TEX264 at the Intersection of Autophagy and DNA repair

John Fielden[1#], Marta Popović[2#], Kristijan Ramadan[1*]

[1]Medical Research Council (MRC) Oxford Institute for Radiation Oncology, Department of Oncology, University of Oxford, Roosevelt Drive, Oxford, OX3 7DQ, UK, [2]Ruđer Bošković Institute, Bijenička cesta 54, 10000 Zagreb, Croatia

[#]Equal contribution

*Correspondence to Kristijan Ramadan; E-mail: kristijan.ramadan@oncology.ox.ac.uk

Key words:

Autophagy, DNA repair, ER-phagy, gyrase inhibitory-like proteins, nucleophagy, TEX264

Abbreviations

ATG8, autophagy related 8; Cdc48, cell division cycle 48; DPC, DNA-protein crosslinks; DSB, DNA double-strand break; ER, endoplasmic reticulum, GyrI, gyrase inhibitory domain; INM, inner nuclear membrane; LRR, leucine-rich repeat; MAP1LC3/LC3, microtubule-associated protein 1 light chain 3; STUBL, SUMO targeted ubiquitin ligase; SUMO, small ubiquitin like modifier; TEX264, testis expressed gene 264; Top1-cc, topoisomerase 1-cleavage complex; UBZ, ubiquitin binding Zn finger domain; VCP/p97, valosin containing protein

Abstract

TEX264 (testes expressed gene 264) is a single-pass transmembrane protein, consisting of an N-terminal hydrophobic region, a gyrase inhibitory (GyrI)-like domain, and a loosely structured C terminus. TEX264 was first identified as an endoplasmic reticulum (ER)-resident Atg8-family binding protein that mediates the degradation of portions of the ER during starvation (reticulophagy). More recently, TEX264 was identified as a cofactor of VCP/p97 ATPase (Cdc48 in yeast) that promotes the repair of covalently trapped TOP1 (DNA topoisomerase 1)-DNA crosslinks. This review summarizes our current knowledge of TEX264 and provides an evolutionary and structural analysis of GyrI proteins. Based on our phylogenetic analysis, we provide evidence that TEX264 is a member of a large superfamily of GyrI-like proteins that evolved in bacteria and are present in metazoans, including invertebrates and chordates. This review summarizes our current knowledge of TEX264 as a genuine factor that directly bridges DNA repair and autophagy, and provides an evolutionary and structural analysis of GyrI proteins.

Introduction

Protein homeostasis is essential for cellular viability. The two major branches of protein homeostasis are autophagy and ubiquitin-dependent proteasomal degradation, which share notable similarities. Both processes enable cells to dispose of excess, aggregated, and damaged organelles or proteins. Dedicated receptor proteins target specific cargo/substrates to facilitate their trafficking into autophagosomes or their presentation to proteasome, and often recognise ubiquitin chains on cargo/substrates [1]. Moreover, certain chaperones, such as the VCP/p97 ATPase, have critical roles in both degradative processes [2].

The autophagic degradation of portions of the ER has recently been recognised as an important response to nutrient deprivation and the accumulation of misfolded ER lumenal proteins [3]. Reticulophagy is mediated by receptor proteins that are tethered to the ER membrane and bind LC3/GABARAP proteins on phagophore membranes. The Gyrase inhibitory (GyrI)-like domain-containing protein TEX264 was recently shown to be a major ER-phagy receptor which is sequestered by phagophores via its interaction with LC3-family members and mediates the autophagic degradation of many ER membrane and lumenal proteins upon starvation [4,5]. More recently, we identified an important role of TEX264 as a substrate adaptor at the inner nuclear membrane (INM), where it helps to preserve genome stability [6]. As a cofactor of the VCP/p97 ATPase, TEX264 promotes the degradation of DNA lesions known as TOP1 (DNA topoisomerase 1) cleavage complexes (TOP1cc), which are composed of TOP1 covalently bound to a single-stranded DNA break, and its evolutionarily conserved GyrI-like domain is critical for this function.

Overall, three recent papers have reported distinct functions of the TEX264 protein with a common theme, whereby TEX264 acts as a membrane-anchored receptor to promote the degradation of ER proteins during ER-phagy or of nuclear substrates during DNA repair. This raises fascinating questions regarding potential overlap between these roles. Here, we review the recent reports on TEX264, discuss its potential role in bridging DNA repair and autophagy, and provide a phylogenetic and structural analysis of the GyrI superfamily of proteins.

TEX264 in ER-phagy

The first reports on TEX264 revealed its critical role as a receptor for ER-phagy, a process by which portions of the ER are sequestered into autophagosomes during nutrient deprivation. Chino *et al.* identified TEX264 in a mass spectrometry analysis of proteins that interact preferentially with wild-type LC3B versus a LC3-interacting region (LIR) binding-defective

variant [4]. An *et al*. meanwhile, identified TEX264 in a global quantitative proteome analysis of proteins whose abundance is decreased upon either MTOR inhibition or amino acid deprivation in an ATG7- and RB1CC1-dependent manner [5]. TEX264 was shown to undergo trafficking from the ER to lysosomes upon nutrient deprivation which was dependent both on canonical autophagy pathway components and a LIR motif in TEX264's C terminus. The long, intrinsically disordered nature of TEX264's C terminus is also crucial for its ER-phagy function. Due to their large size, ribosomes on the ER membrane may prevent the direct association of the ER and phagophore membranes; however, TEX264's long C terminus bridges this spatial gap by extending into the cytosol and binding LC3 on phagophores [4].

Of the seven known mammalian ER-phagy receptors, TEX264 appears to play a major role in regulating ER-phagy flux [7–13]. By comparing the effects of individually depleting TEX264 and other known ER-phagy receptors, it was observed that TEX264 knockdown most dramatically suppressed ER-phagy in HeLa cells [4]. Similarly, based on global quantitative proteome mass spectrometry, it was estimated that approximately 50% of all ER-phagy flux upon starvation is driven by TEX264 in HEK293T cells [5]. A more recent genome-wide CRISPR interference screen of ER-phagy regulators found only a modest reduction in ER-phagy activity upon TEX264 knockdown, which is consistent with there being, at least partial, functional redundancy between different ER-phagy receptors [14]. The extent of ER-phagy flux and the impact of the different receptors may vary between tissues and cell types and could be influenced by the differential expression of ER-phagy receptors, with TEX264 appearing to be the most broadly expressed [4].

During nutrient deprivation, TEX264 loss stabilizes many ER membrane and luminal proteins but does not affect others [5]. This raises important questions as to how TEX264 achieves cargo specificity. One possibility is that the sub-regional differences in TEX264 expression or activation on the ER membrane regulates the differential turnover of cargo. Another is that

specific interactions between TEX264 and proteins on the luminal side of the inner ER membrane enables selective protein degradation [15]. The UFL1 ligase was recently shown to be required for the autophagic degradation of ER sheets [14]. UFL1 is recruited to the ER membrane by DDRGK1, where it UFMylates the oligosaccharyltransferase (OST) complex subunit, RPN1, and the ribosomal protein, RPL26. Depletion of DDRGK1 specifically impairs ER-phagy mediated by receptors on ER sheets, such as TEX264, but not by those on ER tubules. Thus, it will be very important to understand how UFMylation on the ER surface is recognised prior to ER-phagy and how this impacts TEX264's function.

TEX264 in DNA repair

We identified TEX264 in a mass spectrometry analysis of proteins that interact with p97 inside the nucleus [6]. p97 is an ATPase which mediates protein unfolding, typically to present them to the proteasome for degradation [16]. An intriguing aspect of TEX264 was that it possessed a putative p97 interaction motif, known as a SHP box, in its loosely structured C-terminus, which we found mediates its direct interaction with p97 *in vitro*.

As discussed below, the GyrI-like domain of TEX264 suggested it may play a role in regulating topoisomerases, possibly in collaboration with p97. The yeast homolog of p97, Cdc48, was previously implicated in repairing a DNA lesion composed of topoisomerase 1 covalently bound to the 3' end of a single-stranded DNA break, known as a TOP1cc [17]. TOP1ccs impede DNA replication and transcription and defects in their repair contribute to various neurological disorders [18–21]. Abrogating p97 activity in human cells significantly impaired TOP1cc repair [6]. As p97 requires cofactors to be recruited to its substrates, we speculated that TEX264 might fulfil the role of targeting p97 to TOP1ccs. Accordingly, we found that TEX264 is needed to bridge p97 and TOP1 both *in vitro* and *in vivo* (Figure 1) [6].

TEX264 deficient cells accumulate endogenous TOP1ccs, exhibit basal replication stress and DNA damage, and are sensitive to low doses of TOP1cc-stabilising drugs. TEX264's ability to promote TOP1cc repair relies on motifs in/neighbouring its GyrI-like domain as well as a SUMO-interacting motif contained within this domain [6]. As recombinant TEX264 and unmodified TOP1 directly interact *in vitro*, it seems most plausible that SUMO represents an additional signal to enable TEX264 to distinguish transient TOP1ccs from trapped TOP1ccs, which are extensively modified with SUMO [22]. Indeed, in yeast, Cdc48 promotes the repair of SUMOylated TOP1ccs via its SUMO-binding cofactors, Ufd1 and the metalloprotease Wss1 [17,23,24]. In addition, TOP1cc SUMOylation may enhance the binding affinity between TEX264 and TOP1.

In metazoans, the metalloprotease, and p97 cofactor, SPRTN also proteolytically digests TOP1ccs, as well as other DNA-protein crosslinks (DPCs) [25–27]. TEX264 is necessary to bridge the interaction between TOP1 and SPRTN but is dispensable for general DPC repair [6]. Overall, we propose that p97 unfolds the TOP1 protein such that it can be proteolytically digested by SPRTN. The resulting DNA-bound peptide remnant can only then be excised by the phosphodiesterase TDP1, thus completing TOP1cc repair [20,28].

SPRTN bares motifs which enable it to interact with the DNA replication clamp loader, PCNA, via a PIP box and ubiquitinated proteins, via its UBZ, and its role in DPC repair is coupled to DNA replication [26,29–32]. This begs the question of why an additional cofactor is needed for its recruitment to substrates. There is some evidence that SPRTN's PIP box and UBZ are not required for its recruitment to chromatin upon DPC formation and its role in DPC repair [25,33]. This indicates that there must be other modes of recruiting SPRTN to specific DPC substrates. Indeed, the requirement for an additional recruitment factor, such as TEX264, is particularly important in the case of TOP1ccs, which are linked to the 3' end of single-stranded DNA breaks, and therefore would not be directly encountered by the elongating polymerase.

Moreover, owing to SPRTN's small active site, which can only be accessed by flexible peptide substrates, there must be additional factors, such as p97, that enable the processing of bulky DPCs [34]. The involvement of additional substrate-recognition factors for specific DPCs could also be a mechanism of restraining SPRTN's potentially deleterious protein sequence-unspecific cleavage activity by uncoupling DPC recognition from DPC proteolysis.

TEX264 is localised predominantly at the ER and the nuclear periphery, where it is tethered by its N-terminal transmembrane leucine-rich repeat (LRR) [5]. A variant of TEX264 that lacks this LRR redistributes into the cytosol as well as the inner nuclear space [6]. This is consistent with a sub-population of TEX264 being localised to the INM, facing inwards, as well as the ER membrane, facing the cytosol. Interestingly, we detected TEX264 at replication forks by isolation of proteins on nascent DNA (iPOND) and by immunofluorescent co-localisation with nascent DNA [6]. As there is no experimental evidence of alternatively spliced TEX264 isoforms that lack the LRR nor for a cleavage-mediated mechanism for releasing TEX264 from the INM, TEX264 is likely to be acting at replication forks in the vicinity of the nuclear envelope. This is interesting in the context of recent work which demonstrated that TOP1 acts on R-loops at nuclear lamina-associated heterochromatic regions [35]. Indeed, these chromatin regions are highly prone to topological stress and could suggest that TOP1ccs frequently arise in the vicinity of the INM. Further supporting this possibility, it was found that, upon TOP1cc-induced fork stalling, the SLFN11 protein is recruited by RPA1 to DNA replication sites at the nuclear periphery [36].

The SUMO modification machinery is also active at the INM. For example, modification of lamin A/C by SUMO1 in response to DNA damage is proposed to stimulate its interaction with LC3B and promote its clearance by nucleophagy [37]. A role in relocalizing SUMOylated proteins at DNA lesions to the nuclear periphery has been widely described in yeast. For example, in the S/G_2 phases of the cell cycle, mono-SUMOylation of unidentified factors

triggers the relocalization of a persistent DNA double-strand break (DSB) to the INM [38]. Moreover, mono-SUMOylation of various repair proteins promotes the recruitment of collapsed DNA replication forks to the nuclear pore complex [39]. Similarly, in Drosophila, DSBs in heterochromatic DNA regions move to the INM in a SUMO-dependent manner [40]. Whether DNA lesions are relocalized to the nuclear periphery in human cells is less well explored, although the association between the nuclear lamina and various human DNA repair and replication factors is important for maintaining genome stability [41,42].

Because a SUMO-targeted ubiquitin ligase (STUbL) is required for these relocalized lesions to be repaired, it has been speculated that SUMOylated proteins at collapsed replication forks or resected DSBs need to be degraded by the proteasome to ensure appropriate repair [43]. Some of the targets of this STUbL activity are likely to also be Cdc48/p97 substrates, given the cooperative activity of Cdc48 and STUbL in maintaining genome stability, including in the repair of TOP1ccs [17]. Whether TEX264 also acts with p97 to present TOP1ccs to the proteasome is unknown, however, proteasomal proteolysis is thought to largely occur at the nuclear envelope (and rough ER), where p97 also has diverse roles [44,45]. This possibility is further supported by the findings that the human STUbL, RNF4, is required for proteasomal TOP1cc degradation, and is known to mark DNA repair factors for extraction by p97 and SPRTN [46,47].

Intersection of DNA repair and autophagy?

As a transmembrane protein, TEX264 acts as a receptor, either at the ER membrane facing the cytosol (for ER-phagy), or the INM facing the nucleus (for DNA repair). Whether there is any further overlap between its distinct reported roles remains unknown (Figure 1).

Numerous lines of evidence indicate that autophagy contributes to the maintenance of genome stability through the degradation of nuclear proteins, micronuclei, and cytosolic chromatin

fragments. In yeast, the DNA repair protein Sae2 (RBBP8/CtIP in humans) is degraded by autophagy when histone deacetylases are inhibited, resulting in impaired DNA end resection and increased cellular sensitivity to DNA damaging agents [48]. In human cells, the levels of the autophagy cargo receptor SQSTM1/p62 influence DNA repair. For example, nuclear SQSTM1 interacts with, and inhibits, the DNA repair E3 ligase RNF168, resulting in defective homology-dependent DNA repair [49]. Nuclear p62 also promotes the degradation of the DNA repair protein, RAD51, by the proteasome [50]. Thus, the autophagic degradation of nuclear p62 facilitates homologous recombination repair.

Besides these direct roles in DNA repair, autophagy also mediates the degradation of nuclear components in mammalian cells during DNA damage- or oncogene-induced senescence [51,52]. Nuclear autophagy (i.e., nucleophagy) was first described in yeast, where Atg39 mediates the autophagic degradation of the nuclear envelope and inner nuclear membrane proteins in response to starvation [53]. While no human homologue of Atg39 has been identified, recent work has shown that numerous autophagy proteins are present in the nuclei of mammalian cells, including LC3, ATG5, and ATG7 [52,54]. Indeed, nuclear proteins, such as LMNB1 (lamin B1) and SIRT1, undergo stress-induced degradation in a manner that requires their direct interaction with nuclear LC3B and is mediated by the canonical cytosolic autophagy machinery [51,52]. Importantly, mammalian nucleophagy appears to be distinct from yeast nucleophagy in that it is not induced by conventional stresses, such as starvation or MTOR inhibition [52]. Rather, mammalian nucleophagy is triggered during DNA damage- and oncogene-induced replicative senescence and cells that fail to induce nucleophagy escape senescence [52]. A detailed understanding of how nuclear proteins are targeted for autophagic degradation is lacking and, to date, no mammalian counterpart of Atg39 has been identified. It is plausible that such a receptor protein(s) exists to facilitate the shuttling of nuclear components to cytosolic autophagosomes by directly interacting with either nuclear LC3B or

the substrates themselves. Some of the known ER-phagy receptors could possibly also regulate nucleophagy given that the ER and inner nuclear membranes are contiguous. TEX264 potentially fulfils the criteria of a nucleophagy receptor since it localises to the perinuclear ER membrane and nuclear envelope and interacts with LC3 family members [5,6]. Accordingly, it will be interesting to determine if TEX264 and LC3 interact inside the nucleus. Additionally, the ER-phagy receptors, CCPG1, which localises to the perinuclear ER, and C53, which is also present within the nucleus, may be candidate nucleophagy receptors [9,55].

Recent work has shown that cytosolic DNA triggers autophagy which, in turn, drives the clearance of DNA from the cytosol. After replicative stress, damaged chromatin fragments bud from the nucleus into the cytosol and are targeted to the lysosome by p62 [56]. Similarly, micronuclei harbouring damaged chromatin are coated with p62 and subjected to autophagic degradation [57]. Cytosolic DNA species generated by telomeric DNA damage activate autophagy via the CGAS-STING1 pathway, triggering autophagic cell death, presumably through the degradation of vital cellular components [58]. Intriguingly, TOP1ccs on cytosolic chromatin fragments were recently proposed to be crucial for cGAS activation during senescence because they can be directly bound by cGAS, enhancing its binding to DNA [59].

Autophagy has previously been implicated in the SUMO- and Cdc48-dependent repair of TOP1ccs. In yeast, the DPC protease Wss1 forms a complex with Cdc48 and Doa1, another Cdc48 cofactor implicated in selective autophagy [24]. In response to replication stress, Wss1 relocalises to vacuoles, suggesting a link between DPC repair and autophagy. It is unclear if DPCs could be degraded by autophagy. Given the high endogenous cellular concentrations of formaldehyde and its propensity to induce protein-protein as well as DNA-protein crosslinks, it is possible that autophagy helps to evict aggregated protein-protein and DNA-protein crosslinks from the nucleus. This could possibly involve the recycling of the liberated protein fragments generated by SPRTN- or Wss1-mediated DPC proteolysis. Interestingly, tandem-

affinity mass spectrometry data revealed that the interaction between TEX264 and TOP1 is significantly increased during starvation, however the functional relevance of this is unclear [5].

Proximity biotinylation mass spectrometry analysis of TEX264 detected many p97-derived peptides, the number of which was not altered during starvation or by mutating TEX264's LIR, indicating that the association between TEX264 and p97 is unaffected by TEX264's ability to traffic into autophagosomes [5]. p97 is required for autophagic degradation, including ribophagy and mitophagy, because it promotes autophagosome–lysosome fusion, yet the precise mechanisms underlying its role are unclear [2,60]. In cooperation with TEX264, p97 might be required to enable engulfment of the ER membrane by phagophores. Another possibility is that direct p97-dependent extraction of modified ribosomes from the ER membrane is necessary to enable its engulfment. Arguing against this is the observation that VCP/p97 depletion by CRISPR interference enhances ER-phagy [14]. This could result from diminished p97-dependent ER-associated degradation (ERAD), leading to an increased reliance on ER-phagy to clear misfolded proteins from the ER, which could mask any negative impact p97 depletion has on ER-phagy. Nevertheless, p97 could facilitate ER-phagy by, for example, removing ER membrane proteins to expose other substrates for ubiquitination or UFMylation, via a mechanism that would be analogous to its proposed role in mitophagy and the Endo-Lysosomal Damage Response [61–63].

Evolutionary and structural analysis of GyrI proteins

A particularly noteworthy feature of TEX264 is its GyrI-like domain which makes it a member of an evolutionarily ancient superfamily of proteins with diverse functions that include inhibitors of the type II topoisomerase gyrase and transcriptional regulators [64–68]. TEX264's

GyrI-like domain is required for TEX264 to bind TOP1, but its relevance, if any, for ER-phagy is unknown.

Most GyrI proteins have acquired domains that confer a diverse array of additional functions, some of which are illustrated in Figure 2A. For example, the GyrI domain of the transcription factors, Rob and BmrR, is fused to an N-terminal helix-turn-helix (HTH) motif that binds DNA. The prototypical member of the GyrI superfamily is *Escherichia coli* SbmC (renamed GyrI), which was shown to protect cells from microcin b17, a peptide which traps covalent DNA-gyrase intermediates [69]. Interestingly, the expression of SbmC/GyrI is induced in response to both DNA damage and nutrient starvation, potentially providing a distant evolutionary basis for the roles of human TEX264 in autophagy and DNA repair. Subsequent work found that GyrI co-purified with gyrase and suppressed its supercoiling activity, most likely by either sequestering gyrase or inhibiting its binding to DNA [65,67,70]. GyrI also counteracts the cytotoxic effects of quinolones, a non-proteinaceous class of antibiotics that stabilise DNA-gyrase complexes, and other DNA-damaging agents, such as mitomycin C [71]. It will be interesting to know if TEX264 adopts a similar mechanism of action to GyrI to suppress TOP1ccs, specifically, by addressing whether the direct binding of TEX264 to topoisomerases inhibits their decatenation activity on DNA templates *in vitro*. They appear to be distinct mechanisms as TEX264 recruits TOP1cc repair factors, requires p97 activity, and is epistatic with TDP1 in the repair of endogenous TOP1ccs. Moreover, a GyrI-derived 8 amino acid-long peptide that inhibits gyrase does not inhibit TOP1 activity [67]. Interestingly, a subgroup of prokaryotic GyrI-like proteins (but not GyrI itself) was recently shown to possess hydrolase activity [72]. This activity catalyses the hydrolysis of DNA-alkylating agents and thereby confers cellular resistance to cytotoxic xenobiotics. The catalytic activity of these proteins depends on pairs of aromatic and acidic residues, however, TEX264 does not contain corresponding residues required for catalysis.

Intrigued by the fact that homologs of TEX264 are present in vertebrates but absent in established model organisms such as yeast (*Saccharomyces cerevisiae and Schizosaccharomyces pombe),* we decided to investigate the evolutionary history of GyrI domain-containing proteins. Blastp searches using bacterial and human GyrI domain sequences through bacteria, yeast, plants, fungi, invertebrate and chordate species was followed by multiple sequence alignment using the MAFFT alignment algorithm, while alignment quality was assessed using Guidance software (Figure 3) [73,74]. Phylogenetic trees were constructed with maximum likelihood analysis in PhyML (Figure 4) [75,76]. Ftsa (cell division ATPase) proteins were used as an outgroup (Figure S1) because they are functionally different from GyrI domain-containing proteins, yet they contain an SHS2 module like GyrI proteins, making it possible to reach sufficiently good multiple protein alignment for subsequent tree building. The GyrI domain of SbmC/GyrI contains two tandem SHS modules, the second of which encompasses its interaction site with Gyrase, suggesting this domain mediates protein-protein interactions [77]. The SHS2 module of TEX264 is highly conserved across TEX264 orthologs, as well as in *E. coli* SbmC/GyrI, indicating its functional and/or structural importance prior to the evolution of autophagy (Figure S1).

We have found conserved TEX264 orthologs in invertebrate and chordate species (Figure 4 & S1), while they were absent in fungi and plants. Despite being present in bacteria and metazoans, GyrI domain-containing proteins are notably absent in yeast, similar to poly(ADP-ribose) polymerases (PARPs) [78]. Since yeast diverged prior to the evolution of metazoans, this indicates that they either independently lost GyrI domain-containing proteins or that these domains were regained via convergent evolution in the first common ancestor of all metazoans.

Phylogenetic analysis of full-length proteins (Figure 4 & S1) and of the GyrI domain (amino acids 41-185 of human TEX264) showed similar clustering (Figure S2). Bacterial GyrI-domain containing proteins expectedly cluster closer to the TEX264 group than to the Ftsa outgroup

but are quite distant to the TEX264 group and bare substantial differences in protein sequence outside of the GyrI domain. On the other hand, the GyrI domain of bacterial proteins is similar to TEX264 orthologs with several conserved regions (Figure 3), and most importantly structural models of human TEX264 GyrI domain can be constructed with high model confidence (Figure 2B & 3). Given the phylogenetic proximity and similarity in 3D structures of the GyrI domain, bacterial GyrI domain-containing proteins might be regarded as distant ancestors of the GyrI domain in TEX264 orthologs.

Unlike its GyrI-like domain, the C terminus of TEX264 - corresponding to amino acids 186-313 of the full-length protein sequence – substantially diverges in invertebrates and is absent in bacterial GyrI domain-containing proteins (Figure S3). In vertebrates, TEX264's C terminus is conserved, including both the LIR and SHP motifs (Figure S3). However, the C-terminal part diverges considerably in the invertebrate lineage (Figure S3). Specifically, the LIR motif is only partly conserved in tunicates, molluscs, and crustaceans, while it is absent in nematodes, sponges, and a chordate lancelet (*Branchiostoma floridae*; Figure S3). The SHP motif is conserved in higher vertebrate species, from reptiles to mammals, but is divergent in lower vertebrates (fish and amphibians), and completely absent in invertebrate TEX264 orthologs (Figure S3). Bacterial GyrI domain-containing proteins are shorter than TEX264 orthologs and lack a C terminus which would resemble that of TEX264. Likewise, bacterial GyrI domain-containing proteins lack the N-terminal LRR domain which is otherwise highly conserved throughout the animal kingdom (Figure S4).

The model of TEX264's GyrI-like domain shows two antiparallel sheets and two α-helices following the β1-α1-β2-β3-β4-α2-β5-β6 linear arrangement (Figure 2B & 4). Similar to bacterial GyrI proteins, *E. coli* SbmC and Rob2, two similar halves of GyrI domain show pseudo two-fold symmetry (Figure 2B). The N-terminus of TEX264 bares a leucine-rich repeat (LRR) structural motif that forms an α/β horseshoe fold. The LRR motif was modelled with

good-to-high confidence based on the photosystem II reaction centre protein J (6J3Y; Figure 2B). The C-terminal part containing the LC3-interacting region (LIR) and p97 interacting motif (SHP) were modelled with lower confidence based on a *Thermotoga maritima* mannanase (Man5) carbohydrate binding module (CBM) (1of3) and shows at least two β-sheets with good model confidence (Figure 2B).

Future perspectives

Three recent studies have reported distinct roles for the TEX264 protein as a membrane-anchored receptor either for ER-phagy or for nuclear substrates during DNA repair. TEX264 evolved from an ancient superfamily of proteins, orthologs of which are present in bacteria and metazoans. GyrI domain-containing proteins have acquired additional and diverse domains and functions throughout evolution, including transcription regulation, chromatin-remodelling, and protein homeostasis. The fact that GyrI proteins and TEX264 orthologs pre-date the evolution of autophagy hints at a distinct primordial function of these proteins, as illustrated by the role of the TEX264 relative, SbmC, in regulating bacterial gyrase. It will be interesting to understand whether the GyrI-like domain of TEX264 and its interaction with p97 are also important for ER-phagy. It will be also be fascinating to explore the contribution of TEX264 to processes such as nuclear degradation in cell types that undergo extensive organelle loss during differentiation, such as erythroblasts and epidermal keratinocytes [79]. Future studies should also aim to address whether TEX264's role in the nucleus extends beyond TOP1cc repair and whether these roles rely on its ability to promote autophagy and associate with the INM.

Acknowledgments

The K.R. laboratory is funded by Medical Council Research Programme grant (MC_EX_MR/K022830/1). M.P. research group is supported by Croatian Science Foundation Installation Grant (UIP-2017-05-5258) and Project grant (IPS-2020-01-4225), Ruder Boskovic Institute (Zagreb, Croatia) and European Structural and Investment Funds STIM – REI project (KK.01.1.1.01.0003).

Declaration of interest

The authors declare no conflict of interest.

References

1. Khaminets, A., Behl, C. & Dikic, I. Ubiquitin-Dependent And Independent Signals In Selective Autophagy. *Trends in Cell Biology* **26**, 6–16 (2016).

2. Dargemont, C. & Ossareh-Nazari, B. Cdc48/p97, a key actor in the interplay between autophagy and ubiquitin/proteasome catabolic pathways. *Biochimica et Biophysica Acta - Molecular Cell Research* **1823**, 138–144 (2012).

3. Dikic, I. Open questions: Why should we care about ER-phagy and ER remodelling? *BMC Biology* **16**, 131 (2018).

4. Chino, H., Hatta, T., Natsume, T. & Mizushima, N. Intrinsically Disordered Protein TEX264 Mediates ER-phagy. *Molecular Cell* **74**, 909-921.e6 (2019).

5. An, H. *et al.* TEX264 Is an Endoplasmic Reticulum-Resident ATG8-Interacting Protein Critical for ER Remodeling during Nutrient Stress. *Molecular Cell* **74**, 891-908.e10 (2019).

6. Fielden, J. *et al.* TEX264 coordinates p97- and SPRTN-mediated resolution of topoisomerase 1-DNA adducts. *Nature Communications* **11**, 1274 (2020).

7. Khaminets, A. *et al.* Regulation of endoplasmic reticulum turnover by selective autophagy. *Nature* **522**, 354–358 (2015).

8. Stephani, M. et al. A cross-kingdom conserved ER-phagy receptor maintains endoplasmic reticulum homeostasis during stress. *eLife* **9**, e58396 (2020).

9. Smith, M. D. *et al.* CCPG1 Is a Non-canonical Autophagy Cargo Receptor Essential for ER-Phagy and Pancreatic ER Proteostasis. *Developmental Cell* **44**, 217-232.e11 (2018).

10. Chen, Q. *et al.* ATL3 Is a Tubular ER-Phagy Receptor for GABARAP-Mediated Selective Autophagy. *Current Biology* **29**, 846-855.e6 (2019).

11. Liang, J. R., Lingeman, E., Ahmed, S. & Corn, J. E. Atlastins remodel the endoplasmic reticulum for selective autophagy. *Journal of Cell Biology* **217**, 3354–3367 (2018).

12. Delorme-Axford, E., Popelka, H. & Klionsky, D. J. TEX264 is a major receptor for mammalian reticulophagy. *Autophagy* vol. 15 1677–1681 (2019).

13. Fumagalli, F. *et al.* Translocon component Sec62 acts in endoplasmic reticulum turnover during stress recovery. *Nature Cell Biology* **18**, 1173–1184 (2016).

14. Liang, J. R. *et al.* A Genome-wide ER-phagy Screen Highlights Key Roles of Mitochondrial Metabolism and ER-Resident UFMylation. *Cell* **180**, 1160-1177.e20 (2020).

15. Wilkinson, S. Picky Eating at the ER-phagy Buffet. *Trends in Biochemical Sciences* **44**, 731–733 (2019).

16. Bodnar, N. O. & Rapoport, T. A. Molecular Mechanism of Substrate Processing by the Cdc48 ATPase Complex. *Cell* **169**, 722-735.e9 (2017).

17. Nie, M. *et al.* Dual recruitment of Cdc48 (p97)-Ufd1-Npl4 ubiquitin-selective segregase by small ubiquitin-like modifier protein (SUMO) and ubiquitin in SUMO-

targeted ubiquitin ligase-mediated genome stability functions. *Journal of Biological Chemistry* **287**, 29610–29619 (2012).

18. Walker, C. *et al.* C9orf72 expansion disrupts ATM-mediated chromosomal break repair. *Nature Neuroscience* **20**, 1225–1235 (2017).

19. Alagoz, M., Chiang, S. C., Sharma, A. & El-Khamisy, S. F. ATM Deficiency Results in Accumulation of DNA-Topoisomerase I Covalent Intermediates in Neural Cells. *PLoS ONE* **8**, e58239 (2013).

20. El-Khamisy, S. F. *et al.* Defective DNA single-strand break repair in spinocerebellar ataxia with axonal neuropathy-1. *Nature* **434**, 108–113 (2005).

21. Katyal, S. *et al.* Aberrant topoisomerase-1 DNA lesions are pathogenic in neurodegenerative genome instability syndromes. *Nature Neuroscience* **17**, 813–821 (2014).

22. Mao, Y., Desai, S. D. & Liu, L. F. SUMO-1 conjugation to human DNA topoisomerase II isozymes. *Journal of Biological Chemistry* **275**, 26066–26073 (2000).

23. Stingele, J., Schwarz, M. S., Bloemeke, N., Wolf, P. G. & Jentsch, S. A DNA-dependent protease involved in DNA-protein crosslink repair. *Cell* **158**, 327–338 (2014).

24. Balakirev, M. Y. *et al.* Wss1 metalloprotease partners with Cdc48/Doa1 in processing genotoxic SUMO conjugates. *eLife* **4**, e06763 (2015).

25. Maskey, R. S. *et al.* Spartan deficiency causes accumulation of Topoisomerase 1 cleavage complexes and tumorigenesis. *Nucleic Acids Research* **45**, 4564–4576 (2017).

26. Vaz, B. *et al.* Metalloprotease SPRTN/DVC1 Orchestrates Replication-Coupled DNA-Protein Crosslink Repair. *Molecular Cell* **64**, 704–719 (2016).

27. Fielden, J., Ruggiano, A., Popović, M. & Ramadan, K. DNA protein crosslink proteolysis repair: From yeast to premature ageing and cancer in humans. *DNA Repair* **71**, 198–204 (2018).

28. Interthal, H. & Champoux, J. J. Effects of DNA and protein size on substrate cleavage by human tyrosyl-DNA phosphodiesterase 1. *Biochemical Journal* **436**, 559–566 (2011).

29. Mosbech, A. *et al.* DVC1 (C1orf124) is a DNA damage-targeting p97 adaptor that promotes ubiquitin-dependent responses to replication blocks. *Nature Structural and Molecular Biology* **19**, 1084–1092 (2012).

30. Larsen, N. B. *et al.* Replication-Coupled DNA-Protein Crosslink Repair by SPRTN and the Proteasome in Xenopus Egg Extracts. *Molecular Cell* **73**, 574-588.e7 (2019).

31. Halder, S. *et al.* SPRTN protease and checkpoint kinase 1 cross-activation loop safeguards DNA replication. *Nature Communications* **10**, 3142 (2019).

32. Mórocz, M. *et al.* DNA-dependent protease activity of human Spartan facilitates replication of DNA-protein crosslink-containing DNA. *Nucleic Acids Research* **45**, 3172–3188 (2017).

33. Stingele, J. *et al.* Mechanism and Regulation of DNA-Protein Crosslink Repair by the DNA-Dependent Metalloprotease SPRTN. *Molecular Cell* **64**, 688–703 (2016).

34. Li, F., Raczynska, J. E., Chen, Z. & Yu, H. Structural Insight into DNA-Dependent Activation of Human Metalloprotease Spartan. *Cell Reports* **26**, 3336-3346.e4 (2019).

35. Manzo, S. G. *et al.* DNA Topoisomerase I differentially modulates R-loops across the human genome. *Genome Biology* **19**, 100 (2018).

36. Murai, J. *et al.* SLFN11 Blocks Stressed Replication Forks Independently of ATR. *Molecular Cell* **69**, 371-384.e6 (2018).

37. Li, Y. *et al.* Nuclear accumulation of UBC9 contributes to SUMOylation of lamin A/C and nucleophagy in response to DNA damage. *Journal of Experimental and Clinical Cancer Research* **38**, 67 (2019).

38. Horigome, C. *et al.* PolySUMOylation by Siz2 and Mms21 triggers relocation of DNA breaks to nuclear pores through the Slx5/Slx8 STUbL. *Genes and Development* **30**, 931–945 (2016).

39. Whalen, J. M., Dhingra, N., Wei, L., Zhao, X. & Freudenreich, C. H. Relocation of Collapsed Forks to the Nuclear Pore Complex Depends on Sumoylation of DNA Repair Proteins and Permits Rad51 Association. *Cell Reports* **31**, 107635 (2020).

40. Ryu, T., Bonner, M. R. & Chiolo, I. Cervantes and Quijote protect heterochromatin from aberrant recombination and lead the way to the nuclear periphery. *Nucleus* **7**, 485–497 (2016).

41. Li, W. *et al.* The nucleoskeleton protein IFFO1 immobilizes broken DNA and suppresses chromosome translocation during tumorigenesis. *Nature Cell Biology* **21**, 1273–1285 (2019).

42. Cobb, A. M., Murray, T. v., Warren, D. T., Liu, Y. & Shanahan, C. M. Disruption of PCNA-lamins A/C interactions by prelamin A induces DNA replication fork stalling. *Nucleus* **7**, 498–511 (2016).

43. Horigome, C. & Gasser, S. M. SUMO wrestles breaks to the nuclear ring's edge. *Cell Cycle* **15**, 3011–3013 (2016).

44. Olmos, Y., Hodgson, L., Mantell, J., Verkade, P. & Carlton, J. G. ESCRT-III controls nuclear envelope reformation. *Nature* **522**, 236–239 (2015).

45. Enenkel, C., Lehmann, A. & Kloetzel, P. M. GFP-labelling of 26S proteasomes in living yeast: Insight into proteasomal functions at the nuclear envelope/rough ER. *Molecular Biology Reports* **26**, 131–135 (1999).

46. Sun, Y. *et al.* A conserved SUMO-Ubiquitin pathway directed by RNF4/SLX5-SLX8 and PIAS4/SIZ1 drives proteasomal degradation of topoisomerase DNA-protein crosslinks. *BioRxiv* **707661**, 10.1101/707661 (2019).

47. Gibbs-Seymour, I. *et al.* Ubiquitin-SUMO circuitry controls activated fanconi anemia ID complex dosage in response to DNA damage. *Molecular Cell* **57**, 150–164 (2015).

48. Robert, T. *et al.* HDACs link the DNA damage response, processing of double-strand breaks and autophagy. *Nature* **471**, 74–79 (2011).

49. Wang, Y. *et al.* Autophagy Regulates Chromatin Ubiquitination in DNA Damage Response through Elimination of SQSTM1/p62. *Molecular Cell* **63**, 34–48 (2016).

50. Hewitt, G. *et al.* SQSTM1/p62 mediates crosstalk between autophagy and the UPS in DNA repair. *Autophagy* **12**, 1917–1930 (2016).

51. Xu, C. *et al.* SIRT1 is downregulated by autophagy in senescence and ageing. *Nature Cell Biology* **22**, 1170–1179 (2020).

52. Dou, Z. *et al.* Autophagy mediates degradation of nuclear lamina. *Nature* **527**, 105–109 (2015).

53. Mochida, K. *et al.* Receptor-mediated selective autophagy degrades the endoplasmic reticulum and the nucleus. *Nature* **522**, 359–362 (2015).

54. Lee, I. H. *et al.* Atg7 modulates p53 activity to regulate cell cycle and survival during metabolic stress. *Science* **336**, 225–228 (2012).

55. Wang, J., He, X., Luo, Y. & Yarbrough, W. G. A novel ARF-binding protein (LZAP) alters ARF regulation of HDM2. *Biochemical Journal* **393**, 489–501 (2006).

56. Ivanov, A. *et al.* Lysosome-mediated processing of chromatin in senescence. *Journal of Cell Biology* **202**, 129–143 (2013).

57. Rello-Varona, S. *et al.* Autophagic removal of micronuclei. *Cell Cycle* **11**, 170–176 (2012).

58. Nassour, J. *et al.* Autophagic cell death restricts chromosomal instability during replicative crisis. *Nature* **565**, 659–663 (2019).

59. Zhao, B. *et al.* Topoisomerase 1 cleavage complex enables pattern recognition and inflammation during senescence. *Nature Communications* **11**, 908 (2020).

60. Ju, J. S. *et al.* Valosin-containing protein (VCP) is required for autophagy and is disrupted in VCP disease. *Journal of Cell Biology* **187**, 875–888 (2009).

61. McLelland, G. L. *et al.* Mfn2 ubiquitination by PINK1/parkin gates the p97-dependent release of ER from mitochondria to drive mitophagy. *eLife* **7**, e32866 (2018).

62. Tanaka, A. *et al.* Proteasome and p97 mediate mitophagy and degradation of mitofusins induced by Parkin. *Journal of Cell Biology* **191**, 1367–1380 (2010).

63. Papadopoulos, C. *et al.* VCP /p97 cooperates with YOD 1, UBXD 1 and PLAA to drive clearance of ruptured lysosomes by autophagy. *The EMBO Journal* **36**, 135–150 (2017).

64. Zheleznova Heldwein, E. E. & Brennan, R. G. Crystal structure of the transcription activator BmrR bound to DNA and a drug. *Nature* **409**, 378–382 (2001).

65. Nakanishi, A., Oshida, T., Matsushita, T., Imajoh-Ohmi, S. & Ohnuki, T. Identification of DNA gyrase inhibitor (GyrI) in Escherichia coli. *Journal of Biological Chemistry* **273**, 1933–1938 (1998).

66. Romanowski, M. J., Gibney, S. A. & Burley, S. K. Crystal structure of the Escherichia coli SbmC protein that protects cells from the DNA replication inhibitor microcin B17. *Proteins: Structure, Function and Genetics* **47**, 403–407 (2002).

67. Nakanishi, A., Imajoh-Ohmi, S. & Hanaoka, F. Characterization of the interaction between DNA gyrase inhibitor and DNA gyrase of Escherichia coli. *Journal of Biological Chemistry* **277**, 8949–8954 (2002).

68. Kwon, H. J., Bennik, M. H. J., Demple, B. & Ellenberger, T. Crystal structure of the Escherichia coli Rob transcription factor in complex with DNA. *Nature Structural Biology* **7**, 424–430 (2000).

69. Baquero, M. R., Bouzon, M., Varea, J. & Moreno, F. sbmC, a stationary-phase induced SOS Escherichia coli gene, whose product protects cells from the DNA replication inhibitor microcin B17. *Molecular Microbiology* **18**, 301–311 (1995).

70. Chatterji, M. & Nagaraja, V. GyrI: A counter-defensive strategy against proteinaceous inhibitors of DNA gyrase. *EMBO Reports* **3**, 261–267 (2002).

71. Chatterji, M., Sengupta, S. & Nagaraja, V. Chromosomally encoded gyrase inhibitor GyrI protects Escherichia coli against DNA-damaging agents. *Archives of Microbiology* **180**, 339–346 (2003).

72. Yuan, H. *et al.* GyrI-like proteins catalyze cyclopropanoid hydrolysis to confer cellular protection. *Nature Communications* **8**, 1485 (2017).

73. Katoh, K., Misawa, K., Kuma, K. I. & Miyata, T. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* **30**, 3059–3066 (2002).

74. Penn, O. *et al.* GUIDANCE: A web server for assessing alignment confidence scores. *Nucleic Acids Research* **38**, W23–W28 (2010).

75. Guindon, S. & Gascuel, O. A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. *Systematic Biology* **52**, 696–704 (2003).

76. Anisimova, M. & Gascuel, O. Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. *Systematic Biology* **55**, 539–552 (2006).

77. Anantharaman, V. & Aravind, L. The SHS2 module is a common structural theme in functionally diverse protein groups, like Rpb7p, FtsA, GyrI, and MTH1598/Tm1083 superfamilies. *Proteins: Structure, Function and Genetics* **56**, 795–807 (2004).

78. Citarelli, M., Teotia, S. & Lamb, R. S. Evolutionary history of the poly(ADP-ribose) polymerase gene family in eukaryotes. *BMC Evolutionary Biology* **10**, 308 (2010).

79. Rogerson, C., Bergamaschi, D. & O'Shaughnessy, R. F. L. Uncovering mechanisms of nuclear degradation in keratinocytes: A paradigm for nuclear degradation in other tissues. *Nucleus* **9**, 56–64 (2018).

Figure Legends

Figure 1. A model for TEX264 function in the ER and nucleus. TEX264 is anchored at both the ER and INM via its N-terminal single-pass transmembrane domain. TEX264 promotes degradation of portions of the ER during starvation by binding LC3-coated phagophores via its C-terminal LIR. At the INM, TEX264 associates with VCP/p97-SPRTN subcomplexes via its C-terminal SHP box and promotes TOP1cc repair.

Figure 2. Diversity, topology, and structural models of GyrI proteins. (A) Representative schematics of a subset of GyrI superfamily members, from a total of 73 distinct domain organisations. (B) Topology and structural models of human TEX264 protein motifs and domains. All models were created using the SWISS-MODEL workspace and/or Phyre2 server. GyrI domain was modeled with high to very high confidence based on three templates: SbmC, *E. coli* Rob transcription factor2 (1d5y), and an uncharacterized protein from *Chlorobium tepidum* (2kcu). The N-terminus of TEX264 bares a leucine-rich repeat (LRR) structural motif that forms an α/β horseshoe fold. The LRR motif was modelled with good to high confidence based on the photosystem II reaction centre protein J (6J3Y). The C-terminal part containing the LC3-interacting region (LIR) and p97 interacting motif (SHP) were modelled with lower confidence based on a *Thermotoga maritima* mannanase (Man5) carbohydrate binding module (CBM) (1of3) and show at least 2 β-sheets with good model confidence.

Figure 3. GyrI domain sequence alignments of TEX264 orthologs. The GyrI-like domain of human TEX264 corresponds to amino acids 41-185 of the full-length protein. Shown above the alignment (grey line), the SHS2 fold in human TEX264 corresponds to amino acids 21-127. The structure of human TEX264 according to 3D modelling is labelled for the corresponding protein sequence, where α-helices are shown in green and β-sheets in blue, as in the structural model of human TEX264 in Figure 2B. Red lines designate conserved motifs and

domains. TEX264 orthologs are shown in orange (dark: vertebrates; bright: invertebrates), while bacterial GyrI-domain containing proteins are shown in blue. Protein sequences were aligned using the MAFFT (Multiple Alignment using Fast Fourier Transform) alignment algorithm. Alignment quality score was assessed using the Guidance2 server and was 0.752, where 1 is maximum, indicating high alignment quality.

Figure 4. Phylogenetic analysis of TEX264 proteins. TEX264 orthologs in vertebrates are highlighted (vertebrates in orange; invertebrates are colourless) and bacterial GyrI-domain containing proteins are highlighted blue. Ftsa proteins were used as an outgroup (Figure S2). Full length protein sequences were aligned with the MAFFT (Multiple Alignment using Fast Fourier Transform) alignment algorithm. The phylogentic tree was constructed using the maximum likelihood method. The expanded phylogenetic tree with detailed methodology is shown in Figure S2.

Supplementary Figure Legend

Figure S1. Phylogenetic analysis of TEX264 proteins. TEX264 orthologs are shown in orange (dark: vertebrates; bright: invertebrates). Bacterial GyrI domain-containing proteins are shown in blue. Ftsa proteins are used as an outgroup (indicated in red) because they are functionally different from GyrI domain-containing proteins, yet they contain the SHS2 module like GyrI proteins, making it possible to reach sufficiently good multiple protein alignment for tree building. Full-length protein sequences were aligned using MAFFT (Multiple Alignment using Fast Fourier Transform) alignment algorithm. Alignment quality score was assessed using Guidance2 server and was 0.563, where 1 is maximum, indicating sufficiently high alignment quality for tree building. The phylogenetic tree was constructed using the maximum likelihood method (PhyML software with optimised tree topology, LG model, 8 rates of categories, tree searching operation best of Nearest Neighbor Interchange & Subtree Pruning and Regrafting (NNI&SPR)). Branch support Alrt values (Approximate likelihood-ratio test) are shown at tree nodes on a scale of 0-1, where 1 is maximum node confidence.

Figure S2. Phylogenetic analysis of the GyrI domain in TEX264 orthologs. The GyrI domain of human TEX264 corresponds to amino acids 41-185. Vertebrate orthologs are shown in dark orange, invertebrate orthologs in bright orange, and bacterial GyrI domains are shown in blue. Ftsa proteins are used as an outgroup (indicated in red) because they are functionally different from GyrI domain-containing proteins, yet they contain the SHS2 module like GyrI proteins, making it possible to reach sufficiently good multiple protein alignment for subsequent tree building. GyrI domain protein sequences were aligned using MAFFT (Multiple Alignment using Fast Fourier Transform) alignment algorithm. Alignment quality score was assessed using Guidance2 server and was 0.619, where 1 is maximum, indicating sufficiently high alignment quality for tree building. The phylogenetic tree was constructed using the maximum likelihood method (PhyML software with optimised tree topology, LG model, 8

rates of categories, tree searching operation best of Nearest Neighbor Interchange & Subtree Pruning and Regrafting (NNI&SPR)). Branch support Alrt values (Approximate likelihood-ratio test) are shown at tree nodes on a scale of 0-1, where 1 is maximum node confidence.

Figure S3. Amino acid sequence alignment of the C-terminal part of TEX264 orthologs. The C-terminal part of human TEX264 corresponds to amino acids 186-313 in the sequence of the full-length protein. The LIR motif, consisting of amino acids FEEL in vertebrates (in orange), corresponds to amino acids 273-276 in human TEX264. The SHP motif consisting of amino acids GEGPLG in mammals (in black) corresponds to amino acids 280-285 in human TEX264. Protein sequences were aligned using the MAFFT (Multiple Alignment using Fast Fourier Transform) alignment algorithm. Alignment quality score was assessed using the Guidance2 server and was 0.752, where 1 is maximum, indicating high alignment quality.

Figure S4. LRR domain alignment in TEX264 orthologs. The LRR domain in human TEX264 corresponds to amino acids 5-33 of the full-length protein. Conserved residues are highlighted. Protein sequences were aligned using the MAFFT (Multiple Alignment using Fast Fourier Transform) alignment algorithm. Alignment quality score was assessed using the Guidance2 server and was 0.752, where 1 is maximum, indicating high alignment quality.

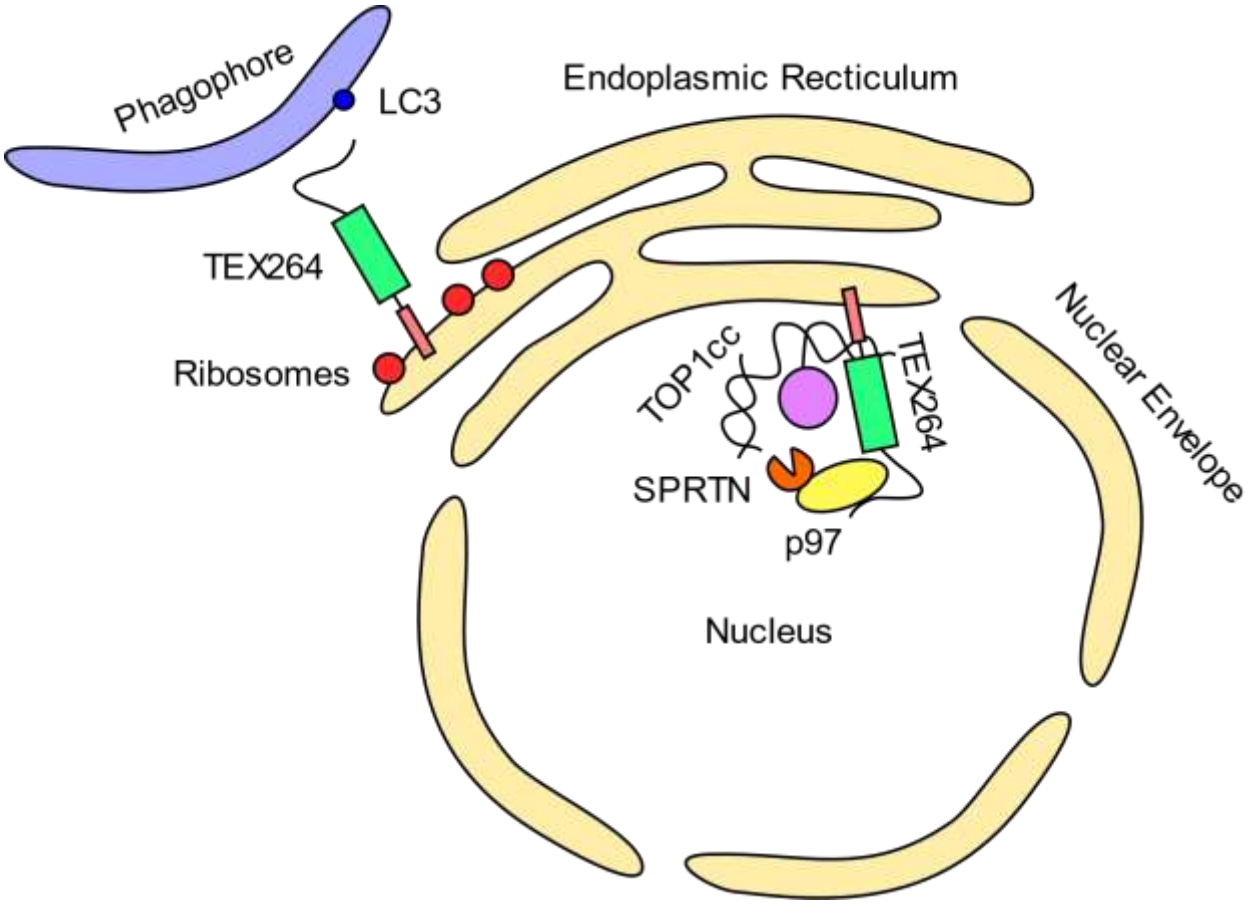

Figure 1. A model for TEX264 function in the ER and nucleus. TEX264 is anchored at both the ER and inner nuclear membranes via its N-terminal single-pass transmembrane domain. TEX264 promotes degradation of portions of the ER during starvation by binding LC3-coated phagophores via its C-terminal LC3-interacting region (LIR). At the INM, TEX264 associates with p97-SPRTN subcomplexes via its C-terminal SHP box and promotes TOP1cc repair.

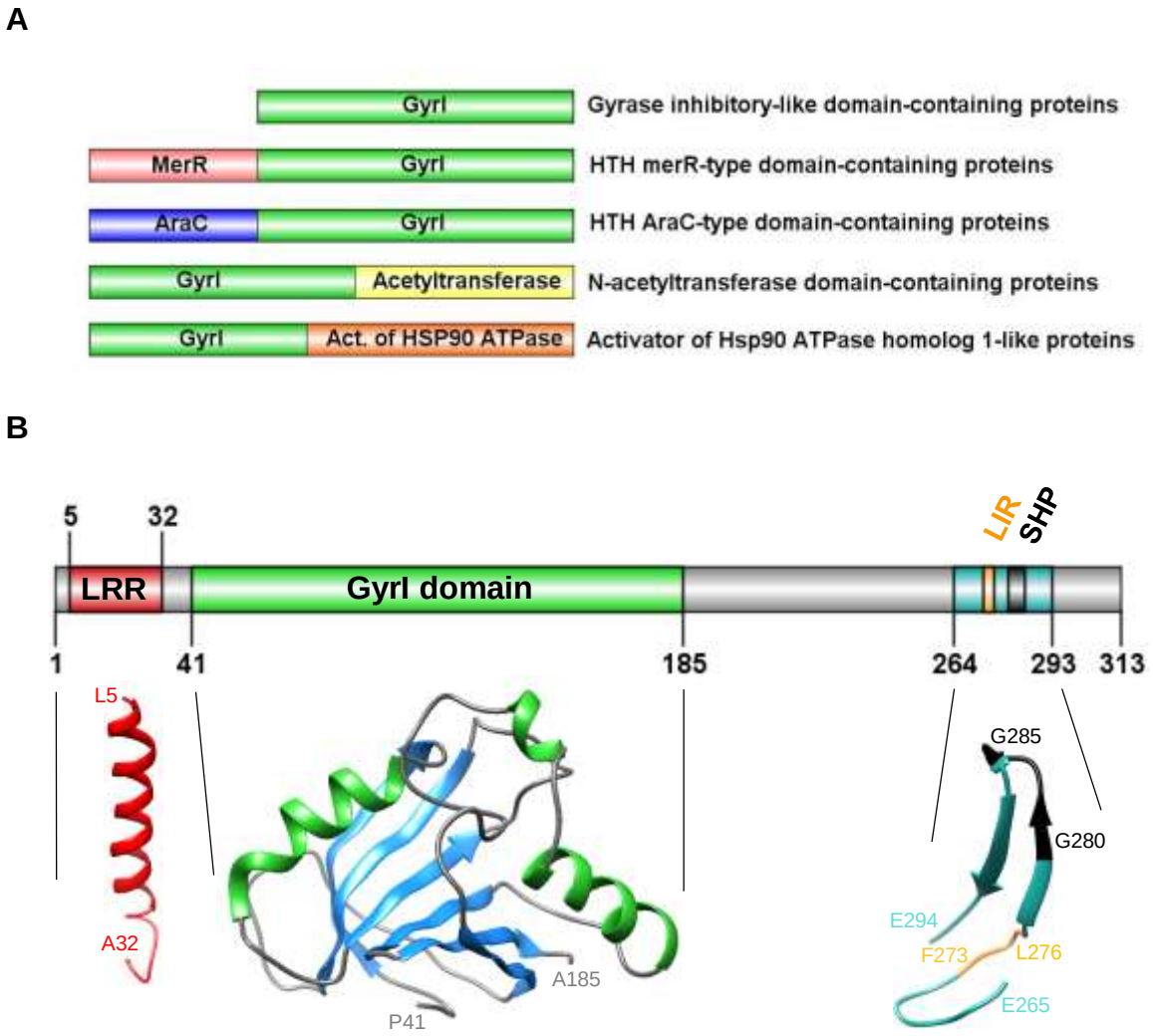

Figure 2. Diversity, topology, and structural models of GyrI proteins.
(A) Representative schematics of a subset of GyrI superfamily members, from a total of 73 distinct domain organisations. (B) Topology and structural models of human TEX264 protein motifs and domains. All models were created using the SWISS-MODEL workspace and/or Phyre2 server. GyrI domain was modeled with high to very high confidence based on three templates: SbmC, *E. coli* Rob transcription factor2 (1d5y) an uncharacterized protein from *Chlorobium tepidum* (2kcu). The N-terminus of TEX264 bares a leucine-rich repeat (LRR) structural motif that forms an α/β horseshoe fold. The LRR motif was modeled with good to high confidence based on the photosystem II reaction center protein J (6J3Y). The C-terminal part containing the LC3-interacting region (LIR) and p97 interacting motif (SHP) were modeled with lower confidence based on a *Thermotoga maritima* mannanase (Man5) carbohydrate binding module (CBM) (1of3) and show at least 2 β-sheets with good model confidence.

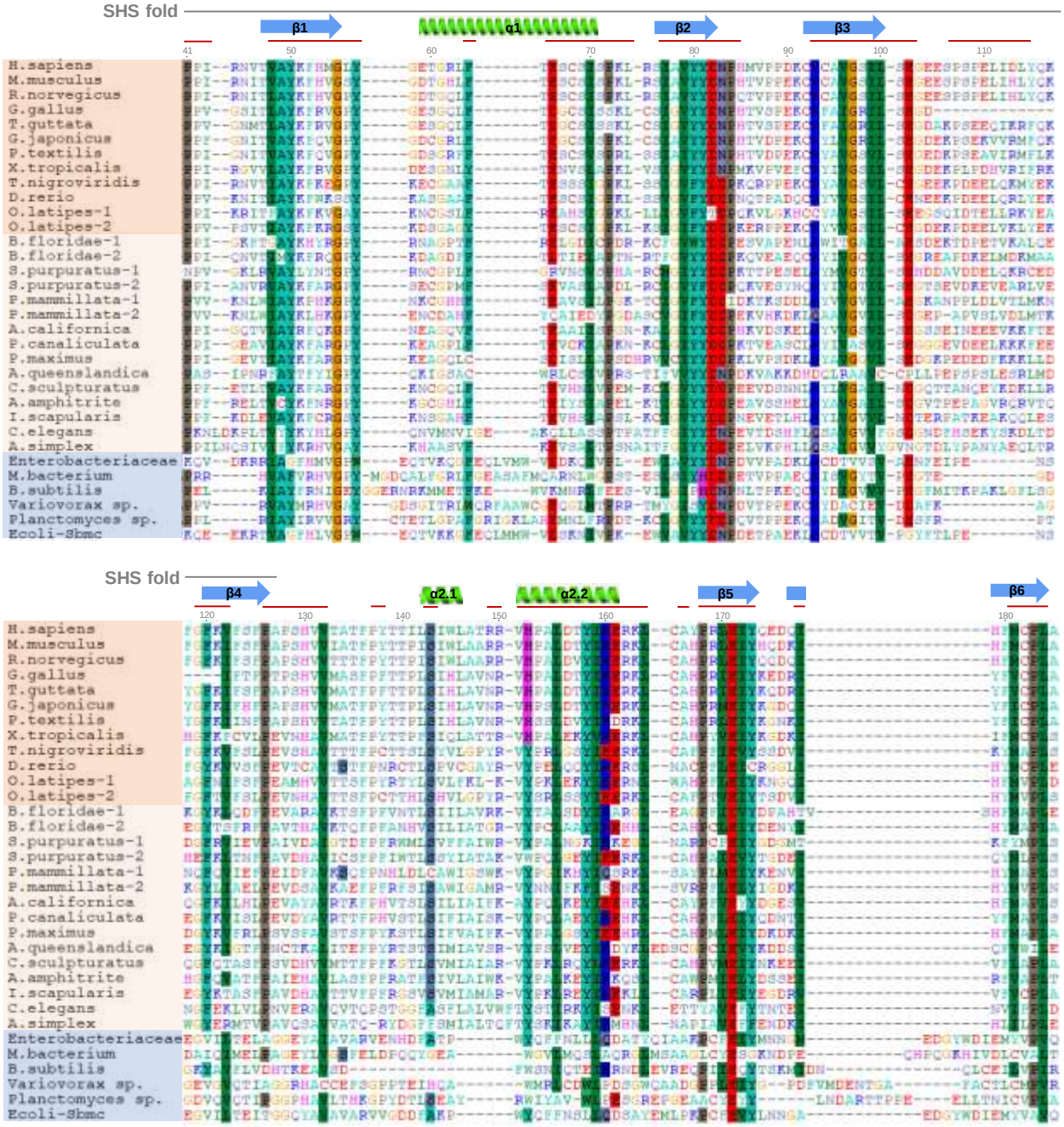

Figure 3. GyrI domain sequence alignments of TEX264 orthologs. The GyrI-like domain of human TEX264 corresponds to amino acids 41-185 of the full-length protein. Shown above the alignment (grey line), the SHS2 fold in human TEX264 corresponds to amino acids 21-127. The structure of human TEX264 according to 3D modelling is labelled for the corresponding protein sequence, where α-helices are shown in green and β-sheets in blue, as in the structural model of human TEX264 in Figure 2B. Red line designates conserved motifs and domains. TEX264 orthologs are shown in orange (dark: vertebrates; bright: invertebrates), while bacterial GyrI-domain containing proteins are shown in blue. Protein sequences were aligned using the MAFFT (Multiple Alignment using Fast Fourier Transform) alignment algorithm. Alignment quality score was assessed using the Guidance2 server and was 0.752, where 1 is maximum, indicating high alignment quality.

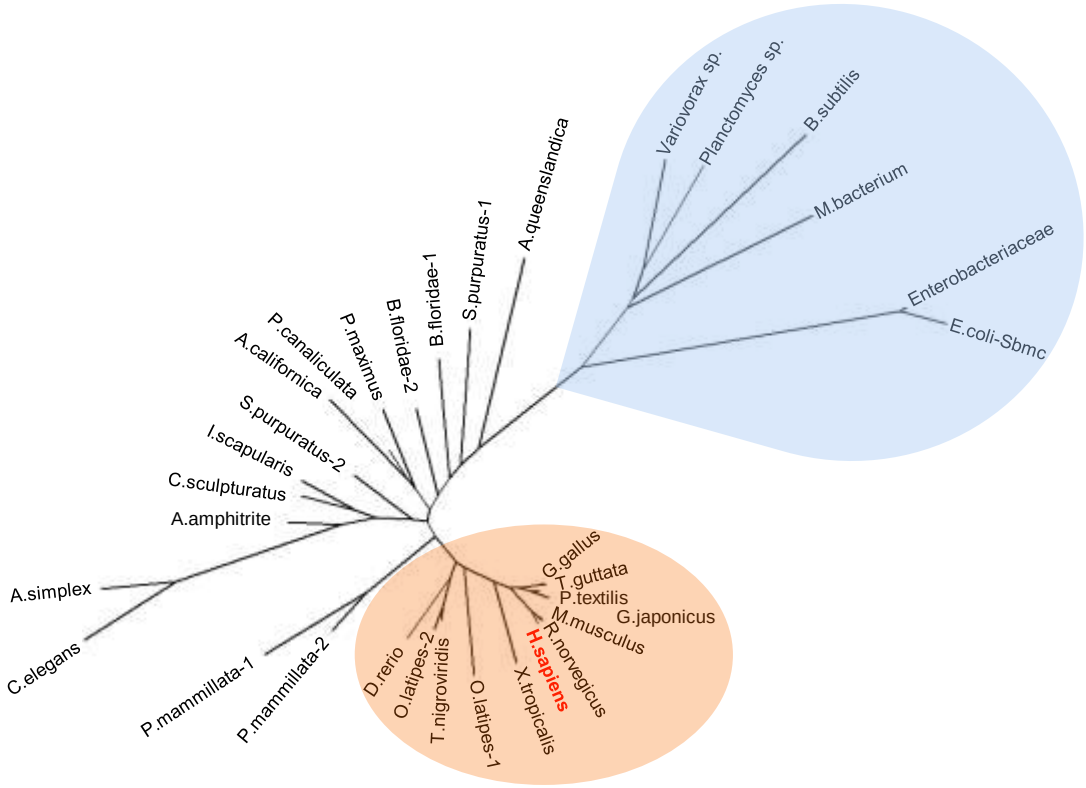

Figure 4. Phylogenetic analysis of TEX264 proteins. TEX264 orthologs in vertebrates are shown in orange, in invertebrates are shown colorless, while bacterial GyrI-domain containing proteins are shown in blue. Ftsa proteins were used as an outgroup (Figure S2). Full length protein sequences were aligned with the MAFFT (Multiple Alignment using Fast Fourier Transform) alignment algorithm. The phylogentic tree was constructed using the Maximum Likelihood method. Expanded phylogenetic tree with detailed methodology is shown in Figure S2.

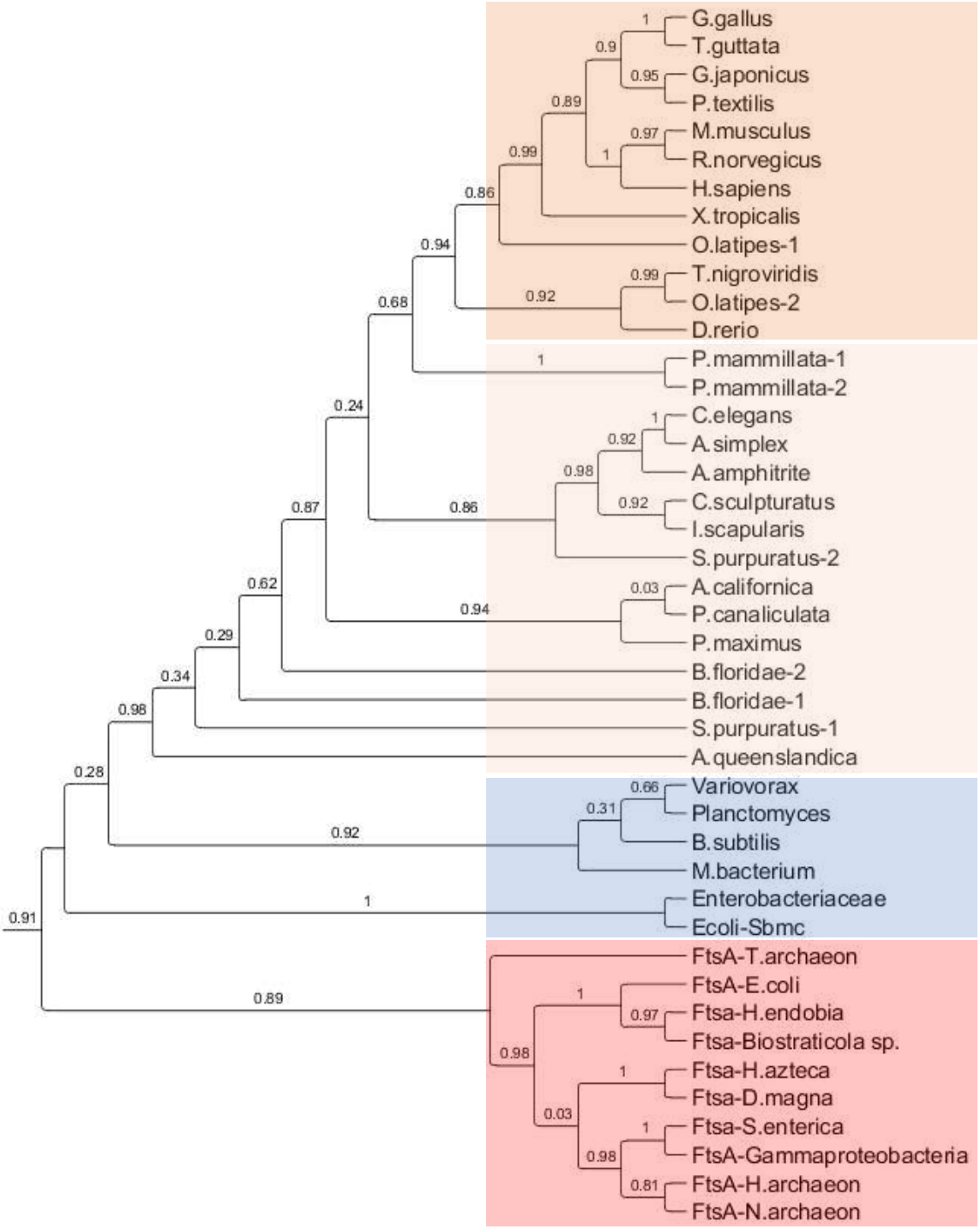

Figure S1. Phylogenetic analysis of TEX264 proteins. TEX264 orthologs are shown in orange (dark: vertebrates; bright: invertebrates). Bacterial GyrI domain-containing proteins are shown in blue. Ftsa proteins are used as an outgroup (indicated in red) because they are functionally different from GyrI domain-containing proteins, yet they contain the SHS2 module like GyrI proteins, making it possible to reach sufficiently good multiple protein alignment for tree building. Full-length protein sequences were aligned using MAFFT (Multiple Alignment using Fast Fourier Transform) alignment algorithm. Alignment quality score was assessed using Guidance2 server and was 0.563, where 1 is maximum, indicating sufficiently high alignment quality for tree building. The phylogenetic tree was constructed using the Maximum Likelihood method (PhyML software with optimised tree topology, LG model, 8 rates of categories, tree searching operation best of NNI&SPR (Nearest Neighbor Interchange & Subtree Pruning and Regrafting). Branch support Alrt values (Approximate likelihood-ratio test) are shown at tree nodes on a scale of 0-1, where 1 is maximum node confidence.

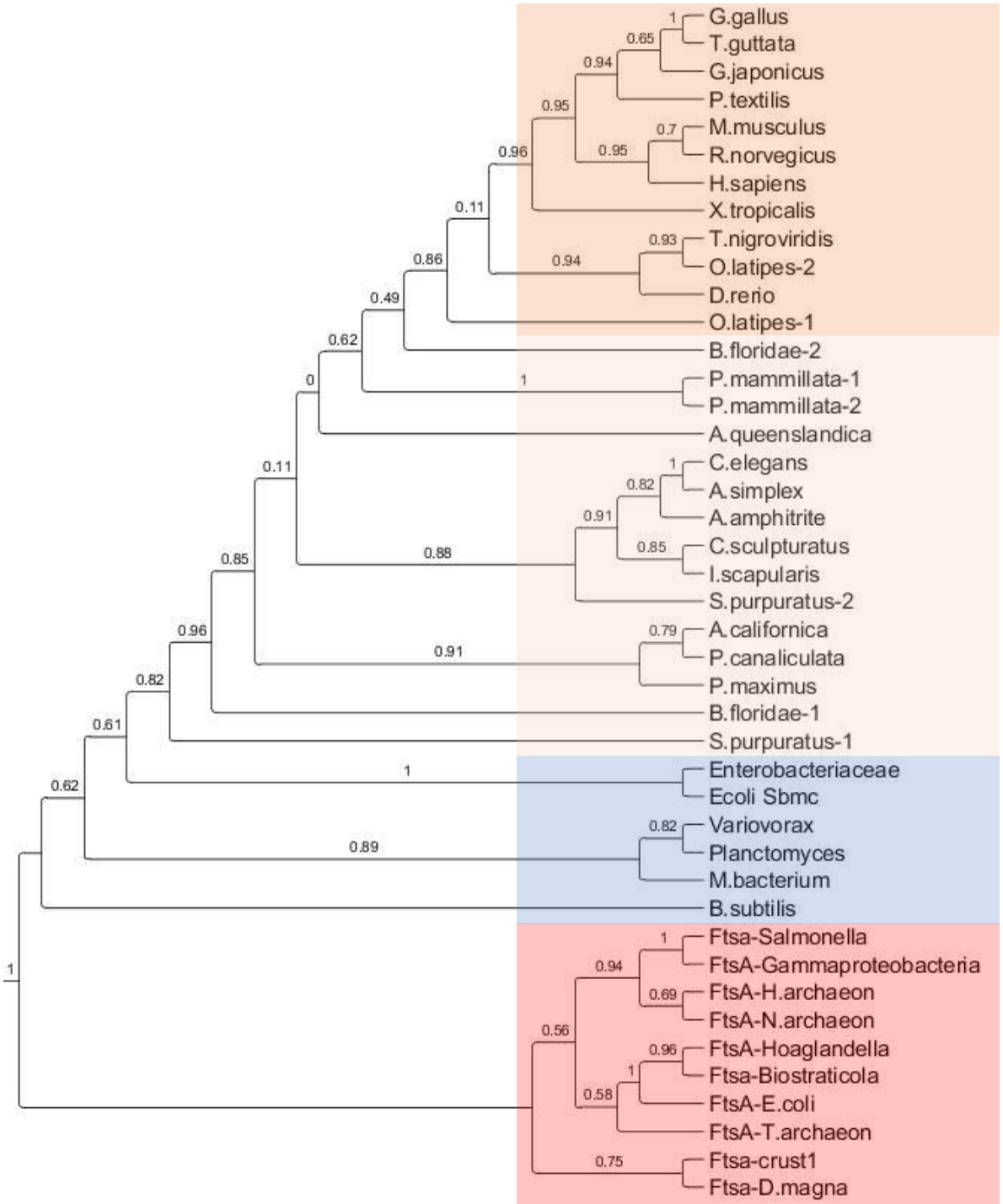

Figure S2. Phylogenetic analysis of GyrI domain in TEX264 orthologs. The GyrI domain of human TEX264 corresponds to amino acids 41-185. Vertebrate orthologs are shown in dark orange, invertebrate orthologs in bright orange, and bacterial GyrI domains are shown in blue. Ftsa proteins are used as an outgroup (indicated in red) because they are functionally different from GyrI domain-containing proteins, yet they contain the SHS2 module like GyrI proteins, making it possible to reach sufficiently good multiple protein alignment for subsequent tree building. GyrI domain protein sequences were aligned using MAFFT (Multiple Alignment using Fast Fourier Transform) alignment algorithm. Alignment quality score was assessed using Guidance2 server and was 0.619, where 1 is maximum, indicating sufficiently high alignment quality for tree building. The phylogenetic tree was constructed using the Maximum Likelihood method (PhyML software with optimised tree topology, LG model, 8 rates of categories, tree searching operation best of NNI&SPR (Nearest Neighbor Interchange & Subtree Pruning and Regrafting). Branch support, Alrt values (Approximate likelihood-ratio test) are shown at tree nodes on a scale of 0-1, where 1 is maximum node confidence.

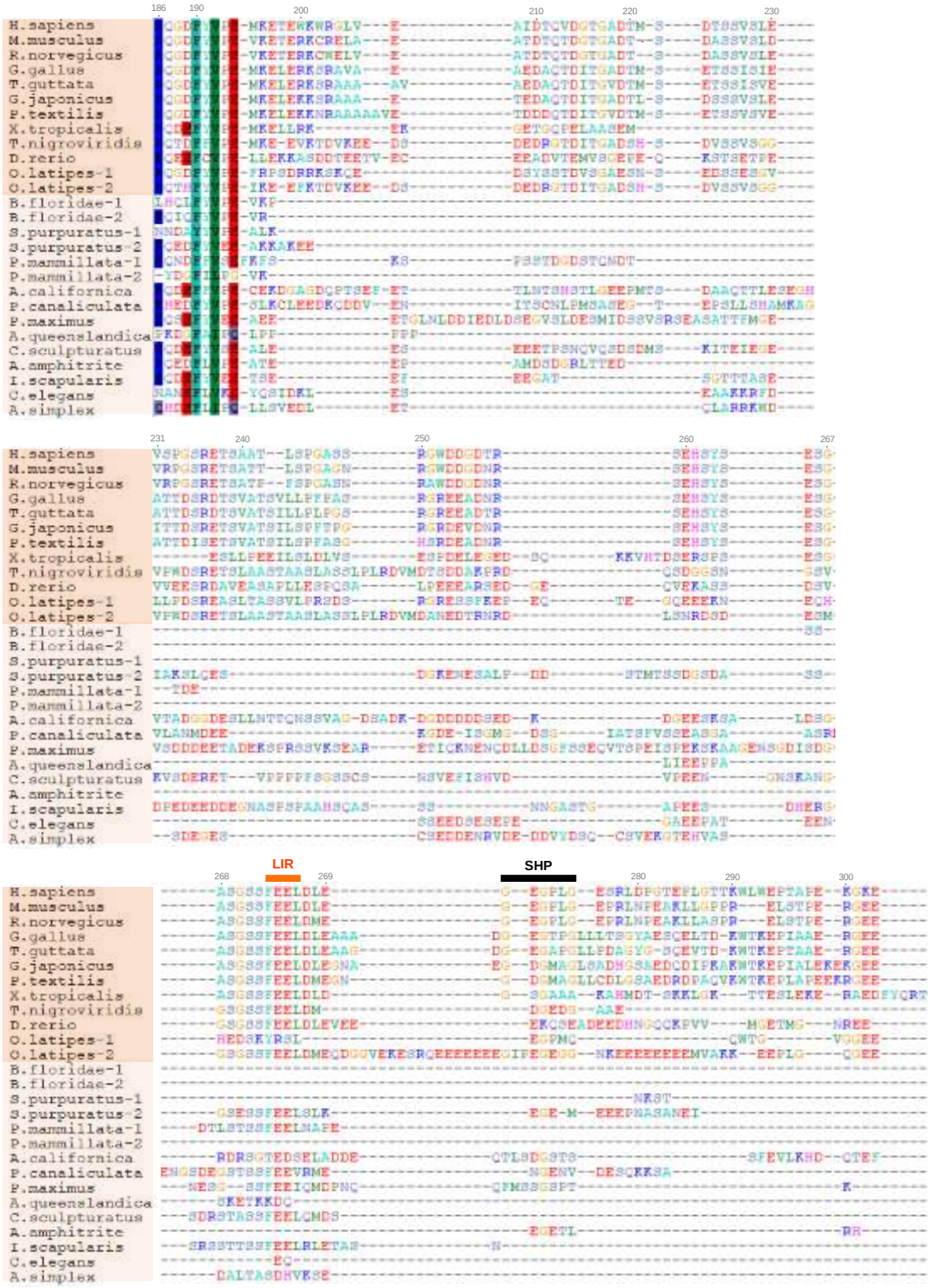

Figure S3. Amino acid sequence alignment of the C-terminal part of TEX264 orthologs. The C-terminal part of human TEX264 corresponds to amino acids 186-313 in the sequence of the full-length protein. The LIR motif, consisting of amino acids FEEL in vertebrates (in orange), corresponds to amino acids 273-276 in human TEX264. The SHP motif consisting of amino acids GEGPLG in mammals (in black) corresponds to amino acids 280-285 in human TEX264. Protein sequences were aligned using the MAFFT (Multiple Alignment using Fast Fourier Transform) alignment algorithm. Alignment quality score was assessed using the Guidance2 server and was 0.752, where 1 is maximum, indicating high alignment quality.

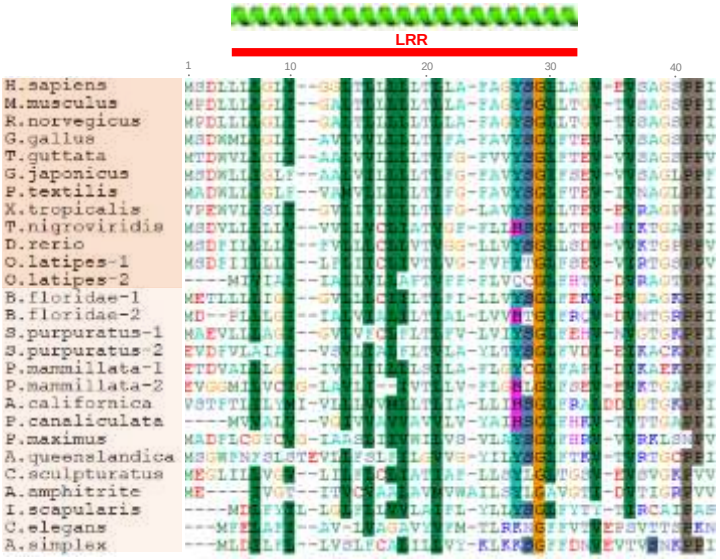

Figure S4. LRR domain alignment in TEX264 orthologs. The LRR domain in human TEX264 corresponds to amino acids 5-33 of the full-length protein. Conserved residues are highlighted. Protein sequences were aligned using the MAFFT (Multiple Alignment using Fast Fourier Transform) alignment algorithm. Alignment quality score was assessed using the Guidance2 server and was 0.752, where 1 is maximum, indicating high alignment quality.