Investigating regulatory signatures of human autophagy related gene 5 (ATG5) through functional *in silico* analysis

Avni Vij [a], Rohit Randhawa [a], Jyoti Parkash [b], Harish Changotra [a,*]

[a] *Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Waknaghat, Solan, 173 234, Himachal Pradesh, India*
[b] *Centre for Animal Sciences, School of Basic and Applied Sciences, Central University Punjab, Mansa Road, Bathinda 151 001, India*

ARTICLE INFO

ABSTRACT

Autophagy is an essential, homeostatic process which removes damaged cellular proteins and organelles for cellular renewal. ATG5, a part of E3 ubiquitin ligase-like complex (Atg12-Atg5/Atg16L1), is a key regulator involved in autophagosome formation - a crucial phase of autophagy. In this study, we used different *in silico* methods for comprehensive analysis of ATG5 to investigate its less explored regulatory activity. We have predicted various physico-chemical parameters and two possible transmembrane models that helped in exposing its functional regions. Twenty four PTM sites and 44 TFBS were identified which could be targeted to modulate the autophagy pathway. Furthermore, LD analysis identified 3 blocks of genotyped SNPs and 2 deleterious nsSNPs that may have damaging impact on protein function and thus could be employed for carrying genome-wide association studies. In conclusion, the information obtained in this study could be helpful for better understanding of regulatory roles of ATG5 and provides a base for its implication in population-based studies.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Autophagy is highly conserved, intracellular, self-degrading process that plays fundamental role in maintaining cellular homeostasis. In mammalian cells, basal level of autophagy is beneficial but the pathway gets amplified under stress conditions such as nutrient starvation, hypoxia, and pathogen infection in order to promote cell survival and adaptation (Levine and Yuan, 2005; Kiel, 2010; Wirawan et al., 2012; Randhawa et al., 2015). Recent reports have demonstrated the involvement of autophagy in the development of cancer, neurodegenerative diseases, cardiac hypertrophy, diabetic nephropathy, infectious diseases, muscles and liver disorders (Huett et al., 2010; Yang et al., 2011; Badadani, 2012; Chiramel et al., 2013; Ghavami et al., 2014; Cursio et al., 2015; Ding and Choi, 2015). Involvement of autophagy in various disorders shows that this process has very important biological and clinical significance. Moreover, recent studies have shown that modulation of this pathway may result in better treatment response in breast cancer (Nagelkerke et al., 2014), infectious diseases and other various diseases (Sir et al., 2010; Rubinsztein et al., 2012; Shoji-Kawata et al., 2013). Therefore, it becomes important to understand each phase of the process so that it could be utilized for better therapeutic/targeted treatment management (Cheng et al., 2013).

The process of autophagy initiates in response to stress signals and proceeds with the formation of a double membranous structure, known as the autophagosome. The autophagosome engulfs cytoplasmic constituents, and then fuses with the lysosome for degradation of its contents (Mizushima, 2007; Glick et al., 2010; Tanida, 2011). The process of formation of autophagosome is regulated by number of autophagy related (*ATG*) genes. With the initial discovery in yeast, >37 *ATG* genes have been identified until now and most of them have homologues in mammalian cell (Ohsumi, 2014). Among them ATG5 is a key regulator involved in autophagosome formation. ATG5 (Lys149) is covalently linked to with ATG12 (C-terminal Gly 186) through reactions mediated by an ubiquitin like system members, ATG7 (E1) and ATG10 (E2). ATG5, after conjugating with ATG12 further joins non-covalently with a multimeric protein, ATG16L1. ATG12-ATG5/ATG16L1 complex possesses E3 like activity, helps in the formation of ATG8-phosphatidylethanolamine (PE) and its targeting to the pre-autophagosomal structure (PAS) which plays a pivotal role in initiation of autophagosome formation (Hwang et al., 2012; Otomo et al., 2013; Walczak and Martens, 2013).

ATG5 is a 141.34 kb gene located on chromosome 6q21, contains 10 exons and codes for a 275 amino acid autophagy protein 5. Changes in ATG5 levels directly or indirectly influence the levels of autophagy pathway. Neonatal mouse deficient in ATG5 dies within a day due to energy depletion indicating the functional significance of autophagy process (Kuma et al., 2004). Moreover, conditional ATG5 knockout mouse not only suffer from nutrient depletion but also develops dysfunctional hepatic tissue and progressive neurodegeneration (Winslow and Rubinsztein, 2008). In another study, systemic mosaic deficiency of ATG5 was shown to lead to the development of multiple liver tumors in mice (Takamura et al., 2011). Moreover, this protein is strongly

down-regulated in colorectal cancer in humans (Cho et al., 2012) and overexpression of ATG5 may contribute to autoimmune demyelination, multiple sclerosis (Alirezaei et al., 2009) and pathogenesis of hepatitis B virus infection (Mukherjee et al., 2014). ATG5 plays important role in cell death progression by interacting with FADD (Pyo et al., 2005), B and T-cell survival and proliferation (Pua et al., 2007; Miller et al., 2008) mitochondrial maintenance (Stephenson et al., 2009), regression of deficits in neurogenesis at embryonal level due to dysfunctional autophagy (Lv et al., 2014), involved in normal adipocyte differentiation (Baerga et al., 2009), primary ciliogenesis (Tang et al., 2013), required in antigen presentation by dendritic cells (Lee et al., 2010), provides cellular immunity to intracellular pathogens (Zhao et al., 2008) and promotes mitotic catastrophe independent of autophagy (Maskey et al., 2013). Not only the expression pattern, but also recent genetic studies reveal that polymorphisms in *ATG5* gene contribute to the pathogenesis and susceptibility of various diseases in human. In this context, case-control and genome-wide association studies (GWAS) in different populations have linked several single nucleotide polymorphisms (SNPs) in ATG5 to the susceptibility of systemic lupus erythematosus (Gateva et al., 2009), asthma (Martin et al., 2012), thyroid carcinoma (Plantinga et al., 2014), rheumatoid arthritis (Orozco et al., 2011), systemic sclerosis (Mayes et al., 2014), and sporadic Parkinson's disease (Chen et al., 2013).

Taking into account the functional relevance and clinical significance of ATG5, it becomes important to characterize it computationally in order to understand more about its less studied characteristics which otherwise are tedious and time consuming. Moreover, *in silico* analysis of this protein will provide a strong base for its prospective implications as in population-based studies. In this study, we have explored ATG5 gene *in silico* to study its evolutionary relationship with other species, sequence feature-based analysis, detected signal peptide and transmembrane regions, identified essential regulatory elements, over-represented transcription factor binding sites, linkage disequilibrium and nsSNPs, putative post-translational modification sites and protein-protein interactions. This study provides data about important features of ATG5 which would be further helpful to unravel its various functions.

2. Materials and methods

The computational methodology followed in this study is shown in the form of a flow diagram (Fig. 1). The study consists of the following main steps: (1) phylogenetic analysis, (2) sequence feature-based analysis (3) detection of signal peptides and transmembrane regions, (4) identification of regulatory elements and over-represented TFBS, (5) linkage disequilibrium (LD) analysis and prediction of non-synonymous Single Nucleotide Polymorphism (nsSNP), (6) elucidation of post-translational modification sites, and (7) protein-protein interaction studies.

2.1. Data mining

The protein sequence for human ATG5 was retrieved from the National Center for Biotechnology Information (NCBI) (Protein ID: NP_004840); the UniProt database (UniProtKB ID: Q9H1Y0). SNPs information for human ATG5 was obtained from the NCBI, dbSNP database.

2.2. Phylogenetic analysis

For the evolutionary relationships analysis, the protein sequence for human ATG5 was retrieved from the UniProt database (UniProtKB ID: Q9H1Y0) and corresponding reviewed sequences for other species from 20 families were also selected. Multiple sequence alignment was performed using ClustalW tool and phylogenetic tree was constructed using Molecular Evolutionary Genetics Analysis 6 (MEGA6) which estimates the rates of molecular evolution and deduce ancestral links accordingly (Tamura et al., 2013). Maximum likelihood (ML) method was applied for the phylogenetic reconstruction. The maximum likelihood method for constructing a phylogenetic tree involves various standard statistical techniques for inference of probability distribution to assign probabilities to particular possible phylogenetic trees. For the verification of the inferred tree, bootstrap analysis was also performed by taking 1000 bootstrap replicates and 70% cutoff value was set to generate statistically significant phylogenetic tree.
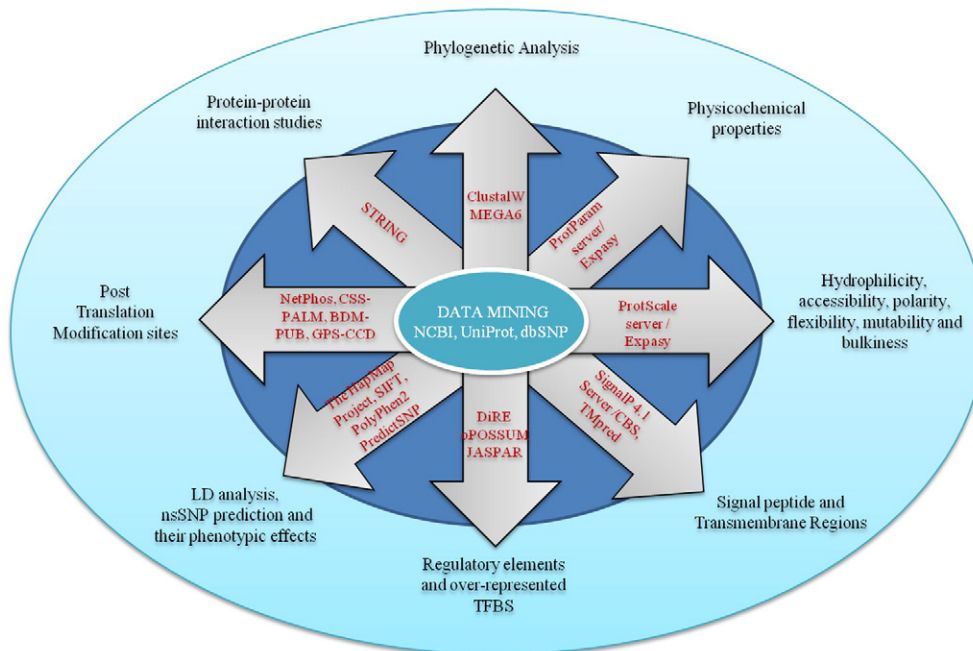

Fig. 1. Schematic representation of the computational methodology followed in the study.

2.3. Sequence feature-based analysis

Sequence based features include the physicochemical, biochemical and structural properties that help to unravel important characteristics that are essential for analysis of protein characters , functions and protein-protein interactions (Bock and Gough, 2001). ProtParam tool (Gasteiger et al., 2005) (http://web.expasy.org/protparam/) was utilized for the computation of various physicochemical properties of human ATG5. Protein sequence was submitted as an input and the computed parameters included molecular weight, theoretical pI, amino acid composition, atomic composition, estimated half-life (Tobias et al., 1991), instability index (Guruprasad et al., 1990) , aliphatic index (Ikai, 1980)and grand average of hydropathicity (GRAVY) (Kyte and Doolittle, 1982) . Biochemical and structural features of ATG5 like hydrophilicity, accessibility, polarity, flexibility, mutability and bulkiness play important role in retaining protein stability and structure (White, 1992; Teng et al., 2010). ProtScale (http://web.expasy.org/protscale/) was explored to predict the protein profile based on amino acid scales like hydrophilicity, accessibility , polarity , flexibility , mutability and bulkiness (Gasteiger et al., 2005). The protein sequence of human ATG5 was provided as an input and the program was run on selected amino acid scale.

2.4. Detection of signal peptides and transmembrane regions

SignalP 4.1 Server (http://www.cbs.dtu.dk/services/SignalP/) was used to predict the signal peptide in human ATG5. Protein sequence was provided as an input (Petersen et al., 2011). The TMpred program (http://www.ch.embnet.org/software/TMPRED_form.html) was used to predict the transmembrane regions and their orientation. The algorithm is based on the statistical analysis of TMbase, a database of naturally occurring transmembrane proteins where prediction is made using a combination of several weight-matrices for scoring.

2.5. Identification of regulatory elements and over-represented transcription factor binding sites (TFBS)

It is through the complex cooperative activity of proximal promoters and distant regulatory elements (REs, for example, enhancers, repressors, silencers and transcription factors) that regulate gene expression in eukaryotic genomes. Identification of REs and TFs enhance the functional annotation of the genome which eventually leads to better understanding of mechanism of gene regulation (Maston et al., 2006; Gotea and Ovcharenko, 2008). Distant Regulatory Elements of co-regulated genes (DiRE) (Gotea and Ovcharenko, 2008) and oPOSSUM 3 (Kwon
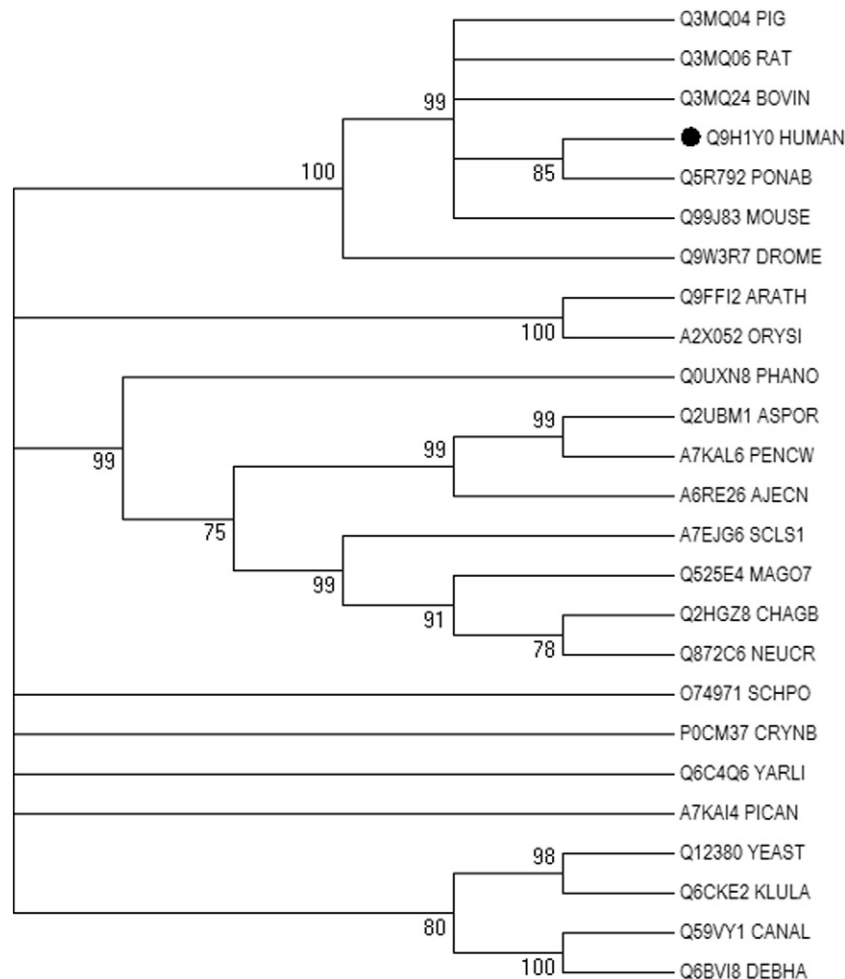

Fig. 2. Phylogentic analysis was done by applying MEGA6. Multiple sequence alignment was performed by ClustalW and evolutionary history was inferred using maximum likelihood method. The phylogentic tree reveals evolutionary relationship among 25 organisms from a total of 20 families.

Table 1
Computed physiological parameters of ATG5.

Sr. no	Physiological parameters	Observations
1.	Number of amino acids	275
2.	Molecular weight	32,447.2 Da
3.	Theoretical pI	5.47
4.	Total number of negatively charged residues (Asp + Glu)	42
5.	Total number of positively charged residues (Arg + Lys)	31
6.	Atomic composition:	
	Carbon C	1480
	Hydrogen H	2254
	Nitrogen N	378
	Oxygen O	418
	Sulfur S	13
7.	Formula	$C_{1480}H_{2254}N_{378}O_{418}S_{13}$
8.	Total number of atoms	4543
9.	Estimated half-life	30 h (mammalian reticulocytes, *in vitro*). >20 h (yeast, *in vivo*). >10 h (*Escherichia coli, in vivo*).
10.	The instability index (II)	45.29
11.	Aliphatic index	82.62
12.	Grand average of hydropathicity (GRAVY)	−0.468

et al., 2012) were used for identification of the REs and over-represented transcription binding sites (TFBS), respectively. Additionally, JASPAR database (Mathelier et al., 2014) was likewise investigated to distinguish different classes, families and sequence logos for TFs and their binding sites for comparable datasets.

2.6. Linkage disequilibrium analysis, prediction of non-synonymous single nucleotide polymorphism (nsSNP) and their phenotypic effects

For linkage disequilibrium analysis, CEU (CEPH−Utah Residents with Northern and Western European Ancestry) population was

selected and the genotype data for the *ATG5* gene was retrieved from The International HapMap project (Thorisson et al., 2005). Haploview version 4.2 was used for linkage disequilibrium (LD) analysis. Parameters analyzed for genetic association were D′ and r^2. The D′ value provides the measure of LD between the blocks and its value closer to zero shows a higher amount of historical recombination between the blocks and r^2 gives the correlation coefficient between the two loci. Haplotypes were constructed using the algorithm implemented in Haploview 4.2. The deleterious/damaging nsSNPs were predicted using Sorting Intolerant From Tolerant (SIFT) (Kumar et al., 2009) and Polymorphism Phenotyping (PolyPhen-2) (Adzhubei et al., 2010) and PredictSNP (Bendl et al., 2014).

2.7. Elucidation of post-translational modification (PTM) sites

PTM sites of phosphorylation, ubiquitination, palmitoylation and calpain-cleavage were predicted in the ATG5 protein using various web servers. The NetPhos algorithm was used for the prediction of phosphorylation sites at serine (S), threonine (T) and tyrosine (Y) residues in the ATG5 amino acid sequence. This algorithm utilizes an artificial neural network (ANN) based method which is trained from Phospho Base, a database of experimentally validated phosphorylated proteins (Blom et al., 1999). The palmitoylation sites were obtained from CSS-PALM, a tool based on the Clustering and Scoring Strategy (CSS) algorithm for the prediction of palmitoylation sites (Ren et al., 2008). An online web server called BDM-PUB (Prediction of Ubiquitination site with Bayesian Discriminant Method; http://bdmpub.biocuckoo.org/) was used to predict ubiquitination sites. Group-based Prediction System-Calpain Cleavage Detector (GPS-CCD) 1.0 tool was utilized for the prediction of calpain-cleavage sites in ATG5 (Liu et al., 2011).

2.8. Protein-protein interaction studies

Search Tool for the Retrieval of Interacting Genes/Proteins-STRING database was explored for predicting functional partners of ATG5 (Szklarczyk et al., 2015). STRING version 10 generates a protein network
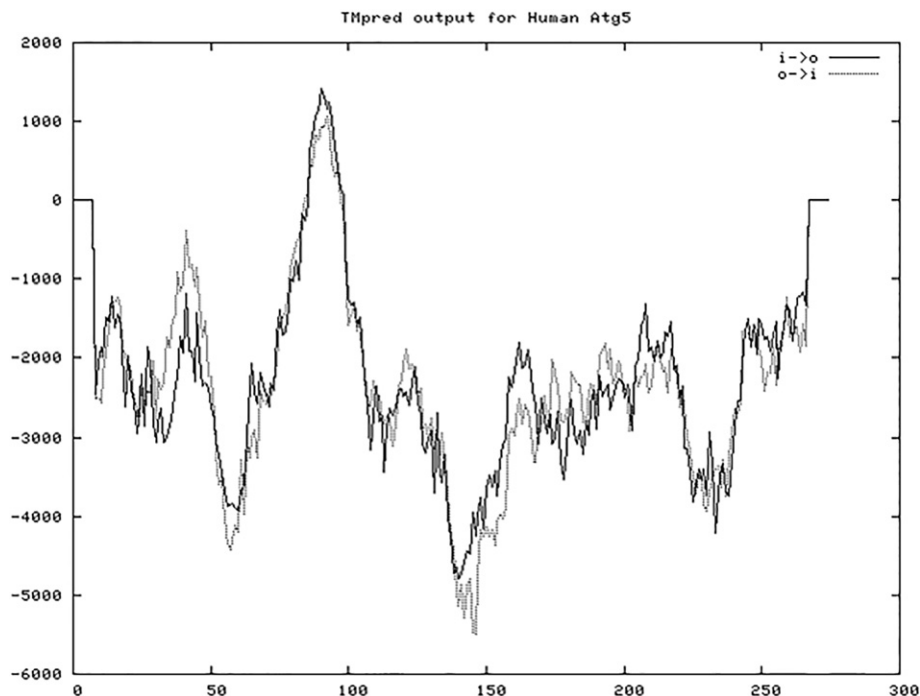

Fig. 3. Predicted transmembrane regions in ATG5 using TMpred server. X axis represents amino acid sequence from N- to C-terminal and Y axis represents scores computed by each server. Score > 500 represents significant possibility for existence of TMD.
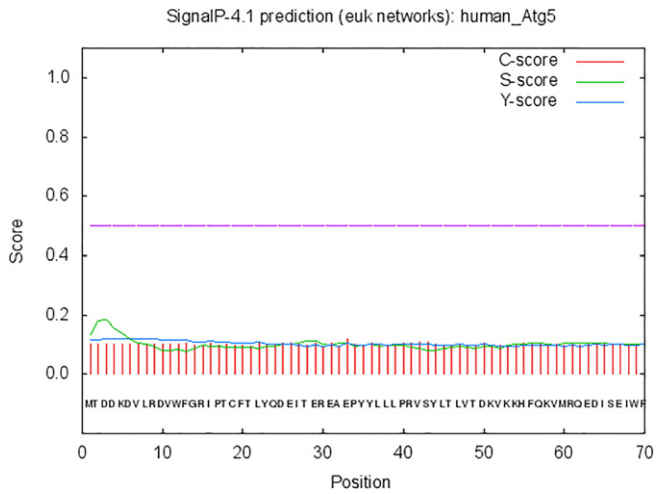
Fig. 4. Prediction of Signal peptide of ATG5 by using SignalP 4.1Server. It can be observed from the figure that all scores are less than the standard value (0.5), no signal peptide is predicted in ATG5. (X axis represents amino acid sequence from N- to C-terminal and Y axis represents scores computed by each server).

view that can be inspected for the interaction evidence, re-adjustments for the score-cutoffs and network size limits can be performed and detailed information about the interacting proteins can be obtained.

3. Results and discussion

3.1. Data retrieval

The amino acid sequence for the ATG5 protein was obtained from NCBI (Protein ID: NP_004840) and UniProt (UniProtKB ID: Q9H1Y0).

Table 2
The list of 10 most important TFs in *ATG5* determined by DiRE.

S. no.	Transcription factor	Occurrence	Importance
1	ISRE	100.00%	1
2	LDSPOLYA	100.00%	0.97812
3	CREL	50.00%	0.63438
4	ZID	50.00%	0.5
5	PXR	50.00%	0.49961
6	STAT6	50.00%	0.49617
7	COUPTF	50.00%	0.49609
8	CDPCR3HD	50.00%	0.49609
9	YY1	100.00%	0.49531
10	NFKAPPAB50	50.00%	0.49375

This 275 amino acid sequence was analyzed by various *in silico* tools. The SNP information was obtained from NCBI, dbSNP. According to dbSNP, human ATG5 contains 7358 Single nucleotide polymorphisms. This SNP dataset was selected as a batch report and saved in the form of their rs IDs for prediction of nsSNPs through various *in silico* tools.

3.2. Phylogentic analysis shows the evolutionary links of ATG5 revealing its conservation profile

The analysis was performed for ATG5 protein sequence of humans (Uniprot ID: Q9H1Y0) and 25 other organisms from 20 different families. Only the reviewed sequences were selected for the analysis (Table S1). In MEGA5, bootstrap test of phylogeny with 1000 bootstrap replicates and 70% cutoff value was set as parameters. ML method was used for phylogenetic tree reconstruction as shown in Fig. 2. Phylogenies are helpful for arranging information of biological diversity, for structuring classifications and for providing knowledge into event that happened amid evolution. The phylogenetic tree obtained represents the evolutionary relationships among the 25 taxa. From this tree
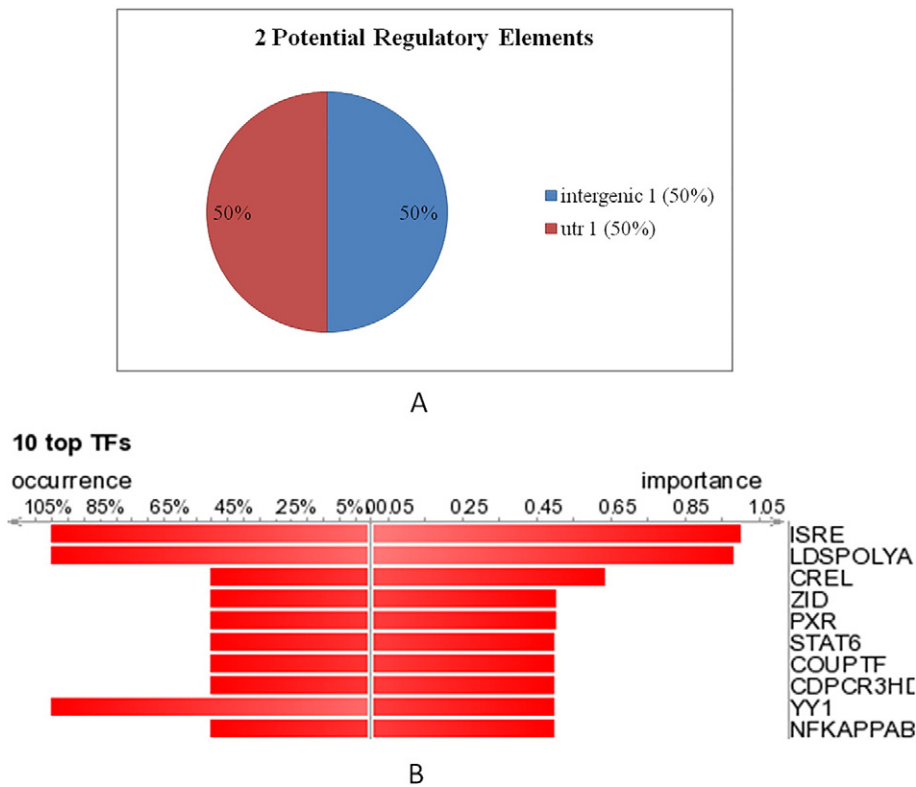

Fig. 5. A summary of output generated by the DiRE server. (a) the graphical representation of the predicted regulatory elements (REs) and their genomic distribution. (b) 10 most important transcription factors (TFs) with their occurrence and importance rates. The 'occurrence' represents the fraction of putative REs that contain a particular TFBS, while the 'importance' is defined as the product of the TF occurrence and its weight.
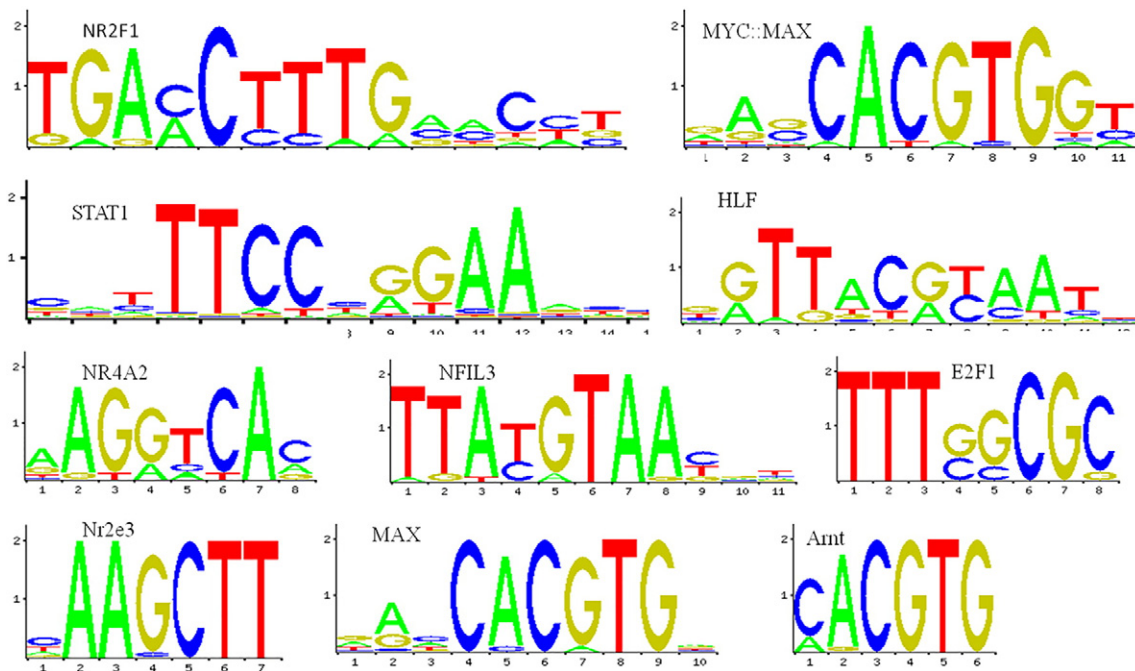
Fig. 6. Sequence logos obtained from JASPAR for the TFBS in *ATG5*.

diagram, we could infer that *Homo sapiens* and *Pongo abelii* have 85% similarity depicting their closely related ancestors. *Mus musculus, Rattus norvegicus, Sus scrofa* and *Bos Taurus* show 99% similarity and thus are closely related to each other as compared to other species. Moreover, *Arabidopsis thaliana* and *Oryza sativa* have 100% similarity and could have the same ancestral origin. *Debaryomyces hansenii* and *Candida albicans* also show 100% similarity revealing their common ancestral origin. Therefore, the data provides information from which it could be inferred that these successions have a tendency to remain closely related in a family as well as with species from other families.

3.3. ATG5 is a hydrophilic protein

ProtParam server was used to compute various physicochemical properties that were deduced from a protein sequence (Gasteiger et al., 2005). The computed parameters and observations are shown in Table 1. The physiological parameters are essential for the protein identification and analysis which plays a central role in the investigation of proteins characteristics applying techniques like two-dimensional (2-D) gels and mass spectrometry (Wilkins et al., 1999). Moreover, properties of a protein, *e.g. in vivo* half life, have been shown to be altered in pathological conditions such as viral infections, neurodegenerative disorders, or cancer (Bojkowska et al., 2011). Therefore, better understanding of physiological parameter/degradation rates of a protein would definitely help in facilitating drug development for various human diseases. The instability index (II) provides an estimate of the stability of a protein in a test tube (Guruprasad et al., 1990). The expanding utilization of recombinantly expressed proteins in the pharmaceutical industry has highlighted issues, for example, their stability amid long-term storage. Therefore, II provides essential information in this aspect.

Here, the instability index was predicted to be 45.29, which classifies ATG5 protein as unstable but aliphatic index, which is defined as the relative volume occupied by aliphatic side chains (alanine, valine, isoleucine, and leucine), classifies this protein as thermostable. Higher value of aliphatic index (82.62) indicates better stability of this protein (Ikai, 1980). Normally, major portion of ATG5 and ATG12 *in vivo* exists as a conjugate (Romanov et al., 2012). Increasing positive score of GRAVY of the linear polypeptide sequence indicates greater hydrophobicity. Negative GRAVY and positive GRAVY values indicate hydrophilic and hydrophobic proteins, respectively. Hydrophobicity is a key determinant in surface properties hence in directing protein-protein interactions. Hydrophobicity also influences amino acid side chain packing and protein folding and therefore an important factor that determine stability of a protein (41). Here, the GRAVY value for ATG5 was found to be -0.468, indicating that it is a hydrophilic protein. Overall, these physico-chemical characteristics of ATG5 may help in unfolding the different physiological functions of the protein. These physiological parameters of ATG5 are shown in Table 1. ProtScale was used to analyze the hydrophilicity, accessibility, polarity, flexibility, mutability and bulkiness of ATG5. The higher score indicates the higher probability of each parameter of ATG5. The hydrophilicity values of this protein are between -1.567 (AA position 83) and 2.022 (AA position 232) (Supplementary Fig. (SF) 1a). Polarity (P) is the dipole-dipole intermolecular interactions between the positively and negatively charged residues. The polarity values are between 0.776 (AA position 93) and 34.243 (AA position 148) (SF 1b). The accessibility values obtained by Janin score are between -0.867 (AA position 148) and 0.356 (AA position 87) (SF 1c). It represents the free energy of transfer from inside to outside of a globular protein while the average flexibility values are between 0.379 (AA position 125) and 0.497 (AA position 232) (SF 1d).

Table 3
The class-wise order of transcription factors recognized in *ATG5* by oPOSSUM tool.

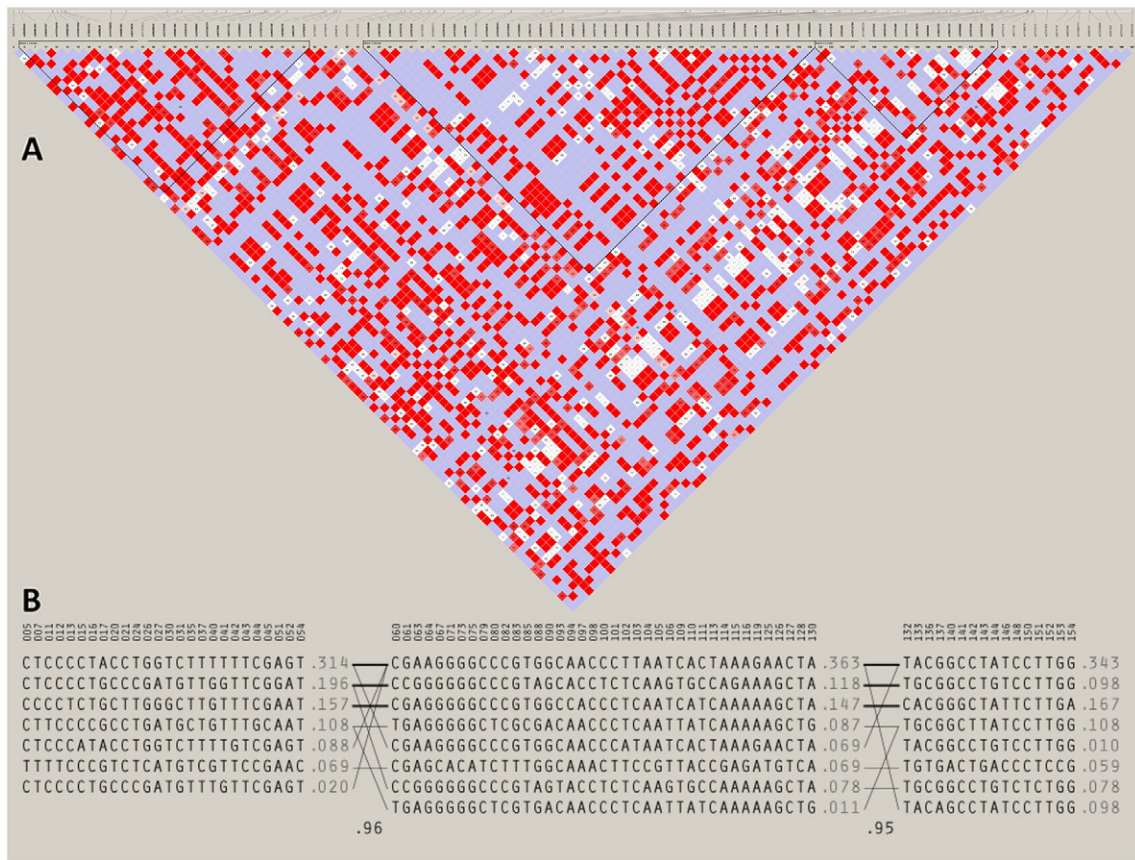| Class name	TFS	Total gene hits	TFBS hits
Zinc-coordinating	NR2F1,NR4A2,Nr2e3,RORA_1	4	7
Zipper-type	MYC,MAX,HLF,NFIL3,MAX,Arnt,Myc,Mycn,CREB1,USF1	9	9
Ig-fold	STAT1,Stat3	2	2
Winged Helix-Turn-Helix	E2F1	1	1

Fig. 7. Linkage disequilibrium (LD) analysis of *ATG5*. (a) Haploblock diagram of the LD plot. Three distinct alternative blocks have been identified, 26 markers in first block of 38 kb size, 39 markers in second block of 58 kb size and 16 markers in third block of 15 kb size. (b) Strong significant correlation was seen between the blocks (correlation of 0.96 has been found among interactions between block (1) and (2) and correlation of 0.95 has been found among interactions between block (2) and (3) of the genotyped SNPs of *ATG5*.

Flexibility of protein structure is important as helps in building up its interactions with other various partners, ligands, proteins, or nucleic acids to form complex structures (Craveur et al., 2015). The relative mutability indicates the probability that a given amino acid can be changed to others during evolution (Teng et al., 2010). The relative mutability values are between 49.667 (AA position 83) and 97.111 (AA position 64) (SF 1e). The bulkiness values are between 12.633 (AA position 229) and 18.997 (AA position 38) (SF 1f). Bulkiness reflects the function of an individual residues in the entire protein configuration and hence it may influence the local conformation of protein (Zimmerman et al., 1968).

3.4. ATG5 contains predicted 2 transmembrane models and no signal peptide

The transmembrane regions of ATG5 were predicted using TMpred Server (Stoffel, 1993). The results suggest that there are two possible models for transmembrane helices predicted in ATG5 (Fig. 3). For the first transmembrane helix, which is from inside to outside, the maximum score is 1421 and extends from 83 AA to 104 AA is the strongly preferred model. Second possible transmembrane helix with outside to inside orientation, whose maximum score is 1053, extends from 83 AA to 102 AA is an alternative model. Prediction scores >500 are considered to be significant in TMpred server, therefore, ATG5 contains 2 possible predicted helices and there is a strong preference of inside to outside orientation that of the first helix. Prediction of transmembrane regions helps in defining membrane topology of proteins. Studies show that ATG5 alone has membrane binding capability *in vitro*. This membrane binding by ATG5 is essential *in vivo* for autophagy pathway though not required for the localization of the ATG12–ATG5/ATG16L1 complex to the PAS (Romanov et al., 2012). Determining the sub-cellular localization of a protein is one of the important steps towards understanding its function. Therefore, information regarding the presence and absence of a signal peptide in a protein could reveal its sub-cellular location and various functions (Emanuelsson et al., 2007). SignalP 4.1Server was used to predict the signal peptide of ATG5. Fig. 4 shows all the scores obtained from SignalP 4.1 Server, - including C-score (raw cleavage site score), S-score (signal

Table 4
The non-synonymous Single Nucleotide Polymorphisms in the coding region of ATG5, two with putative damaging effects.

SNP INFORMATION		SIFT		PolyPhen-2		PredictSNP			
						PolyPhen-1	nsSNPAnalyzer	PhD-SNP	MAPP
SNP ID	Amino acid change	Score	Predicted impact	Probability score	Predicted impact	Prediction	Prediction	Prediction	Prediction
rs34793250	M129 V	1.00	TOLERATED	0.000	BENIGN	Deleterious	Disease	Neutral	Deleterious
rs77859116	I65V	0.21	TOLERATED	0.000	BENIGN	Neutral	Neutral	Neutral	Neutral
rs115576116	A95D	0.32	TOLERATED	0.000	BENIGN	Neutral	Neutral	Deleterious	Neutral

Table 5
Putative phosphorylation sites predicted in ATG5.

S. No.	Name	Position	Context sequence	Score	Prediction
1	ATG5	66	QEDISEIWF	0.982	*S*
2	ATG5	106	VHFKSFPEK	0.946	*S*
3	ATG5	225	EVCPSAIDP	0.990	*S*
4	ATG5	259	SEHLSYPDN	0.963	*S*
5	ATG5	28	QDEITEREA	0.957	*T*
6	ATG5	75	EYEGTPLKW	0.682	*T*
7	ATG5	192	RIYQTTTER	0.827	*T*
8	ATG5	193	IYQTTTERP	0.597	*T*
9	ATG5	194	YQTTTERPF	0.981	*T*
10	ATG5	36	AEPYYLLLP	0.907	*Y*
11	ATG5	260	EHLSYPDNF	0.506	*Y*

Table 7
Predicted ubiquitination sites in ATG5.

S. no.	Name	Peptide	Position	Score
1	ATG5	***MTDD**K**DVLRDVW	5	0.93
2	ATG5	TLVTDKV**K**KHFQKVM	53	0.82
3	ATG5	KVKKHFQ**K**VMRQEDI	58	1.23
4	ATG5	DLLHCPS**K**DAIEAHF	118	0.59
5	ATG5	AHFMSCM**K**EADALKH	130	0.83
6	ATG5	MKEADAL**K**HKSQVIN	136	1.00
7	ATG5	EADALKH**K**SQVINEM	138	1.45

Asterisks represent absence of amino acid in sequence. BDM-PUB tool use 15 amino acids for prediction with lysine in the center and for peptide with less than 7 amino acids at either N or C terminal it uses "*".

peptide score) and Y-score (combined cleavage site score), less than the standard value (0.5) indicating no signal peptide existing in ATG5 sequence. In autophagy, the feature that helps in final recruitment of ATG12–ATG5/ATG16L1 complex to the PAS is yet to be identified. TECPR1 has recently been demonstrated to recruit ATG5-ATG12 conjugate and mediate in autophagosome maturation in mammals (Chen et al., 2012).

3.5. Regulatory elements and over-represented TFBS in ATG5

The web based server DiRE was used to identify the distant regulatory elements (REs) in the *ATG5* gene. The program was run with the default value for random set of genes as 5000. Through this analysis, two potential regulatory elements were found in *ATG5* gene. One intergenic and other is the untranslated region (3′UTR); both correspond to 50% of total regulatory region in *ATG5* as shown in Fig. 5a. A total of 44 candidate transcription factors (TFs) were also predicted in the UTR and intergenic region of *ATG5*. The detailed description of these REs (chromosome location, type, score, locus and TFs) is shown in Table S2. The top ten transcription factors with their rates of occurrence and importance are shown Fig. 5b (Graphical representation) and Table 2. Based on high importance rate (1) and occurrence rate (100%), interferon stimulated response element (ISRE) is of utmost significance when compared to the rest of the TFs. The Interferon-Sensitive Responsive Element (ISRE) is largely responsible for the constitutive expression of Programmed death-1(PD-1) receptor and also for the IFN-α-mediated up regulation of PD-1 (Cho et al., 2008). Also, there exists a link between autophagy process and innate immune signaling against viral infection where in Atg5–Atg12 conjugate negatively regulates the type I IFN production pathway promoting RNA virus replication in host cells (Jounai et al., 2007). oPOSSUM tool, which determines the TFBS families employing two statistical measures-the Z-score and Fisher score, was applied to find the over represented TFBS in the promoter region of *ATG5*. The predicted 16 TFs bind to a total of 19 sites are shown in Table S3. These were recognized by standard search parameters (>8 bits; >75% as threshold of position specific scoring matrices) of JASPAR (Fig. 6). The class wise representation of these TFs is shown in Table 3 and majority of over-represented TFBS in *ATG5* belong to Zipper-type class. This information is thus very helpful in unraveling the different ways to modulate the expression of *ATG5* and in understanding the genetic regulation of autophagy. The TFs and TFBS which directly influence the expression of *ATG5* can be targeted to achieve desired alterations. Number of transcription factors like ATF4, CHOP, CHF, FOXO1, GATA4, p63, p73 and SMADs have been shown to target *ATG5*

and regulate stress induced autophagy (Pietrocola et al., 2013). Therefore, prediction and annotation of RE, TFs and TFBS in a gene becomes very essential in order to understand their role in genetic regulation of a process.

3.6. Linkage disequilibrium analysis reveals significant genetic variants, haplotypes and predicted deleterious nsSNP in ATG5

For genome-phenotype investigation, groups of genetic variants that are acquired together in linkage disequilibrium (LD) are especially valuable. Thus, we have used other *in silico* method to identify such genetic variants and haplotypes in ATG5 that exist in linkage disequilibrium. For this, we performed haplotype analysis on SNPs genotyped in CEU population on chromosome 6 (Fig. 7a). Within the same loci on LD plot three distinct alternative blocks have been identified, and a strong correlation was seen between the blocks and there is involvement of 26 markers in first block of 38 kb of size, 39 markers in second block of 58 kb of size and 16 markers in third block of 15 kb of size, with significant statistical support. Overall correlation of 0.96 has been found among interactions between blocks 1 and 2 and correlation of 0.95 has been found among interactions between blocks 2 and 3 of the genotyped SNPs which is quite significant Fig. 7b. All important SNPs are mentioned in the Supplementary Table S4.

Because nsSNPs are coding variations and changes in the amino acid could influence protein stability, function and could have the major effect on human health in contrast to synonymous-SNPs in different locations of the genome (David and Sternberg, 2015).

According to the various SNP databases a large portion of every single genetic change identified with human disease owe to nsSNPs. Subsequently, these polymorphisms are thought to be damaging nsSNPs in light of the fact that they cause prompt phenotypic outcome (Kucukkal et al., 2015; Ohi et al., 2015). In any case, not all nsSNPs are connected with diseases. Some nsSNPs, called tolerant nsSNPs, keep up protein function despite the change in protein structure (Rodriguez-Casado, 2012). Therefore, there is a necessity to predict the deleterious nsSNPs in a protein so that these can be selected for further analysis in association studies.

We have used various *in silico* tools like Sorting Intolerant From Tolerant (SIFT), Polymorphism Phenotype (PolyPhen2), PredictSNP (PolyPhen1, nsSNPAnalyzer, PhD-SNP and MAPP) to predict the nsSNPs and their effects on the ATG5 protein function. The predicted impact of change in amino acid may be tolerated or damaging/deleterious based

Table 6
Palmitoylation sites predicted in ATG5.

S. no.	ID	Position	Peptide
1	ATG5	19	WFGRIPTCFTLYQDE
2	ATG5	128	IEAHFMSCMKEADAL

Table 8
Calpain-cleavage sites predicted in ATG5.

S. no.	Name	Peptide	Position	Score
1	ATG5	MKEADAL**K | H**KSQVIN	136	0.6615
2	ATG5	PFRIYOTT **| T**ERPFIQ	193	0.6846
3	ATG5	PDNFLHI**S | I**PQPTD	268	0.7808
4	ATG5	PDNFLHIS**I | I**PQPTD*	269	0.6500

Asterisk represents pseudo aminoacid.

on the tolerance index and probability scores. We have identified 3 nsSNPs (rs34793250, M129V; rs77859116, I65V; and rs115576116, A95D) out of which two nsSNPs rs34793250 (M129V) and rs115576116 (A95D) could have damaging/deleterious effects on the ATG5 protein (Table 4). These predicted nsSNPs could further be explored to study their impact on ATG5 protein structure and function. Moreover, these nsSNPs could be selected to carry out population based studies to find the role played by them, if any, in the susceptibility of various diseases as well as their involvement in intervening autophagy pathway. These methodologies expand the likelihood of success in recognizing polymorphisms of utmost significance in ATG5.

3.7. Numerous post-translational modification sites predicted in ATG5 and their annotation over the protein domain structure

Posttranslational modifications (PTMs) are considered to be very essential for the regulation of various cellular processes and play important role in influencing protein interaction network (Duan and Walther, 2015). Dysregulation of PTMs results in modulation of a pathway, autophagy here, that could play a role in the pathogenesis of a disease (Mastrangelo et al., 2015). A total of 11 putative phosphorylation sites were predicted in ATG5 using NetPhos algorithm (Table 5). NetPhos is a neural network-based method for predicting potential phosphorylation sites at serine (S), threonine (T) or tyrosine (Y) residues in a protein sequences. The prediction score ≥ 0.5 was considered as phosphorylated. In a recent study, the p38-mediated phosphorylation of ATG5 at the predicted threonine 75 position has been shown to lead to autophagy inhibition in a Gadd45β/MEKK4-dependent manner (Keil et al., 2013).

By using CSS-PALM server, we identified 2 putative palmitoylation sites in ATG5 as shown in Table 6. Palmitoylation enhances surface hydrophobicity and membrane affinity of protein substrates by adding palmitic acid. Palmitoylation modulates proteins trafficking, compartmentalization, and membrane-tethering of many proteins. Moreover, protein palmitoylation has been implicated in numerous cellular processes, including signaling, apoptosis, and neuronal transmission (Mingming et al., 2014). These could further be explored for their impact on functional impact on ATG5.

BDMPUB web server was used to predict ubiquitination sites in ATG5 by Bayesian Discriminant Method. Ubiquitination is a major post-translational modification that regulates several biological processes at different cellular levels, e.g., protein trafficking, cell cycle, and immune response (Zohaib et al., 2015). Dysregulation of ubiquitination has expansive consequences and has a role in the onset and progression of cancer, metabolic syndromes, neurodegenerative diseases, inflammatory disorders, autoimmunity, muscular dystrophies and infection (Doris Popovic, 2014). A total of 7 ubiquitination sites were predicted in ATG5 (Table 7).

Calpain-mediated cleavage is the most crucial post-translational modification (PTMs) of proteins that play an important role in many biological processes like cell death/apoptosis, cytoskeletal remodeling, and the cell cycle. GPS-CCD 1.0 tool was utilized for the prediction of calpain-cleavage sites in ATG5. Calpain mediated cleavage of ATG5 switches autophagy to apoptosis establishing important link between autophagy and apoptosis and highlights the importance of calpain activation and cleavage (Yousefi et al., 2006). A total of 4 Calpain–cleavage sites were predicted in ATG5 as shown in Table 8. Also, the overall distribution of these predicted PTM sites were represented over the protein domain structure with the help of Illustrator for Biological Sequences (IBS) tool (Liu et al., 2015) is shown in Fig. 8.

3.8. ATG5 interactions with other proteins

In order to explore the function of a protein it is very essential to consider its interactions with other proteins. The STRING database provides a significant assessment and integration of protein–protein interactions which includes direct as well as indirect relations. The main functional partners of ATG5 are shown in Table S5 along with their scores and the STRING network view is shown in Fig. 9. These proteins interact with ATG5 and mediate the process of autophagy as well as other linked pathways. Moreover, aberrations in these interactions may lead to pathogenic conditions (Mai et al., 2012; Pareja and Colombo, 2013; Randall-Demllo et al., 2013; Ryter et al., 2014; Vural and Kehrl, 2014). Furthermore, studies on these protein interactions may provide imperative knowledge about mechanisms, physiology and pathology not only within but also beyond autophagy.
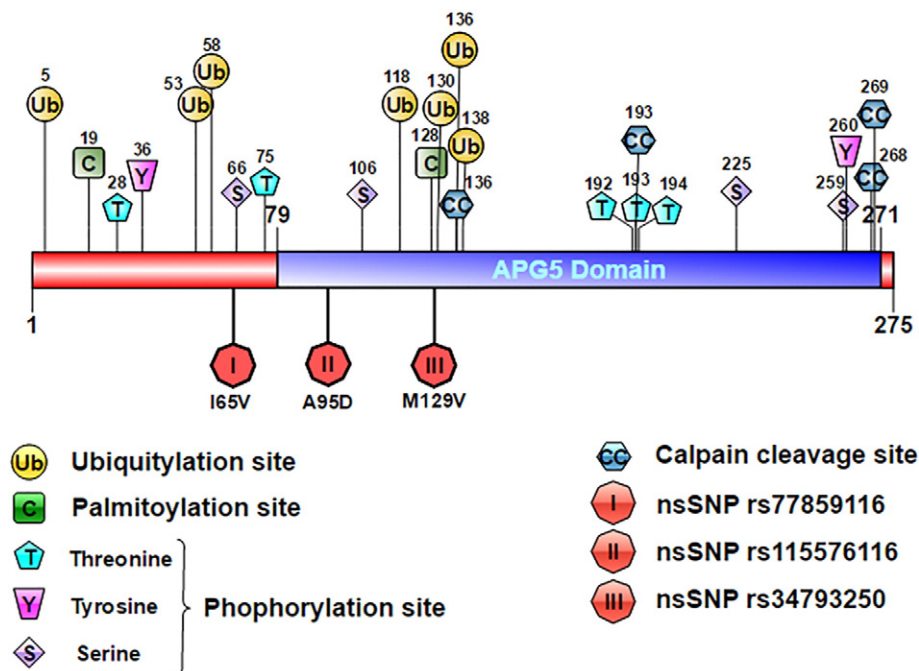

Fig. 8. Putative functional sites in the ATG5 protein. Schematic diagram representing the approximate location of the predicted PTM sites (phosphorylation, ubiquitylation, palmitoylation, and Calpain-cleavage sites) and the 3 putative nsSNPs in ATG5.
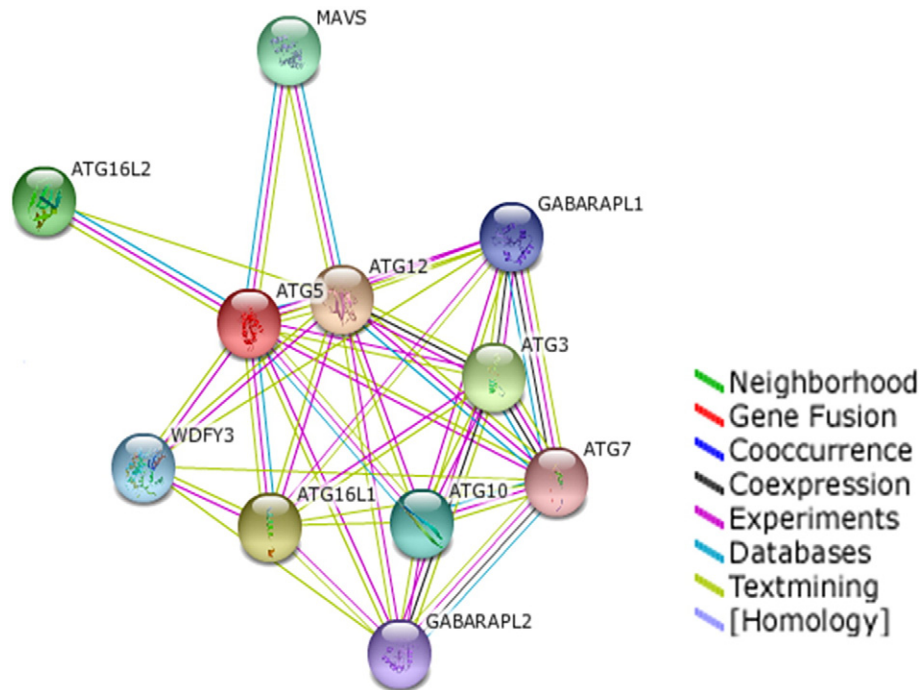
Fig. 9. The STRING network view for ATG5 representing its various functional partners. Colored lines between the proteins indicate the various types of interaction evidence.

4. Conclusion

Autophagy, an essential regulatory mechanism, helps to renovate the cell cytosol by removing damaged proteins and organelles. The involvement of autophagy in various diseases underscores its importance and recent data regarding its therapeutic potential utter the need for systematic understanding of this pathway. This is possible only if extensive study is carried out to understand each component that drives the process so that it could be targeted to modulate for the alleviation of various disorders. ATG5 is a core autophagy gene that plays a major role in its regulation. In this study, we have investigated its evolutionary links and various physico-chemical parameters that help to reveal different physiological properties of the gene. Though no signal peptide was found in the protein but two possible models for transmembrane helices were predicted that served to recognize functional regions of the protein. We have identified 24 putative post-translational modification sites, 44 important TFBS which are distributed throughout the protein. Majority of these TFBS belonged to Zipper-type class of TFs. ISRE, has been found to be the most significant TF based on its occurrence and importance rates. Besides this LD analysis uncovered 3 distinct alternative blocks of the genotyped SNPs of 38 kb, 58 kb and 15 kb of size. These blocks showed strong correlation among them and involved 16, 26 and 39 markers respectively, with significant statistical support. We have also predicted 2 deleterious nsSNPs (rs rs34793250, rs 115576116) that might have damaging impact on protein function and ultimately contribute to an individual's susceptibility to various diseases. Overall, this comprehensive *in silico* analysis provides information that helps unfolding significant physiological properties and regulatory activity of ATG5.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.mgene.2016.07.012.

Conflict of interest

The authors declare that they do not have any conflict of interest.

References

Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., Sunyaev, S.R., 2010. A method and server for predicting damaging missense mutations. Nat. Methods 7, 248–249.

Alirezaei, M., Fox, H.S., Flynn, C.T., Moore, C.S., Hebb, A.L., Fausto, R.F., Bhan, V., Kiosses, W.B., Whitton, J.L., Robertson, G.S., Crocker, S.J., 2009. Elevated ATG5 expression in autoimmune demyelination and multiple sclerosis. Autophagy 5, 152–158.

Badadani, M., 2012. Autophagy mechanism, regulation, functions, and disorders. ISRN Cell Biol. 2012 11 pages.

Baerga, R., Zhang, Y., Chen, P.H., Goldman, S., Jin, S., 2009. Targeted deletion of autophagy-related 5 (atg5) impairs adipogenesis in a cellular model and in mice. Autophagy 5, 1118–1130.

Bendl, J., Stourac, J., Salanda, O., Pavelka, A., Wieben, E.D., Zendulka, J., Brezovsky, J., Damborsky, J., 2014. PredictSNP: robust and accurate consensus classifier for prediction of disease-related mutations. PLoS Comput. Biol. 10, e1003440.

Blom, N., Gammeltoft, S., Brunak, S., 1999. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. J. Mol. Biol. 294, 1351–1362.

Bock, J.R., Gough, D.A., 2001. Predicting protein–protein interactions from primary structure. Bioinformatics 17, 455–460.

Bojkowska, K., Santoni de Sio, F., Barde, I., Offner, S., Verp, S., Heinis, C., Johnsson, K., Trono, D., 2011. Measuring in vivo protein half-life. Chem. Biol. 18, 805–815.

Chen, D., Fan, W., Lu, Y., Ding, X., Chen, S., Zhong, Q., 2012. A mammalian autophagosome maturation mechanism mediated by TECPR1 and the Atg12-Atg5 conjugate. Mol. Cell 45, 629–641.

Chen, D., Zhu, C., Wang, X., Feng, X., Pang, S., Huang, W., Hawley, R.G., Yan, B., 2013. A novel and functional variant within the ATG5 gene promoter in sporadic Parkinson's disease. Neurosci. Lett. 538, 49–53.

Cheng, Y., Ren, X., Hait, W.N., Yang, J.M., 2013. Therapeutic targeting of autophagy in disease: biology and pharmacology. Pharmacol. Rev. 65, 1162–1197.

Chiramel, A.I., Brady, N.R., Bartenschlager, R., 2013. Divergent roles of autophagy in virus infection. Cell 2, 83–104.

Cho, H.Y., Lee, S.W., Seo, S.K., Choi, I.W., Choi, I., 2008. Interferon-sensitive response element (ISRE) is mainly responsible for IFN-alpha-induced upregulation of programmed death-1 (PD-1) in macrophages. Biochim. Biophys. Acta 1779, 811–819.

Cho, D.H., Jo, Y.K., Kim, S.C., Park, I.J., Kim, J.C., 2012. Down-regulated expression of ATG5 in colorectal cancer. Anticancer Res. 32, 4091–4096.

Craveur, P., Joseph, A.P., Esque, J., Narwani, T.J., Noel, F., Shinada, N., Goguet, M., Leonard, S., Poulain, P., Bertrand, O., Faure, G., Rebehmed, J., Ghozlane, A., Swapna, L.S., Bhaskara, R.M., Barnoud, J., Teletchea, S., Jallu, V., Cerny, J., Schneider, B., Etchebest, C., Srinivasan, N., Gelly, J.C., de Brevern, A.G., 2015. Protein flexibility in the light of structural alphabets. Front. Mol. Biosci. 2, 20.

Cursio, R., Colosetti, P., Codogno, P., Cuervo, A.M., Shen, H.M., 2015. The role of autophagy in liver diseases: mechanisms and potential therapeutic targets. Biomed Res Int 2015, 480508.

David, A., Sternberg, M.J., 2015. The contribution of missense mutations in core and rim residues of protein-protein interfaces to human disease. J. Mol. Biol.

Ding, Y., Choi, M.E., 2015. Autophagy in diabetic nephropathy. J. Endocrinol. 224, R15–R30.

Doris Popovic, D.V.I.D., 2014. Ubiquitination in disease pathogenesis and treatment. Nat. Med. 20.

Duan, G., Walther, D., 2015. The roles of post-translational modifications in the context of protein interaction networks. PLoS Comput. Biol. 11, e1004049.

Emanuelsson, O., Brunak, S., von Heijne, G., Nielsen, H., 2007. Locating proteins in the cell using TargetP, SignalP and related tools. Nat. Protoc. 2, 953–971.

Gasteiger, E., H., C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., Bairoch, A., 2005. Protein identification and analysis tools on the ExPASy Server. In: Walker, J.M. (Ed.), The Proteomics Protocols Handbook. Humana Press.

Gateva, V., Sandling, J.K., Hom, G., Taylor, K.E., Chung, S.A., Sun, X., Ortmann, W., Kosoy, R., Ferreira, R.C., Nordmark, G., Gunnarsson, I., Svenungsson, E., Padyukov, L., Sturfelt, G., Jonsen, A., Bengtsson, A.A., Rantapaa-Dahlqvist, S., Baechler, E.C., Brown, E.E., Alarcon, G.S., Edberg, J.C., Ramsey-Goldman, R., McGwin Jr., G., Reveille, J.D., Vila, L.M., Kimberly, R.P., Manzi, S., Petri, M.A., Lee, A., Gregersen, P.K., Seldin, M.F., Ronnblom, L., Criswell, L.A., Syvanen, A.C., Behrens, T.W., Graham, R.R., 2009. A large-scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus. Nat. Genet. 41, 1228–1233.

Ghavami, S., Shojaei, S., Yeganeh, B., Ande, S.R., Jangamreddy, J.R., Mehrpour, M., Christoffersson, J., Chaabane, W., Moghadam, A.R., Kashani, H.H., Hashemi, M., Owji, A.A., Los, M.J., 2014. Autophagy and apoptosis dysfunction in neurodegenerative disorders. Prog. Neurobiol. 112, 24–49.

Glick, D., Barth, S., Macleod, K.F., 2010. Autophagy: cellular and molecular mechanisms. J. Pathol. 221, 3–12.

Gotea, V., Ovcharenko, I., 2008. DiRE: identifying distant regulatory elements of co-expressed genes. Nucleic Acids Res. 36, W133–W139.

Guruprasad, K., Reddy, B.V., Pandit, M.W., 1990. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. Protein Eng. 4, 155–161.

Huett, A., Goel, G., Xavier, R.J., 2010. A systems biology viewpoint on autophagy in health and disease. Curr. Opin. Gastroenterol. 26, 302–309.

Hwang, S., Maloney, N.S., Bruinsma, M.W., Goel, G., Duan, E., Zhang, L., Shrestha, B., Diamond, M.S., Dani, A., Sosnovtsev, S.V., Green, K.Y., Lopez-Otin, C., Xavier, R.J., Thackray, L.B., Virgin, H.W., 2012. Nondegradative role of Atg5-Atg12/ Atg16L1 autophagy protein complex in antiviral activity of interferon gamma. Cell Host Microbe 11, 397–409.

Ikai, A., 1980. Thermostability and aliphatic index of globular proteins. J. Biochem. 88, 1895–1898.

Jounai, N., Takeshita, F., Kobiyama, K., Sawano, A., Miyawaki, A., Xin, K.Q., Ishii, K.J., Kawai, T., Akira, S., Suzuki, K., Okuda, K., 2007. The Atg5 Atg12 conjugate associates with innate antiviral immune responses. Proc. Natl. Acad. Sci. U. S. A. 104, 14050–14055.

Keil, E., Hocker, R., Schuster, M., Essmann, F., Ueffing, N., Hoffman, B., Liebermann, D.A., Pfeffer, K., Schulze-Osthoff, K., Schmitz, I., 2013. Phosphorylation of Atg5 by the Gadd45beta-MEKK4-p38 pathway inhibits autophagy. Cell Death Differ. 20, 321–332.

Kiel, J.A., 2010. Autophagy in unicellular eukaryotes. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 365, 819–830.

Kucukkal, T.G., Petukh, M., Li, L., Alexov, E., 2015. Structural and physico-chemical effects of disease and non-disease nsSNPs on proteins. Curr. Opin. Struct. Biol. 32, 18–24.

Kuma, A., Hatano, M., Matsui, M., Yamamoto, A., Nakaya, H., Yoshimori, T., Ohsumi, Y., Tokuhisa, T., Mizushima, N., 2004. The role of autophagy during the early neonatal starvation period. Nature 432, 1032–1036.

Kumar, P., Henikoff, S., Ng, P.C., 2009. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat. Protoc. 4, 1073–1081.

Kwon, A.T., Arenillas, D.J., Worsley Hunt, R., Wasserman, W.W., 2012. oPOSSUM-3: advanced analysis of regulatory motif over-representation across genes or ChIP-Seq datasets. G3 (Bethesda) 2, 987–1002.

Kyte, J., Doolittle, R.F., 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157, 105–132.

Lee, H.K., Mattei, L.M., Steinberg, B.E., Alberts, P., Lee, Y.H., Chervonsky, A., Mizushima, N., Grinstein, S., Iwasaki, A., 2010. In vivo requirement for Atg5 in antigen presentation by dendritic cells. Immunity 32, 227–239.

Levine, B., Yuan, J., 2005. Autophagy in cell death: an innocent convict? J. Clin. Invest. 115, 2679–2688.

Liu, Z., Cao, J., Gao, X., Ma, Q., Ren, J., Xue, Y., 2011. GPS-CCD: a novel computational program for the prediction of calpain cleavage sites. PLoS One 6, e19001.

Liu, W., Xie, Y., Ma, J., Luo, X., Nie, P., Zuo, Z., Lahrmann, U., Zhao, Q., Zheng, Y., Zhao, Y., Xue, Y., Ren, J., 2015. IBS: an illustrator for the presentation and visualization of biological sequences. Bioinformatics 31, 3359–3361.

Lv, X., Jiang, H., Li, B., Liang, Q., Wang, S., Zhao, Q., Jiao, J., 2014. The crucial role of Atg5 in cortical neurogenesis during early brain development. Sci. Rep. 4, 6010.

Mai, S., Muster, B., Bereiter-Hahn, J., Jendrach, M., 2012. Autophagy proteins LC3B, ATG5 and ATG12 participate in quality control after mitochondrial damage and influence lifespan. Autophagy 8, 47–62.

Martin, L.J., Gupta, J., Jyothula, S.S., Butsch Kovacic, M., Biagini Myers, J.M., Patterson, T.L., Ericksen, M.B., He, H., Gibson, A.M., Baye, T.M., Amirisetty, S., Tsoras, A.M., Sha, Y., Eissa, N.T., Hershey, G.K., 2012. Functional variant in the autophagy-related 5 gene promotor is associated with childhood asthma. PLoS One 7, e33454.

Maskey, D., Yousefi, S., Schmid, I., Zlobec, I., Perren, A., Friis, R., Simon, H.U., 2013. ATG5 is induced by DNA-damaging agents and promotes mitotic catastrophe independent of autophagy. Nat. Commun. 4, 2130.

Maston, G.A., Evans, S.K., Green, M.R., 2006. Transcriptional regulatory elements in the human genome. Annu. Rev. Genomics Hum. Genet. 7, 29–59.

Mastrangelo, A., Colasanti, T., Barbati, C., Pecani, A., Sabatinelli, D., Pendolino, M., Truglia, S., Massaro, L., Mancini, R., Miranda, F., Spinelli, F.R., Conti, F., Alessandri, C., 2015. The role of posttranslational protein modifications in rheumatological diseases: focus on rheumatoid arthritis. J. Immunol. Res. 2015, 712490.

Mathelier, A., Zhao, X., Zhang, A.W., Parcy, F., Worsley-Hunt, R., Arenillas, D.J., Buchman, S., Chen, C.Y., Chou, A., Ienasescu, H., Lim, J., Shyr, C., Tan, G., Zhou, M., Lenhard, B., Sandelin, A., Wasserman, W.W., 2014. JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. Nucleic Acids Res. 42, D142–D147.

Mayes, M.D., Bossini-Castillo, L., Gorlova, O., Martin, J.E., Zhou, X., Chen, W.V., Assassi, S., Ying, J., Tan, F.K., Arnett, F.C., Reveille, J.D., Guerra, S., Teruel, M., Carmona, F.D., Gregersen, P.K., Lee, A.T., Lopez-Isac, E., Ochoa, E., Carreira, P., Simeon, C.P., Castellvi, I., Gonzalez-Gay, M.A., Zhernakova, A., Padyukov, L., Alarcon-Riquelme, M., Wijmenga, C., Brown, M., Beretta, L., Riemekasten, G., Witte, T., Hunzelmann, N., Kreuter, A., Distler, J.H., Voskuyl, A.E., Schuerwegh, A.J., Hesselstrand, R., Nordin, A., Airo, P., Lunardi, C., Shiels, P., van Laar, J.M., Herrick, A., Worthington, J., Denton, C., Wigley, F.M., Hummers, L.K., Varga, J., Hinchcliff, M.E., Baron, M., Hudson, M., Pope, J.E., Furst, D.E., Khanna, D., Phillips, K., Schiopu, E., Segal, B.M., Molitor, J.A., Silver, R.M., Steen, V.D., Simms, R.W., Lafyatis, R.A., Fessler, B.J., Frech, T.M., Alkassab, F., Docherty, P., Kaminska, E., Khalidi, N., Jones, H.N., Markland, J., Robinson, D., Broen, J., Radstake, T.R., Fonseca, C., Koeleman, B.P., Martin, J., 2014. Immunochip analysis identifies multiple susceptibility loci for systemic sclerosis. Am. J. Hum. Genet. 94, 47–61.

Miller, B.C., Zhao, Z., Stephenson, L.M., Cadwell, K., Pua, H.H., Lee, H.K., Mizushima, N.N., Iwasaki, A., He, Y.W., Swat, W., Virgin, H.W.t., 2008. The autophagy gene ATG5 plays an essential role in B lymphocyte development. Autophagy 4, 309–314.

Mingming, L., .F.a, P., .W ., Y., 2014. Lipidomics in health and diseases - beyond the analysis of lipids. Glycomics Lipidomics 5.

Mizushima, N., 2007. Autophagy: process and function. Genes Dev. 21, 2861–2873.

Mukherjee, R.M., VS ., G., Balkumar Reddy, P., Rao, P.N., Reddy, D.N., 2014. Escalated expression of autophagy related gene Atg5 in hepatitis B virus infection. J. Microbiol. Res. Rev. 2, 68–73.

Nagelkerke, A., Sieuwerts, A.M., Bussink, J., Sweep, F.C., Look, M.P., Foekens, J.A., Martens, J.W., Span, P.N., 2014. LAMP3 is involved in tamoxifen resistance in breast cancer cells through the modulation of autophagy. Endocr. Relat. Cancer 21, 101–112.

Ohi, K., Ursini, G., Li, M., Shin, J.H., Ye, T., Chen, Q., Tao, R., Kleinman, J.E., Hyde, T.M., Hashimoto, R., Weinberger, D.R., 2015. DEGS2 polymorphism associated with cognition in schizophrenia is associated with gene expression in brain. Transl. Psychiatry 5, e550.

Ohsumi, Y., 2014. Historical landmarks of autophagy research. Cell Res. 24, 9–23.

Orozco, G., Eyre, S., Hinks, A., Bowes, J., Morgan, A.W., Wilson, A.G., Wordsworth, P., Steer, S., Hocking, L., Thomson, W., Worthington, J., Barton, A., 2011. Study of the common genetic background for rheumatoid arthritis and systemic lupus erythematosus. Ann. Rheum. Dis. 70, 463–468.

Otomo, C., Metlagel, Z., Takaesu, G., Otomo, T., 2013. Structure of the human ATG12 ~ ATG5 conjugate required for LC3 lipidation in autophagy. Nat. Struct. Mol. Biol. 20, 59–66.

Pareja, E., Colombo, M.I., 2013. Autophagic clearance of bacterial pathogens: molecular recognition of intracellular microorganisms. Front. Cell Infect. Microbiol. 3, 54.

Petersen, T.N., Brunak, S., von Heijne, G., Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods 8, 785–786.

Pietrocola, F., Izzo, V., Niso-Santano, M., Vacchelli, E., Galluzzi, L., Maiuri, M.C., Kroemer, G., 2013. Regulation of autophagy by stress-responsive transcription factors. Semin. Cancer Biol. 23, 310–322.

Plantinga, T.S., van de Vosse, E., Huijbers, A., Netea, M.G., Joosten, L.A., Smit, J.W., Netea-Maier, R.T., 2014. Role of genetic variants of autophagy genes in susceptibility to non-medullary thyroid cancer and patients outcome. PLoS One 9, e94086.

Pua, H.H., Dzhagalov, I., Chuck, M., Mizushima, N., He, Y.W., 2007. A critical role for the autophagy gene Atg5 in T cell survival and proliferation. J. Exp. Med. 204, 25–31.

Pyo, J.O., Jang, M.H., Kwon, Y.K., Lee, H.J., Jun, J.I., Woo, H.N., Cho, D.H., Choi, B., Lee, H., Kim, J.H., Mizushima, N., Oshumi, Y., Jung, Y.K., 2005. Essential roles of Atg5 and FADD in autophagic cell death: dissection of autophagic cell death into vacuole formation and cell death. J. Biol. Chem. 280, 20722–20729.

Randall-Demllo, S., Chieppa, M., Eri, R., 2013. Intestinal epithelium and autophagy: partners in gut homeostasis. Front. Immunol. 4, 301.

Randhawa, R., Sehgal, M., Singh, T.R., Duseja, A., Changotra, H., 2015. Unc-51 like kinase 1 (ULK1) in silico analysis for biomarker identification: a vital component of autophagy. Gene 562, 40–49.

Ren, J., Wen, L., Gao, X., Jin, C., Xue, Y., Yao, X., 2008. CSS-palm 2.0: an updated software for palmitoylation sites prediction. Protein Eng. Des. Sel. 21, 639–644.

Rodriguez-Casado, A., 2012. In silico investigation of functional nsSNPs –an approach to rational drug design. Res. Reports Med. Chem. 31–42.

Romanov, J., Walczak, M., Ibiricu, I., Schuchner, S., Ogris, E., Kraft, C., Martens, S., 2012. Mechanism and functions of membrane binding by the Atg5-Atg12/Atg16 complex during autophagosome formation. EMBO J. 31, 4304–4317.

Rubinsztein, D.C., Codogno, P., Levine, B., 2012. Autophagy modulation as a potential therapeutic target for diverse diseases. Nat. Rev. Drug Discov. 11, 709–730.

Ryter, S.W., Mizumura, K., Choi, A.M., 2014. The impact of autophagy on cell death modalities. Int. J. Cell Biol. 2014, 502676.

Shoji-Kawata, S., Sumpter, R., Leveno, M., Campbell, G.R., Zou, Z., Kinch, L., Wilkins, A.D., Sun, Q., Pallauf, K., MacDuff, D., Huerta, C., Virgin, H.W., Helms, J.B., Eerland, R., Tooze, S.A., Xavier, R., Lenschow, D.J., Yamamoto, A., King, D., Lichtarge, O., Grishin, N.V., Spector, S.A., Kaloyanova, D.V., Levine, B., 2013. Identification of a candidate therapeutic autophagy-inducing peptide. Nature 494, 201–206.

Sir, D., Tian, Y., Chen, W.L., Ann, D.K., Yen, T.S., Ou, J.H., 2010. The early autophagic pathway is activated by hepatitis B virus and required for viral DNA replication. Proc. Natl. Acad. Sci. U. S. A. 107, 4383–4388.

Stephenson, L.M., Miller, B.C., Ng, A., Eisenberg, J., Zhao, Z., Cadwell, K., Graham, D.B., Mizushima, N.N., Xavier, R., Virgin, H.W., Swat, W., 2009. Identification of Atg5-dependent transcriptional changes and increases in mitochondrial mass in Atg5-deficient T lymphocytes. Autophagy 5, 625–635.

Stoffel, K.H.W., 1993. TMbase - A database of membrane spanning proteins segments. Biol. Chem. 374 (166) (Hoppe-Seyler).

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K.P., Kuhn, M., Bork, P., Jensen, L.J., von Mering, C., 2015. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res. 43, D447–D452.

Takamura, A., Komatsu, M., Hara, T., Sakamoto, A., Kishi, C., Waguri, S., Eishi, Y., Hino, O., Tanaka, K., Mizushima, N., 2011. Autophagy-deficient mice develop multiple liver tumors. Genes Dev. 25, 795–800.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.

Tang, Z., Lin, M.G., Stowe, T.R., Chen, S., Zhu, M., Stearns, T., Franco, B., Zhong, Q., 2013. Autophagy promotes primary ciliogenesis by removing OFD1 from centriolar satellites. Nature 502, 254–257.

Tanida, I., 2011. Autophagosome formation and molecular mechanism of autophagy. Antioxid. Redox Signal. 14, 2201–2214.

Teng, S., Srivastava, A.K., Wang, L., 2010. Sequence feature-based prediction of protein stability changes upon amino acid substitutions. BMC Genomics 11 (Suppl. 2), S5.

Thorisson, G.A., Smith, A.V., Krishnan, L., Stein, L.D., 2005. The international HapMap project web site. Genome Res. 15, 1592–1593.

Tobias, J.W., Shrader, T.E., Rocap, G., Varshavsky, A., 1991. The N-end rule in bacteria. Science 254, 1374–1377.

Vural, A., Kehrl, J.H., 2014. Autophagy in macrophages: impacting inflammation and bacterial infection. Scientifica (Cairo) 2014, 825463.

Walczak, M., Martens, S., 2013. Dissecting the role of the Atg12-Atg5-Atg16 complex during autophagosome formation. Autophagy 9, 424–425.

White, S.H., 1992. Amino acid preferences of small proteins. Implications for protein stability and evolution. J. Mol. Biol. 227, 991–995.

Wilkins, M.R., Gasteiger, E., Bairoch, A., Sanchez, J.C., Williams, K.L., Appel, R.D., Hochstrasser, D.F., 1999. Protein identification and analysis tools in the ExPASy server. Methods Mol. Biol. 112, 531–552.

Winslow, A.R., Rubinsztein, D.C., 2008. Autophagy in neurodegeneration and development. Biochim. Biophys. Acta 1782, 723–729.

Wirawan, E., Vanden Berghe, T., Lippens, S., Agostinis, P., Vandenabeele, P., 2012. Autophagy: for better or for worse. Cell Res. 22, 43–61.

Yang, Z.J., Chee, C.E., Huang, S., Sinicrope, F.A., 2011. The role of autophagy in cancer: therapeutic implications. Mol. Cancer Ther. 10, 1533–1541.

Yousefi, S., Perozzo, R., Schmid, I., Ziemiecki, A., Schaffner, T., Scapozza, L., Brunner, T., Simon, H.U., 2006. Calpain-mediated cleavage of Atg5 switches autophagy to apoptosis. Nat. Cell Biol. 8, 1124–1132.

Zhao, Z., Fux, B., Goodwin, M., Dunay, I.R., Strong, D., Miller, B.C., Cadwell, K., Delgado, M.A., Ponpuak, M., Green, K.G., Schmidt, R.E., Mizushima, N., Deretic, V., Sibley, L.D., Virgin, H.W., 2008. Autophagosome-independent essential function for the autophagy protein Atg5 in cellular immunity to intracellular pathogens. Cell Host Microbe 4, 458–469.

Zimmerman, J.M., Eliezer, N., Simha, R., 1968. The characterization of amino acid sequences in proteins by statistical methods. J. Theor. Biol. 21, 170–201.

Zohaib, A., Duan, X., Zhu, B., Ye, J., Wan, S., Chen, H., Liu, X., Cao, S., 2015. The role of ubiquitination in regulation of innate immune signaling. Curr. Issues Mol. Biol. 18, 1–10.