserted serine residue was also a strong site for phosphorylation.

Another important kinase identified to be different in humans was LTK. Human LTK was found to contain an exclusive serine phosphorylation site with a high score of 0.995. This site was absent in the non-human primates and more interestingly the region overlapped with glycine repeats that are found to be missing in human (Fig. 4b). An interpro scan of the LTK human sequence further validated this variation by the absence of ATP binding site and TM helix domain in humans compared to the LTK sequence of other hominids. These differences among the LTK sequences of hominids and the association of LTK with Alzheimer's disease in humans could be hence a possible factor for increased disease susceptibility in humans.

Then, we carried out exhaustive literature search for validating the association of these kinases in neurodegenerative diseases, investigated previously. The findings are briefly summarized below. Microtubule-associated protein (MAPK)/microtubule affinity-regulating kinases (MARKs) signaling cascade plays a crucial role in synaptic plasticity and in neurodegenerative diseases.[54] Phosphorylation of human tau protein by MARK2 was demonstrated to show that MARK2 binds to the N-terminal tail of Tau
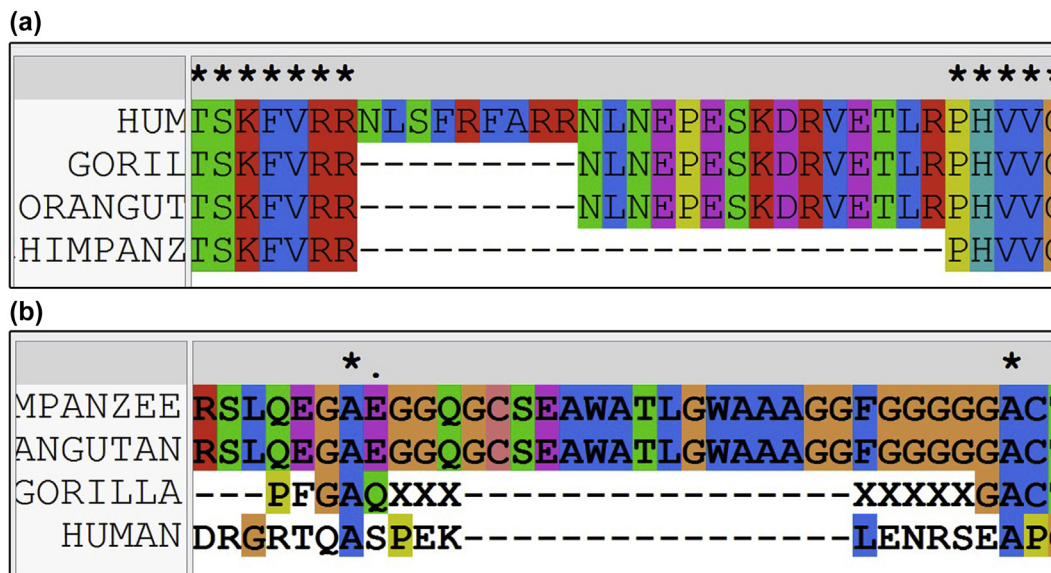

Fig. 4   a) Phosphorylation site inserted in human MARK2. **b)** Phosphorylation site variation and missed glycine repeats in human LTK.

and selectively phosphorylates three major and five minor serine residues in the repeat domain and C-terminal tail.[55]

Similarly, the involvement of protein tyrosine kinases (PTKs) such as LTK and their phosphorylation in the pathology of Alzheimer's disease have been investigated before. Even though, the direct association between LTK and AD has not been investigated before, we include LTK (alias PTK1) in our study as data from previous reports suggest that protein tyrosine kinases are suspected to be involved in the list of PK's altered in AD.[56] Interestingly, our studies through sequence and network analysis reinforce this hypothesis.

JNK3/MAPK10 is a group of serine—threonine protein kinases known to be involved in neuronal death. MAPK10 in the brain tissue and cerebrospinal fluids showed a significant increase in patients with Alzheimer disease and has been linked to cognitive decline.[57] The involvement of the kinase in progression of Huntington's disease has also been indicated.[58]

It is proven that impairment in kinases can lead to aberrated phosphorylation resulting in the pathogenesis of neurodegenerative diseases.[59] The identification of the presence of specific phosphorylation sites specific to human kinases thus indicates possible changes in their activity. Further since kinases catalyze the phosphorylation of a wide variety of other proteins any change in the activity of kinase could also result in changes in the activity of other interacting proteins too.

The pathway enrichment analysis of these kinases and their targets indicating signaling pathways and neurodegenerative disease pathways are a prelude to such effects. Thus the effect of these kinases might not be immediately evident but manifested indirectly through the activity of their targets. It is important to link such genetic differences to their possible effects in the structure and activity of kinases.

Conclusion

A comparison of kinome sequences of four related genomes has provided new insights into kinship evolutions and variability. We have compared 47 human protein kinases associated with Alzheimer's disease, Parkinson disease and Huntington's diseases with their respective orthologs in chimpanzee, gorilla and orangutan. The high degree of conservation of protein kinases associated with neurodegenerative diseases among the hominids impelled us to analyze the minor changes also in more detail. Some interesting differences associated with insertion of new phosphorylation sites in humans whose effects might be further extended to their targets taking part in important signaling pathways were revealed. This suggests the contribution of genetic changes in protein kinases in the incidence of neurodegenerative diseases. However we cannot rule out the role of remaining genes other than kinases for the different susceptibility between hominids. Structural analysis could bring out a significant number of neurodegenerative disease level mutations that might fall at structurally equivalent positions within the catalytic core. Our study provides a pathway to perform experimental analysis of human and other three lineage kinomes

in order to enhance our understanding on their specific biological roles.

Conflicts of interest

The authors declare that they have no conflict of interest in the publication.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.gendis.2016.04.004.

References

1. Bertram L, Tanzi RE. The genetic epidemiology of neurodegenerative disease. *J Clin Invest*. 2005;115:1449—1457.
2. Wetzel EB, Schwarzenbacher R, Lipton SA. Molecular pathways to neurodegeneration. *Nat Med*. 2004;10:S2—S9.
3. Delacourte A, Buee L. Tau pathology: a marker of neurodegenerative disorders. *Curr Opin Neurol*. 2000;13:371—376.
4. Perl DP. Neuropathology of Alzheimer's disease. *Mt Sinai J Med*. 2010;77:32—42.
5. Maslow K. Alzheimer's disease facts and figures. *Alzheimer's Dement*. 2008;4:110—133.
6. Cruts M, Van Broeckhoven C. Molecular genetics of Alzheimer's disease. *Ann Med*. 1998;30:560—565.
7. Schapira AH, Jenner P. Etiology and pathogenesis of Parkinson's disease. *Mov Disord*. 2011;26:1049—1055.
8. Xu J, Kao SY, Lee F, Song W, Jin LW, Yankner BA. Dopamine-dependent neurotoxicity of alpha-synuclein: a mechanism for selective neurodegeneration in Parkinson disease. *Nat Med*. 2002;8:600—606.
9. Cookson MR. Parkin's substrates and the pathways leading to neuronal damage. *Neuromol Med*. 2003;3:1—13.
10. Rubinsztein DC. Lessons from animal models of Huntington's disease. *Trends Genet*. 2002;18:202—209.
11. Lill CM, Bertram L. Towards unveiling the genetics of neurodegenerative diseases. *Semin Neurol*. 2012;31:531—541.
12. Chico LK, Van Eldik LJ, Watterson DM. Targeting protein kinases in central nervous system disorders. *Nat Rev Drug Discov*. 2009;8:892—909.
13. Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. *Science*. 2002;298:1912—1934.
14. Plowman GD, Sudarsanam S, Bingham J, Whyte D, Hunter T. The protein kinases of *Caenorhabditis elegans* a model for signal transduction in multicellular organisms. *Proc Natl Acad Sci*. 1999;96:13603—13610.
15. Hanks SK, Hunter T. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J*. 1995;9:576—596.
16. Kim IC, Kim DS, Kim DW, et al. Comparative genomics of T-complex protein 10 like in humans and chimpanzees. *Genom Inform*. 2005;3:61—65.
17. Anna W, Marie S, Lucia C, Tomas FB. Comparative genomic analysis of human and chimpanzee indicates a key role for indels in primate evolution. *J Mol Evol*. 2006;63:682—690.
18. Krishanpal A, Juliette M, Narayanaswamy S. Comparative kinomics of human and chimpanzee reveals unique kinship and functional diversity generated by new domain combinations. *BMC Genomics*. 2008;9:625.
19. Puente XS, Velasco G, Fernández AG, Bertranpetit J, King MC, Otín CL. Comparative analysis of cancer genes in the human and chimpanzee genomes. *BMC Genom*. 2006;26:7—15.

20. Kimura N, Nakamura S, Goto N, et al. Senile plaques in an aged western lowland gorilla. *Exp Anim*. 2001;50:77—81.

21. Gearing M, Tigges J, Mori H, Mirra SS. Beta-Amyloid (A beta) deposition in the brains of aged orangutans. *Neurobiol Aging*. 1997;18:139—146.

22. Toledano A, Alvarez MI, López-Rodríguez AB, Toledano-Díaz A, Fernández-Verdecia CI. Does Alzheimer's disease exist in all primates? Alzheimer pathology in non-human primates and its pathophysiological implications (I). *Neurologia*. 2012;27: 354—369.

23. Toledano A, Alvarez MI, López-Rodríguez AB, Toledano-Díaz A, Fernández-Verdecia CI. Does Alzheimer's disease exist in all primates? Alzheimer pathology in non-human primates and its pathophysiological implications (II). *Neurologia*. 2014;29: 42—55.

24. Hardy JA, Higgins GA. Alzheimer's disease: the amyloid cascade hypothesis. *Science*. 1992;256:184—185.

25. Karran E, Mercken M, De Strooper B. The amyloid cascade hypothesis for Alzheimer's disease: an appraisal for the development of therapeutics. *Nat Rev Drug Discov*. 2011;10: 698—712.

26. Taylor JP, Hardy J, Fischbeck KH. Toxic proteins in neurodegenerative disease. *Science*. 2002;296:1991—1995.

27. Krainc D. Clearance of mutant proteins as a therapeutic target in neurodegenerative diseases. *Arch Neurol*. 2010;67:388—392.

28. Ramanan VK, Saykin AJ. Pathways to neurodegeneration: mechanistic insights from GWAS in Alzheimer's disease, Parkinson's disease, and related disorders. *Am J Neurodegener Dis*. 2013;2:145.

29. Saranya J, Jeyakumar N. Classification and domain analysis of protein kinases in hominids. *Curr Sci*. 2016;110:828—838.

30. Suresh S, Saranya J, Raja K, Jeyakumar N. HomoKinase: a curated database of human protein kinases. *ISRN Comput Biol*. 2013;2013:5.

31. OMIM 2013. Available at: http://www.ncbi.nlm.nih.gov/omim.

32. Hoffmann R. A wiki for the life sciences where authorship matters. *Nat Genet*. 2008;40:1047—1051.

33. Larkin MA, Blackshields G, Brown NP, et al. Clustal W and Clustal X version 2.0. *Bioinformatics*. 2007;23:2947—2948.

34. EMBOSS Web Pages. Available at: http://www.sanger.ac.uk/software/EMBOSS.

35. Blom N, Gammel토ft S, Brunak S. Sequence and structure based prediction of eukaryotic protein phosphorylation sites. *J Mol Biol*. 2007;294:1351—1362.

36. Franceschini A, Szklarczyk D, Frankild S, et al. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res*. 2013;41(Database issue): D808—D815.

37. Peri S, Navarro JD, Kristiansen TZ, et al. Human protein reference database as a discovery resource for proteomics. *Nucleic Acids Res*. 2004;32(Database issue): D497—D501.

38. Winter AG, Wildenhain J, Tyers M. BioGRID REST Service, BiogridPlugin2 and BioGRID WebGraph: new tools for access to interaction data at BioGRID. *Bioinformatics*. 2011;27: 1043—1044.

39. Wang J, Duncan D, Shi Z, Zhang B. Web-based gene set analysis toolkit (WebGestalt): update 2013. *Nucleic Acids Res*. 2013;41: W77—W83.

40. Kanehisa M, Goto S, Furumichi M, Tanabe M, Hirakawa M. KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res*. 2010;38(suppl 1):D355—D360.

41. Cerami EG, Gross BE, Demir E, et al. Pathway commons, a web resource for biological pathway data. *Nucleic Acids Res*. 2011; 39(suppl 1):D685—D690.

42. Kelder T, Pico AR, Hanspers K, Van Iersel MP, Evelo C, Conklin BR. Mining biological pathways using WikiPathways web services. *PLoS One*. 2009;4:e6447.

43. Hewett M, Oliver DE, Rubin DL, et al. PharmGKB: the pharmacogenetics knowledge base. *Nucleic Acids Res*. 2002;30: 163—165.

44. Vitale A, Manciocco A, Alleva E. The 3R principle and the use of non-human primates in the study of neurodegenerative diseases: the case of Parkinson's disease. *Neurosci Biobehav Rev*. 2009;33:33—47.

45. Chan AW. Transgenic nonhuman primates for neurodegenerative diseases. *Reprod Biol Endocrinol*. 2004;16:39.

46. Sibal LR, Samson KJ. Nonhuman primates: a critical role in current disease research. *ILAR J*. 2001;42:74—84.

47. Marvanova M, Me Nager J, Bezard E, Bontrop RE, Pradier L, Wong G. Microarray analysis of nonhuman primates: validation of experimental models in neurological disorders. *FASEB J*. 2003;17:929—931.

48. Bingaman KD, Bakay RA. The primate model of Parkinson's disease: its usefulness limitations and importance in directing future studies. *Prog Brain Res*. 2000;127:267—297.

49. Krupa A, Srinivasan N. The repertoire of protein kinases encoded in the draft version of the human genome: atypical variations and uncommon domain combinations. *Genome Biol*. 2002;3. RESEARCH0066.

50. Krupa A, Srinivasan N. Diversity in domain architectures of Ser/Thr kinases and their homologues in prokaryotes. *BMC Genom*. 2005;6:129.

51. Krupa A, Anamika, Srinivasan N. Genome-wide comparative analyses of domain organisation of repertoires of protein kinases of *Arabidopsis thaliana* and *Oryza sativa*. *Gene*. 2006; 380:1—13.

52. Hunter T. Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. *Cell*. 1995;80: 225—236.

53. Shenolikar S. Protein phosphorylation: hormones drugs and bioregulation. *FASEB J*. 1988;2:2753—2764.

54. Corrêa SA, Eales KL. The role of p38 MAPK and its substrates in neuronal plasticity and neurodegenerative disease. *J Signal Transduct*. 2012;2012, 649079.

55. Schwalbe M, Biernat J, Bibow S, et al. Phosphorylation of human tau protein by microtubule affinity-regulating kinase 2. *Biochemistry*. 2013;52:9068—9079.

56. Shapiro IP, Masliah E, Saitoh T. Altered protein tyrosine phosphorylation in Alzheimer's disease. *J Neurochem*. 1991;56: 1154—1162.

57. Gourmaud S, Paquet C, Dumurgier J, et al. Increased levels of cerebrospinal fluid JNK3 associated with amyloid pathology: links to cognitive decline. *J Psychiatry Neurosci*. 2015;40:151.

58. Muth F, Gunther M, Bauer SM, et al. Tetra-substituted pyridinylimidazoles as dual inhibitors of p38a mitogen-activated protein kinase and c-Jun N-terminal kinase 3 for potential treatment of neurodegenerative diseases. *J Med Chem*. 2014; 58:443—456.

59. Wagey RT, Krieger C. Abnormalities of protein kinases in neurodegenerative diseases. *Prog Drug Res*. 1998;51:133—183.