

Exon 3 of the *NUMB* gene emerged in the chordate lineage coopting the NUMB protein to the regulation of MDM2

Stefano Confalonieri^{*,†}, Ivan Nicola Colaluca[†], Andrea Basile^{†,§}, Salvatore Pece^{†,§},
Pier Paolo Di Fiore^{*,†,§}.

[†] IEO, European Institute of Oncology IRCCS, Milan, Italy.

[§] Department of Oncology and Hemato-Oncology, Università degli Studi di Milano, 20142 Milan, Italy.

Running Title

The *NUMB* gene evolved to regulate MDM2 functions.

Key words: NUMB, MDM2, exon gain, chordate evolution, PTB domain.

***To whom correspondence should be addressed:**

Pier Paolo Di Fiore, European Institute of Oncology IRCCS, Via Ripamonti 435, 20141, Milan, Italy. Ph: +39 02 94375198. E-mail: pierpaolo.difiore@ieo.it

Stefano Confalonieri, European Institute of Oncology IRCCS, Via Ripamonti 435, 20141, Milan, Italy. Ph: +39 02 94372755. E-mail: stefano.confalonieri@ieo.it

Abstract

MDM2 regulates a variety of cellular processes through its dual protein:protein interaction and ubiquitin ligase activities. One major function of MDM2 is to bind and ubiquitinate P53, thereby regulating its proteasomal degradation. This function is in turn controlled by the cell fate determinant NUMB, which binds to and inhibits MDM2 via a short stretch of 11 amino acids, contained in its phosphotyrosine-binding (PTB) domain, encoded by exon 3 of the *NUMB* gene. The NUMB-MDM2-P53 circuitry is relevant to the specification of the stem cell fate and its subversion has been shown to be causal in breast cancer leading to the emergence of cancer stem cells. While extensive work on the evolutionary aspects of the MDM2/P53 circuitry has provided hints as to how these two proteins have evolved together to maintain conserved and linked functions, little is known about the evolution of the *NUMB* gene and, in particular, how it developed the ability to regulate MDM2 function. Here, we show that *NUMB* is a metazoan gene, which acquired exon 3 in the common ancestor of the Chordate lineage, first being present in the Cephalochordate and Tunicate subphyla, but absent in invertebrates. We provide experimental evidence showing that since its emergence, exon 3 conferred to the PTB domain of NUMB the ability to bind and to regulate MDM2 functions.

Introduction

The *NUMB* family of genes in vertebrates comprises two genes, *NUMB* and *NUMBL* (Gulino *et al.* 2010; Pece *et al.* 2011), the former being the best characterized. The NUMB protein is involved in diverse cellular phenotypes, including cell fate determination and maintenance of stem cell compartments, regulation of cell polarity and adhesion, and migration (Nishimura and Kaibuchi 2007; Bresciani *et al.* 2010; Pece *et al.* 2011; Sato *et al.* 2011; Zobel *et al.* 2018). Four major isoforms of NUMB have been described in mammals, produced by alternative splicing of two exons (Ex3 and Ex9), which mediate distinct cellular and developmental functions (Dho *et al.* 1999; Bechara *et al.* 2013; Rajendran *et al.* 2016). In particular, Ex3 encodes for a short stretch of 11 amino acids embedded in the phosphotyrosine-binding (PTB) domain, located at the N-terminus of the protein, thus yielding two alternative versions of the domain: an Ex3-containing and an Ex3-devoid PTB domain (henceforth, PTB-long and PTB-short, respectively) (Dho *et al.* 1999). The comparative structural features of these two PTB versions have recently been resolved (Colaluca *et al.* 2018). In contrast, NUMBL does not contain sequences resembling the alternatively spliced Ex3 or Ex9.

At the biological level, NUMB and NUMBL display both redundant and non-redundant functions. *NUMBL* KO mice do not display any overt phenotype, with the exception of a reduction in female fertility (Petersen *et al.* 2002). Conversely, *NUMB* KO mice are embryonic lethal and the phenotype is worsened in double NUMB/NUMBL KOs (Petersen *et al.* 2002). This might be due, in part, to the restricted pattern of expression of NUMBL (developing nervous system) *vs.* the more abundant and ubiquitous distribution of NUMB (Zhong *et al.* 1997). Biological differences between the two proteins might also underscore their different

impact on a number of signaling pathways: an issue presently unclear given the scarce characterization of NUMBL at the cellular and biochemical level vs. the wealth of evidence available for NUMB. A schematic representation of human NUMB and NUMBL, with their relevant protein domains and splicing patterns is shown in Figure 1A.

At the biochemical level, NUMB is involved in the regulation of several signaling circuitries, including Notch-, Hedgehog- and P53-dependent ones (Pece *et al.* 2004; Anderson *et al.* 2005; Di Marcotullio *et al.* 2006; Colaluca *et al.* 2008; Liu *et al.* 2011b; Pece *et al.* 2011; Tosoni *et al.* 2015; Guo *et al.* 2017). Many of the functions of NUMB have been directly or indirectly ascribed to its ability to regulate endocytic circuitries (Santolini *et al.* 2000; Sato *et al.* 2011), with particular emphasis on internalization, recycling routes and cell migration (Zobel *et al.* 2018). These endocytic properties of NUMB are evidenced by its numerous interactions with components of the endocytic machinery, including the endocytic adaptor AP2 (Santolini *et al.* 2000; Hutterer and Knoblich 2005; Song and Lu 2012), endocytic accessory proteins harboring EH domains (NUMB is an EH-binding protein) (Salcini *et al.* 1997; Smith *et al.* 2004), and others (Krieger *et al.* 2013; Wei *et al.* 2017). While these proteins bind to different regions of NUMB, their binding is apparently independent of the NUMB isoform, possibly attesting to an ancestral endocytic function(s) of NUMB that predates the appearance of its alternatively spliced isoforms.

In contrast, the binding of NUMB to the E3 ubiquitin ligase, MDM2, which results in inhibition of MDM2 enzymatic activity, is isoform-specific, depending on the presence of Ex3 in the PTB domain (Juven-Gershon *et al.* 1998; Colaluca *et al.* 2008; Colaluca *et al.* 2018). The net result of NUMB binding to MDM2 is stabilization of P53 due to inhibition of MDM2-dependent ubiquitination and ensuing proteasomal degradation of P53 (Colaluca *et al.* 2008). This is

relevant to breast cancer in which loss of NUMB expression results in reduction of P53 levels and P53-mediated responses, including sensitivity to genotoxic drugs and maintenance of homeostasis in the stem cell compartment (Pece *et al.* 2004; Tosoni *et al.* 2015; Tosoni *et al.* 2017). Surprisingly, this function of NUMB is entirely controlled by the Ex3-coded sequence, as demonstrated by our recent study showing that this short fragment is both necessary and sufficient for the interaction with MDM2 and for the regulation of P53 stability (Colaluca *et al.* 2018). This result raises interesting questions about the evolution of NUMB, and in particular of the Ex3-encoded sequence, in relationship to MDM2.

The first *MDM-like* gene appeared in Placozoa, where it is able to bind to and to regulate the levels of P53 (Siau *et al.* 2016; Aberg *et al.* 2017). This ancestral gene duplicated to generate *MDM2* and *MDM4* in vertebrates (Momand *et al.* 2011; Aberg *et al.* 2017), acquiring also several P53-unrelated functions (Carrillo *et al.* 2015; Saadatzaheh *et al.* 2017). The *NUMB* gene has been hypothesized to be a metazoan invention (Gazave *et al.* 2009); however, nothing is known about the appearance of the Ex3-containing NUMB isoforms (Dho *et al.* 1999). Since the P53/MDM2 signaling pathway dates back to early metazoan time and has then tightly co-evolved (Momand *et al.* 2011; Lane and Verma 2012; Aberg *et al.* 2017), it is tempting to speculate that also the *NUMB* gene followed a similar evolutionary route, as partner and regulator of MDM2 functions. Thus, important lessons could be learned by following the evolutionary trajectory of the mammalian *NUMB* Ex3, which could also be relevant to human health, since we previously demonstrated that the relative expression of Ex3-containing NUMB isoforms is clinically relevant in predicting the outcome of breast cancers (Colaluca *et al.* 2018). The present study was undertaken to investigate this issue.

Materials and Methods

IDENTIFICATION OF PTB DOMAIN-CONTAINING PROTEINS AND PHYLOGENETIC ANALYSIS OF NUMB HOMOLOGS

Protein, ESTs and mRNA sequence databases were downloaded from NCBI or ENSEMBL databases, from the organism specific databases (Borowiec *et al.* 2015), from Compagen (<http://www.compagen.org/>) or from the Joint Genome Institute (<http://genome.jgi.doe.gov/>) databases. Searches were carried out on 31 complete proteomes (Table S1). Each proteome was made non-redundant with CD-HIT (Li and Godzik 2006). The HMMER software version 3.1b2 (Eddy 2011) was used to search for PTB or IRS domain-containing proteins, with the SMART profile for the PTB or IRS domains. Domain annotations were obtained with the PFAM-A domain collection (Pfam 31.0 release) (Finn *et al.* 2016) (Table S3). ClustalX was used for all multiple alignments (Thompson *et al.* 2002) and visualized with Jalview (Waterhouse *et al.* 2009). Evolutionary tree analyses were carried out using the Minimum Evolution method with MEGA7 (Kumar *et al.* 2016). The bootstrap consensus tree inferred from 50 replicates is taken to represent the evolutionary history of the taxa analyzed. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates were collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (50 replicates) are shown next to the branches. The evolutionary distances were computed using the JTT matrix-based method and are in the units of the number of amino acid substitutions per site. The ME tree was searched using the Close-Neighbor-Interchange (CNI) algorithm at a search level of 1. The Neighbor-joining algorithm was used to generate the initial tree. Exon annotations were made with the GeneWise software (Birney *et al.* 2004), using the genome and the protein

databases from the same species, when available. To find possible Ex3-like sequences, if not present in the protein database, genome loci were additionally searched with GeneWise using the Human or the PTB-long domain-containing NUMB from the closest species.

BIOCHEMICAL STUDIES AND REAGENTS

Mammalian expression vectors for P53 and MDM2 were a gift from K. Helin (Cell Biology Program, Memorial Sloan Kettering Cancer Center, New York). Mammalian expression vectors encoding various NUMB fragments were engineered in a pcDNA vector in-frame with a FLAG-tag at the C-terminus. The *Ciona* cDNAs for MDM2 and NUMB was synthesized from Proteogenix and designed on the sequences AK116170 for MDM2, and AB210594 and XM_009859573 for NUMB. NUMB-PTB point mutations (Colaluca *et al.* 2018) were generated by the QuikChange strategy (Agilent Technologies) according to the manufacturer's instructions. The gene modifications and inserts were confirmed by DNA sequencing. Procedures for immunoblotting (IB) and immunoprecipitation (IP) were as described previously (Pece *et al.* 2004) according to standard procedures. For the *in vivo* p53 ubiquitination assay (Figure 3D), 10-cm plates of H1299 cells were transfected with 1 µg P53-coding vector, 1.3 µg MDM2-coding vector, 5 µg His6-Ub-coding vector and 5 µg NUMB-coding vector. H1299 were chosen since they are P53 null. After 48 h, ubiquitin conjugates were purified as described previously by (Colaluca *et al.* 2018). Antibodies were anti-FLAG M2-agarose affinity gel from Sigma-Aldrich, anti-Mdm2 (OP46) from EMD Millipore, anti-p53 (DO-1) from Santa Cruz Biotechnology, Inc., and anti-FLAG and anti-HA from Cell Signaling Technology, Inc.

DATA AVAILABILITY

Figure S1 contains the analysis of the PTB domain containing protein found in the analyzed organisms Figure S2 contains the alignment of the NUMB and NUMBL proteins found Figure S3 contains the alignment of vertebrate NUMB and NUMBL proteins together with the gene structure analysis. Figure S4 contains the alignment of the NumbF domain of NUMB and NUMBL proteins. Figure S5 contains the analysis of the two *Oikopleura dioica* NUMB genes and proteins Figure S6 contains the analysis of the NUMB and NUMBL proteins found in the Sea and Arctic Lamprey. File S1 contains two supplementary tables and legend for all supplementary figures and tables. Table S3 is an excel file containing contains the detailed output of the HMMER search (version 3.1b1) for each organism, together with accession numbers, protein domain composition and closest homologs identified by reverse BLAST search for each identified protein.

Supplementary tables and figures are available at FigShare:

<https://gsajournals.figshare.com/s/b221acedbeceb60d4906>

Results and Discussion

IDENTIFICATION OF NUMB HOMOLOGS IN EVOLUTION

The defining characteristic of NUMB (and NUMBL) is the presence of a PTB domain. To shed light on the evolution of the Ex3-coded sequence of NUMB, which is contained in its PTB domain, we identified all proteins potentially containing a PTB domain in representative and fully sequenced metazoans (Borowiec *et al.* 2015). We included in the search, proteins containing an insulin receptor substrate-like PTB domain (IRS), which is a variant of the canonical PTB domain (Eck *et al.* 1996; Uhlik *et al.* 2005). In addition to metazoans, we analyzed the proteome of the two choanoflagellates, *Monosiga brevicollis* and *Salpingoeca rosetta*, since it has been reported that choanoflagellates, the closest known unicellular relatives of animals, contain a surprising variety and high number of tyrosine kinase and PTB domains (Manning *et al.* 2008; Pincus *et al.* 2008). Finally, we included in our analysis the amoeboid protozoan *Dictyostelium discoideum* and the plant *Arabidopsis thaliana* as controls (see Table S1 for a complete list of organisms included in this study). Each identified protein was analyzed individually for domain composition using the PFAM and SMART protein domain databases and BLASTed against all vertebrate proteins to identify the true ortholog. This method has been reported to be more reliable than using automatic annotation software, such as Ensembl and InParanoid (Liu *et al.* 2011a).

By this approach, we confirmed the identity of the sequences retrieved with the initial HMM-profile searches (obtained using the HMMER 3.1b2 software) and defined each true ortholog of all human PTB- or IRS-containing proteins (Figure S1; see also Tables S2 and S3 for protein descriptions and the detailed HMM profile search results and domain annotation). NUMB

homologs were identified by a reciprocal best hits BLAST search in each organism, and by the presence of the NumbF domain (see below), and of DPF (Aspartate, Proline, Phenylalanine) and NPF (Arginine-Proline-Phenylalanine) motifs, which mediate NUMB (and NUMBL) interaction with the AP2 protein complex and EPS15, respectively (Salcini *et al.* 1997; Owen *et al.* 2000)(Figure 1A).

NUMB IS A METAZOAN GENE AND IS DUPLICATED IN THE VERTEBRATE LINEAGE

NUMB first appeared in evolution in Metazoa (Figure 1B and Figure S2), being present as a single gene in Ctenophora, as previously reported (Gazave *et al.* 2009), with the two distinct paralogs, *NUMB* and *NUMBL*, emerging only in vertebrates. The *NUMB* and *NUMBL* genes exhibit very similar exon–intron organizations with highly conserved intron positions and phases (Figure S3). In addition to the N-terminal PTB domain, NUMB proteins display a central conserved domain (NumbF, NUMB-Family protein domain) and a C-terminal proline-rich region (PRR). The NumbF domain is a region of unknown function, present only in the NUMB family of proteins (both NUMB and NUMBL), which is well conserved among NUMB homologs (Figure S4) and has been shown to be involved in the binding to RAB7A/B and ERBB2 (Hirai *et al.* 2017), while the PRR is poorly conserved (Figure S2). *NUMB* seems to be present only in one species of Porifera, the Slime sponge *Oscarella carmela*. We found no evidence of the existence of a *NUMB* gene in three other Porifera proteomes and genomes analyzed (Table 1), possibly due to incomplete genome assembly/annotation of the relative genomes, or to selective loss of the gene, similar to what has been observed for other genes in

demosponges (Riesgo *et al.* 2014). Notably, partial or complete gene loss is responsible for the lack of *MDM* and *P53/P63/P73* genes in Porifera (Aberg *et al.* 2017).

Most of the proteins identified were predictions from the genomes, and the existence of a gene, or of a predicted exon, does not necessarily imply that the mRNA will be expressed, or a particular exon included or excluded in the mature mRNA transcript. Thus, we investigated whether the identified NUMB proteins are actually transcribed. As summarized in Table 1, we analyzed all available transcript databases, i.e., mRNAs, ESTs, TSA (Transcriptome Shotgun Assembly) or RNA-Seq data for each species; the presence of a NUMB transcript was confirmed in all species with the exception of *Fugu rubripes*, *Petromyzon marinus* (Sea Lamprey), *Capitella telata*, *Trichoplax adherens*, *Amphimedon queenslandica*, *Xestospongia testudinaria*, *Stylissa carteri*, and non-metazoan organisms as expected. An EST for the *NUMB* gene of *Capitella capitata* is present in the database, suggesting the existence of a NUMB transcript in Polychaeta.

We were able to confirm that the *NUMB* gene and transcript first appeared in metazoans, but they do not seem to be present in all species of sponges. Invertebrates contain a single *NUMB* homolog, while a *NUMBL* paralog appears only in the vertebrate lineage. As in the case of *MDM2/MDM4* and *P53/63/73*, the invertebrate NUMB proteins are more similar to one another than to either of their vertebrate orthologs (Belyi *et al.* 2010; Tan *et al.* 2017). Indeed, in a phylogenetic analysis, only vertebrate NUMB and NUMBL proteins clearly separated into two distinct branches of the tree (Figure 1B).

Exceptions were noted in the Tunicate *Oikopleura dioica* and in Jawless Fishes, in particular:

- 1). In *Oikopleura dioica*, two predicted NUMB-related proteins are present belonging to two distinct genes that are both expressed, as supported by the presence of several ESTs (Figure S5).

At the protein level, the two paralogs display a low level of similarity (24% identity) and they seem more similar to human NUMB than to human NUMBL (Figure S5).

2). In the Jawless fish, Sea Lamprey (*Petromyzon marinus*), three genes encoding putative *NUMB* homologs are present. All three proteins are predicted (NUMB, ENSPMAP00000001971; NUMBL, ENSPMAP00000006450; NUMB-Like, ENSPMAP00000002145), and contain a PTB and a NumbF domain, in addition to NPF and DPF motifs (Figure 1B); however, ENSPMAP00000002145 displays a partial PTB and does not possess the starting methionine (Figure S6). In order to verify whether this occurrence is peculiar to the Sea Lamprey, we searched for NUMB homologs in the Japanese Lamprey (*Lethenteron camtschaticum*) Database (<http://jlampreygenome.imcb.a-star.edu.sg/>) and found three proteins, with similar domains and motif content (Accession numbers: NUMB, JL4438; NUMBL, JL5974; NUMB-Like, JL2165). Notably, the NUMB-Like predicted protein is a fragment, containing a partial PTB domain, lacking the starting methionine and the NumbF domain, but yet very similar to the Sea Lamprey NUMB-Like (Figure S6). Each Lamprey orthologs cluster together, as expected, but not with the respective human ortholog (Figure 1B and Figure S6). The absence of any mRNA or EST sequences does not allow verification of whether any of the three genes are actually transcribed (Table 1).

THE PTB-LONG ISOFORM OF NUMB IS A CHORDATE INVENTION

All PTB domains encompass a minimal fold containing two orthogonally arranged β -sheets composed of seven anti-parallel β -strands. The β -sheets pack against two α -helices, $\alpha 2$ and $\alpha 3$ (Li *et al.* 1998, Uhlik, 2005 #38; Colaluca *et al.* 2018) (Figure 2A). The PTB domains of NUMB and NUMBL are well conserved from their first appearance in Ctenophora (46% and 48%

identity between the Human and *Mnemiopsis leidyi* and *Pleurobrachia bachei* NUMB PTB domains, respectively), especially in the second half of the domain, encompassing strand $\beta 4$ to helix $\alpha 3$. At the genomic level, this sequence is coded by Ex5 and Ex6 of the *NUMB* genes (Figure 2B). Moreover, strand $\beta 5$ together with helix $\alpha 3$ (Figure 2C) plays a central role in the canonical binding properties to NPx(p)Y-containing peptides (Uhlik *et al.* 2005). Thus, it is not surprising that this region is the most conserved one in the domain. With the exception of *Caenorhabditis elegans*, *Drosophila melanogaster*, *Ciona intestinalis* and *Oikopleura dioica* (and possibly of Sea Lamprey, see below), the PTB domain of NUMB is coded by 6 or 5 exons, depending on the presence or absence of the exon corresponding to human Ex3 (Figure 2B).

Analysis of the PTB domain sequences of the various NUMB orthologs indicated that a PTB-long isoform first appeared in the chordate lineage, being present in *Branchiostoma* and *Ciona* (Cephalochordate and Urochordate/Tunicate) and absent in invertebrate proteins (Figure 2B, Figure S2). *NUMBL* orthologs do not encode for a PTB-long isoform (Figure S3) and there is no evidence supporting the existence of an alternative Ex3-like sequence in the genomes or transcriptomes analyzed.

Typically, the alternatively spliced Ex3 is represented by a short sequence of 33 nucleotides encoding for 11 amino acids (Figure 2B and Table 1). Exceptions were:

- 1). *Oikopleura dioica*, a larvacean tunicate similar to *Ciona intestinalis*, in which there are two distinct *NUMB-like* genes, both encoding a PTB-short-containing protein lacking the alternative exon. The *Oikopleura dioica NUMB* gene structure (GSOIDT00015583001) is remarkably different compared to other species (Figure 2B), and also with respect to the other *NUMB-Like* genes (GSOIDT00008916001) present in its genome (Figure S5). The two genes are very small (~2 Kb), and Ex2 and Ex4 (numbering refers to the human gene) are fused in a single exon.

These occurrences may not be surprising, since the genomes of tunicates exhibit great plasticity (Seo *et al.* 2001; Berna and Alvarez-Valin 2014) and the *Oikopleura dioica* genome was subject to massive intron loss (Denoeud *et al.* 2010). Of note, the *Oikopleura dioica* genome does not contain any P53- or MDM-related genes (data not shown).

2). *Ciona intestinalis* and *Branchiostoma floridae*, in which Ex3 (numbering refers to the human *NUMB* gene) is 36 and 39 nucleotides long, respectively (Figure 2B).

3). *Petromyzon marinus* (Sea Lamprey), representing the sole exception to the 33-nucleotide rule in vertebrates. The predicted NUMB protein displays a PTB insert of 14, rather than 11 amino acids, which - according to the ENSEMBL genome database - is coded by two short exons of 35 and 7 nucleotides, but only if the non-canonical splice site consensus is used in the protein prediction. If the canonical splice site rule (GT-AG) is employed, a similar exon organization of the *NUMB* gene is obtained (Figure 2B), but with two stop codons in the coding sequence. There are no mRNA sequences (or ESTs) to confirm the exact sequence of the transcript and thus of the protein. In addition, in this region of the genome there are several gaps in the assembly, thereby rendering it impossible to determine the reliability of this predicted protein sequence. Of note, the same predicted protein sequence and genomic organization (data not shown) are shared by the Japanese Lamprey (*Lethenteron camtschaticum*) (Figure S6).

4). *Takifugu rubripes*, in which the *NUMB* gene lacks Ex3, but this may be due to the presence of a gap in the genome assembly between Ex2 and Ex4. No Fugu mRNA or ESTs are available for *NUMB* (Table 1) to resolve this issue.

THE APPEARANCE OF A NUMB PTB-LONG ISOFORM IS CONNECTED WITH THE ABILITY TO BIND AND TO REGULATE MDM2

Since the PTB-long isoform of NUMB first appeared in the chordate lineage, possibly in a common ancestor of Cephalochordates and Urochordates/Tunicates, we investigated whether in the tunicate, *Ciona intestinalis*, the different NUMB isoforms have evolved to display differential binding to MDM2 and possibly to regulate its functions, as occurs in human (Colaluca *et al.* 2018). We choose this organism since it belongs to the earliest branch in the Chordate phylum (Corbo *et al.* 2001) and the *MDM2* and *NUMB* cDNA are available (Satou *et al.* 2002; Karakostis *et al.* 2016).

We transfected mammalian cells with constructs encoding one of the two versions of the *Ciona* NUMB PTB – Ex3-containing (c-PTB-long) and Ex3-devoid (c-PTB-short) (Figure 3A) – together with the *Ciona* MDM2. Only the c-PTB-long was able to coimmunoprecipitate efficiently with *Ciona* MDM2 (Figure 3B), suggesting that, since its initial appearance in evolution, the sequence coded by Ex3 confers to NUMB the ability to bind MDM2.

We then investigated whether the *Ciona* Ex3-coded sequence can substitute for the human Ex3-encoded sequence. To this end, we swapped the human Ex3 sequence for the *Ciona* one in the human NUMB PTB (Figure 3A). When expressed in HEK293 cells, this chimera (h-PTB/Ci-Ex3) coimmunoprecipitated with human MDM2 with an efficiency comparable to that of the WT human PTB-long (Figure 3C). In contrast, the human NUMB PTB-short and a human NUMB PTB-long mutant (PTB-long-R/K>E), previously engineered to abolish the interaction with MDM2 (Colaluca *et al.* 2018), were unable to coimmunoprecipitate with MDM2 (Figure 3C).

Next, we tested whether the *Ciona* Ex3 is also functionally equivalent to its human counterpart, by assessing the ability to inhibit MDM2-mediated P53 ubiquitination, as previously demonstrated (Colaluca *et al.* 2018). To this end, we set up an *in vivo* ubiquitination assay, by overexpressing in living H1299 cells, the human P53 and MDM2 proteins. Under these

conditions, P53 polyubiquitination was readily detectable (compare the first two lanes of Figure 3D). As previously reported (Colaluca *et al.* 2018), the cotransfection of human PTB-long, but not of human PTB-short, reverted this effect (Figure 3D). Similarly, the cotransfection of the construct encoding the human/*Ciona* chimera, h-PTB/Ci-Ex3, inhibited the polyubiquitination of p53, to an extent comparable to that of human WT PTB-long (Figure 3D). We concluded that the function of the NUMB Ex3 sequence to bind and to inhibit the E3 ligase activity of MDM2 is conserved throughout evolution.

In summary, our findings show that Ex3 of NUMB, and hence the PTB-long isoforms of the protein, is a chordate invention. Since its initial appearance, which probably occurred in a common ancestor of Cephalochordates and Urochordates/Tunicates, Ex3 confers to NUMB the ability to bind and to regulate the E3 ligase activity of MDM2. MDM2 possesses in mammals many functions related to its protein:protein interaction abilities (Riley and Lozano 2012) and although the best studied function of MDM2 involves the direct control of p53, additional substrates of its E3 ligase activity exist, most of which however converge on the regulation of the p53 pathway (Riley and Lozano 2012). Thus, one appealing possibility is that the emergence of NUMB Ex3 may be directly connected to the regulation of the MDM2/P53 pathway in evolution. So far, few studies have explored the functions of MDM2 or P53 proteins in non-vertebrate organisms. The finding that in the ancient invertebrate metazoan, *Trichoplax adhaerens*, an MDM2-related protein can bind to *Trichoplax* P53 and induce its degradation (Siau *et al.* 2016), suggests that the signaling pathway of P53- and MDM-related protein families is ancient and co-evolved, although in some organisms, one or both players have been lost (Belyi *et al.* 2010; Momand *et al.* 2011; Aberg *et al.* 2017). In agreement with the current hypothesis of the evolution of the P53 protein family, the P53/63/73 ancestral function was that of monitoring

the genetic quality in germline cells, as has been shown in *Caenorhabditis elegans* (Derry *et al.* 2001), in the sea anemone *Nematostella vectensis* (Pankow and Bamberger 2007) and in *Ciona intestinalis* (Noda 2011). Later in evolution, P63 and P73 preserved the ancestral functions in the development of tissues and organs (Levrero *et al.* 2000), becoming partly independent from MDM2 with respect to ubiquitination and degradation (Balint *et al.* 1999; Zeng *et al.* 1999; Little and Jochemsen 2001; Kadakia *et al.* 2001; Galli *et al.* 2010; Armstrong *et al.* 2016), whereas P53 developed new functions, becoming the guardian of the somatic genome and a tumor suppressor, with MDM2 acting as its primary “director” (Belyi *et al.* 2010; Aberg *et al.* 2017). The emergence of a novel NUMB splicing isoform as a partner and regulator of MDM2 in Chordata, suggests a specific role for NUMB in the regulation of MDM2 signaling pathways, which in higher Metazoa culminate in the fine-tuning of P53 activity in somatic tissues.

Acknowledgments

We thank R. Gunby for critically editing the manuscript. **Funding:** this work was supported by grants from the Associazione Italiana per la Ricerca sul Cancro (AIRC; IG 11904 to S.P., IG 18988 to P.P.D.F., and MCO 10.000 to S.P. and P.P.D.F.); the Italian Ministry of University and Scientific Research (MIUR) to P.P.D.F., the Italian Ministry of Health to P.P.D.F. and I. N. C. (RF-2016-02361540). Author contribution: Conceptualization, S. Confalonieri and P.P. Di Fiore; Investigation, S. Confalonieri, A. Basile and I.N. Colaluca; Supervision, S. Pece, and P.P. Di Fiore; Funding acquisition, I.N. Colaluca, S. Pece, and P.P. Di Fiore; Writing, all authors.

Figure Legends

Figure 1. Evolution of the NUMB protein family.

A. Structure and splicing isoforms of human NUMB and NUMBL proteins. PTB domains are in green with the portion encoded by Ex3 highlighted in yellow, the NumbF domain is in blue, NPF and DPF/W motifs, responsible for binding to endocytic proteins, are in red and gray respectively. The positions of Ex3 and Ex9 coded inserts are indicated. **B.** Minimum Evolution Phylogenetic tree of the NUMB Protein Family. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test are shown next to the branches. Protein domains and motifs are depicted as in (A). Branches in red show NUMBL proteins. Only the longest isoform in each organism is shown.

Figure 2. Ex3 of the human *NUMB* gene is a Chordate invention.

A. Structure of the PTB-long domain of human NUMB (Colaluca *et al.* 2018); the Ex3 encoded sequence is highlighted in red. **B.** Exon and gene length of various *NUMB* homologs for which the genomic sequence is available; since most of the proteins are predicted from the genome, we calculated the length of each gene from the starting methionine to the stop codon of the protein. Exons coding for the PTB domain are in gray. (1) The *Takifugu rubripes* genome sequence of the *NUMB* locus has a gap in the assembly between Ex2 and Ex4 and therefore the length and presence of Ex3 is undetermined (N.A.: not available). (2) The *Petromyzon marinus* NUMB protein was predicted *ab initio* by the ENSEMBL pipeline with no support of any mRNA or EST, and the locus contains several gaps in the assembly: exons may be longer (+) or contain stop codons (*). Therefore, both the protein sequence and the genomic organization may not be accurate. (3) *Oikopleura dioica*, *Ciona intestinalis*, *Caenorhabditis elegans* and *Drosophila*

melanogaster display a different genomic organization of the *NUMB* locus with respect to the vertebrate and other invertebrate loci, despite a very high similarity of the NUMB encoded proteins (Figure S2). (4) Ex4 (Ex5 with respect to the human gene) is not present in the *Hydra vulgaris* genome, possibly due to a gap in the assembly, thus exact length of flanking exons is not certain (N.A.: not available). (5) The first exon of the *Oscarella carmela* *NUMB* gene is incomplete, due to a gap in the assembly of the genome, and the genomic sequence of last exon(s) is not available (N.A.: not available). Exon phase numbers are reported between corresponding exons. C. Multiple alignment of the PTB domains of the various NUMB proteins. The alignment was manually adjusted to reflect the splicing position of invertebrate NUMB PTB domains between Ex2 and Ex4 (exon numbering is with respect to the human *NUMB* gene). Above the alignment, the secondary structure is depicted as determined by the 3D crystal structure, and residues involved in the binding to (p)Y-containing peptides are indicated by an asterisk (Colaluca *et al.* 2018). Splice site positions and exon phase within the human protein sequence are indicated by arrowheads. Gaps evidenced in red in the *Oikopleura* and *Caenorhabditis* proteins in the alignment, indicate the portion of the protein coded by Ex2 and Ex4 (with respect to the human gene) that is produced by a single exon (underlined in B). See also Figure S2 for the complete protein alignment.

Figure 3. Ex3 of the *Ciona NUMB* gene can substitute for mammalian Ex3 to direct binding to MDM2 and regulation of P53.

A. Schematic representation of the constructs used. c-PTB-long, *Ciona intestinalis* NUMB PTB long isoform; c-PTB-short, *Ciona intestinalis* NUMB PTB short isoform; h-PTB-long, human NUMB PTB long isoform; h-PTB-short, human NUMB PTB short isoform; h-PTB/Ci-Ex3, human NUMB short PTB isoform with the insertion of the *Ciona* Ex3 coded protein sequence; h-

PTB-long R,K>E, human NUMB PTB long isoform with R69E-K70E-K73E-K78E substitutions (Colaluca *et al.* 2018). **B.** HEK293 cells were individually transfected with the indicated FLAG-tagged NUMB *Ciona* PTB (with or without the *Ciona* Ex3: c-PTB-long and c-PTB-short, respectively) and HA-tagged *Ciona* MDM2. IP (anti-FLAG) and IB were as shown. Intervening lanes were spliced out; i.e., long exposure. A digitally-enhanced image (curves tool in Photoshop CS6: shadow 152; highlight 165) of the anti-FLAG protein extracts is also shown. **C.** HEK293 cells were transfected with the indicated FLAG-tagged PTB constructs and HA-tagged human MDM2. IP (anti-FLAG) and IB were as shown (i.e., long exposure). **D.** MDM2-mediated P53 *in vivo* ubiquitination in H1299 cells transfected as indicated on the top. The Histidine-ubiquitinated proteins were purified as described in Materials and Methods and immunoblotted with the P53 antibody. The corresponding lysates were immunoblotted as indicated. A ratio Mdm2/p53 1.3:1 was used in these competition ubiquitination assays; at this protein ratio, only p53 ubiquitination and not its degradation is appreciable, as previously shown (Colaluca *et al.* 2018).

Table 1. Protein and transcript analysis of NUMB in evolution.

Species	Protein Analysis			Transcript Analysis			
	N. of Proteins	NUMB Protein	NUMB PTB-Long Isoform	N. of cDNA	N. of ESTs or RNA-SEQ data	NUMB mRNA	NUMB PTB- Long Isoform mRNA
<i>Homo sapiens</i>	31,283	YES	YES	9,774,024	8,704,790	YES	YES
<i>Taeniopygia guttata</i> (Song Bird)	17,758	YES	YES	79,179	92,142	YES	NO
<i>Gallus gallus</i>	20,979	YES	YES	387,867	600,434	YES	YES
<i>Anolis carolinensis</i> (Lizard)	20,396	YES	YES	8,588,979	156,802	YES	YES
<i>Xenopus tropicalis</i>	22,785	YES	YES	78,985	1,271,480	YES	YES
<i>Fugu rubripes</i>	21,423	YES	N.A. ^a	97,191	566,629	NO	NO
<i>Danio rerio</i>	27,033	YES	YES	220,453	1,488,339	YES	YES
<i>Callorhinchus milii</i> (Elephant shark)	18,806	YES	YES	91,515	109,965	YES	YES
<i>Petromyzon marinus</i> (Sea Lamprey)	11,442	YES	YES	28,215	120,731	NO	NO
<i>Ciona intestinalis</i>	15,453	YES	YES	38,404	1,205,674	YES	YES
<i>Oikopleura dioica</i>	18,020	YES	NO	5,972	207,355	YES	NO
<i>Branchiostoma floridae</i>	39,596	YES	YES	30,270	334,502+ RNA-Seq	YES	YES
<i>S. kowalevskii</i> (Acorn worm)	20,381	YES	NO	85,678	202,190	YES	NO
<i>S. purpuratus</i> (purple sea urchin)	25,045	YES	NO	187,092	141,833	YES	NO
<i>Lottia gigantea</i> (Owl Limpet)	34,510	YES	NO	32,851	252,091	YES	NO
<i>Capitella teleta</i> (polychaete worm)	45,370	YES	NO	20,969	138,404	NO ^b	NO
<i>Caenorhabditis Elegans</i>	20,981	YES	NO	150,848	538,795	YES	NO
<i>Drosophila melanogaster</i>	15,322	YES	NO	320,780	847,942	YES	NO
<i>Nematostella vectensis</i> (sea anemone)	44,353	YES	NO	532,252	163,314	YES	NO
<i>Hydra vulgaris</i>	17,266	YES	NO	172,012	184,731	YES	NO
<i>Trichoplax adhaerens</i>	11,520	YES	NO	14,453	11,498	NO	NO
<i>Amphimedon queenslandica</i> (sponge)	13,721	NO	NO	1,042	83,040	NO	NO
<i>Oscarella carmela</i> (slime sponge)	11,152	YES	NO	22	11,176 + RNA-Seq	YES	NO
<i>Xestospongia testudinaria</i> (barrel sponge)	21,880	NO	NO	N.A.	22,337	NO	NO
<i>Stylissa carteri</i> (elephant ear sponge)	24,925	NO	NO	N.A.	26,967	NO	NO
<i>Mnemiopsis leidyi</i> (Sea Walnut)	16,548	YES	NO	5,493	15,752 + RNA-Seq	YES	NO
<i>Pleurobrachia bachei</i> (Sea gooseberry)	19,524	YES	NO	115	298 + RNA-Seq	YES	NO
<i>Salpingoeca rosetta</i>	11,731	NO	NO	N.A.	7	N.A.	N.A.
<i>Monosiga Brevicollis</i>	9,196	NO	NO	N.A.	29,595	N.A.	N.A.
<i>Dictyostelium discoideum</i>	11,914	NO	NO	N.A.	N.A.	N.A.	N.A.

NUMB protein sequences retrieved with the initial HMM search were analyzed for the presence of a PTB-long isoform at the protein and transcript level: numbers of total sequences available are indicated. ^a No ESTs or mRNA sequence were found for the Fugu NUMB transcript, therefore the existence of Ex3 cannot be assessed. ^b An EST corresponding to the NUMB predicted transcript of *Capitella telata* is present among the ESTs of *Capitella capitata* (JGI_CAWU12746.fwd, 100% ID), suggesting the existence of a NUMB transcript in Polychaeta Worms. N.A., no sequence available for the analysis.

References

- Aberg, E., F. Saccoccia, M. Grabherr, W. Y. J. Ore, P. Jemth *et al.*, 2017 Evolution of the p53-MDM2 pathway. *BMC Evol Biol* **17**: 177.
- Anderson, A. C., E. A. Kitchens, S. W. Chan, C. St Hill, Y. N. Jan *et al.*, 2005 The Notch regulator Numb links the Notch and TCR signaling pathways. *J Immunol* **174**: 890-897.
- Armstrong, S. R., H. Wu, B. Wang, Y. Abuetabh, C. Sergi *et al.*, 2016 The Regulation of Tumor Suppressor p63 by the Ubiquitin-Proteasome System. *Int J Mol Sci* **17**.
- Balint, E., S. Bates and K. H. Vousden, 1999 Mdm2 binds p73 alpha without targeting degradation. *Oncogene* **18**: 3923-3929.
- Bechara, E. G., E. Sebestyen, I. Bernardis, E. Eyraas and J. Valcarcel, 2013 RBM5, 6, and 10 differentially regulate NUMB alternative splicing to control cancer cell proliferation. *Mol Cell* **52**: 720-733.
- Belyi, V. A., P. Ak, E. Markert, H. Wang, W. Hu *et al.*, 2010 The origins and evolution of the p53 family of genes. *Cold Spring Harb Perspect Biol* **2**: a001198.
- Berna, L., and F. Alvarez-Valin, 2014 Evolutionary genomics of fast evolving tunicates. *Genome Biol Evol* **6**: 1724-1738.
- Birney, E., M. Clamp and R. Durbin, 2004 GeneWise and Genomewise. *Genome Res* **14**: 988-995.
- Borowiec, M. L., E. K. Lee, J. C. Chiu and D. C. Plachetzki, 2015 Extracting phylogenetic signal and accounting for bias in whole-genome data sets supports the Ctenophora as sister to remaining Metazoa. *BMC Genomics* **16**: 987.
- Bresciani, E., S. Confalonieri, S. Cermenati, S. Cimbro, E. Foglia *et al.*, 2010 Zebrafish numb and numblake are involved in primitive erythrocyte differentiation. *PLoS One* **5**: e14296.
- Carrillo, A. M., A. Bouska, M. P. Arrate and C. M. Eischen, 2015 Mdmx promotes genomic instability independent of p53 and Mdm2. *Oncogene* **34**: 846-856.
- Colaluca, I. N., A. Basile, L. Freiburger, V. D'Uva, D. Disalvatore *et al.*, 2018 A Numb-Mdm2 fuzzy complex reveals an isoform-specific involvement of Numb in breast cancer. *J Cell Biol* **217**: 745-762.

- Colaluca, I. N., D. Tosoni, P. Nuciforo, F. Senic-Matuglia, V. Galimberti *et al.*, 2008 NUMB controls p53 tumour suppressor activity. *Nature* **451**: 76-80.
- Corbo, J. C., A. Di Gregorio and M. Levine, 2001 The ascidian as a model organism in developmental and evolutionary biology. *Cell* **106**: 535-538.
- Denoeud, F., S. Henriot, S. Mungpakdee, J. M. Aury, C. Da Silva *et al.*, 2010 Plasticity of animal genome architecture unmasked by rapid evolution of a pelagic tunicate. *Science* **330**: 1381-1385.
- Derry, W. B., A. P. Putzke and J. H. Rothman, 2001 *Caenorhabditis elegans* p53: role in apoptosis, meiosis, and stress resistance. *Science* **294**: 591-595.
- Dho, S. E., M. B. French, S. A. Woods and C. J. McGlade, 1999 Characterization of four mammalian numb protein isoforms. Identification of cytoplasmic and membrane-associated variants of the phosphotyrosine binding domain. *J Biol Chem* **274**: 33097-33104.
- Di Marcotullio, L., E. Ferretti, A. Greco, E. De Smaele, A. Po *et al.*, 2006 Numb is a suppressor of Hedgehog signalling and targets Gli1 for Itch-dependent ubiquitination. *Nat Cell Biol* **8**: 1415-1423.
- Eck, M. J., S. Dhe-Paganon, T. Trub, R. T. Nolte and S. E. Shoelson, 1996 Structure of the IRS-1 PTB domain bound to the juxtamembrane region of the insulin receptor. *Cell* **85**: 695-705.
- Eddy, S. R., 2011 Accelerated Profile HMM Searches. *PLoS Comput Biol* **7**: e1002195.
- Finn, R. D., P. Coghill, R. Y. Eberhardt, S. R. Eddy, J. Mistry *et al.*, 2016 The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res* **44**: D279-285.
- Galli, F., M. Rossi, Y. D'Alessandra, M. De Simone, T. Lopardo *et al.*, 2010 MDM2 and Fbw7 cooperate to induce p63 protein degradation following DNA damage and cell differentiation. *J Cell Sci* **123**: 2423-2433.
- Gazave, E., P. Lapebie, G. S. Richards, F. Brunet, A. V. Ereskovsky *et al.*, 2009 Origin and evolution of the Notch signalling pathway: an overview from eukaryotic genomes. *BMC Evol Biol* **9**: 249.
- Gulino, A., L. Di Marcotullio and I. Screpanti, 2010 The multiple functions of Numb. *Exp Cell Res* **316**: 900-906.
- Guo, Y., K. Zhang, C. Cheng, Z. Ji, X. Wang *et al.*, 2017 Numb(-/low) Enriches a Castration-Resistant Prostate Cancer Cell Subpopulation Associated with Enhanced Notch and Hedgehog Signaling. *Clin Cancer Res* **23**: 6744-6756.

- Hirai, M., Y. Arita, C. J. McGlade, K. F. Lee, J. Chen *et al.*, 2017 Adaptor proteins NUMB and NUMBL promote cell cycle withdrawal by targeting ERBB2 for degradation. *J Clin Invest* **127**: 569-582.
- Hutterer, A., and J. A. Knoblich, 2005 Numb and alpha-Adaptin regulate Sanpodo endocytosis to specify cell fate in *Drosophila* external sensory organs. *EMBO Rep* **6**: 836-842.
- Juven-Gershon, T., O. Shifman, T. Unger, A. Elkeles, Y. Haupt *et al.*, 1998 The Mdm2 oncoprotein interacts with the cell fate regulator Numb. *Mol Cell Biol* **18**: 3974-3982.
- Kadakia, M., C. Slader and S. J. Berberich, 2001 Regulation of p63 function by Mdm2 and MdmX. *DNA Cell Biol* **20**: 321-330.
- Karakostis, K., A. Ponnuswamy, L. T. Fusee, X. Bailly, L. Laguerre *et al.*, 2016 p53 mRNA and p53 Protein Structures Have Evolved Independently to Interact with MDM2. *Mol Biol Evol* **33**: 1280-1292.
- Krieger, J. R., P. Taylor, A. S. Gajadhar, A. Guha, M. F. Moran *et al.*, 2013 Identification and selected reaction monitoring (SRM) quantification of endocytosis factors associated with Numb. *Mol Cell Proteomics* **12**: 499-514.
- Kumar, S., G. Stecher and K. Tamura, 2016 MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* **33**: 1870-1874.
- Lane, D. P., and C. Verma, 2012 Mdm2 in evolution. *Genes Cancer* **3**: 320-324.
- Leverro, M., V. De Laurenzi, A. Costanzo, J. Gong, J. Y. Wang *et al.*, 2000 The p53/p63/p73 family of transcription factors: overlapping and distinct functions. *J Cell Sci* **113** (Pt 10): 1661-1670.
- Li, S. C., C. Zwahlen, S. J. Vincent, C. J. McGlade, L. E. Kay *et al.*, 1998 Structure of a Numb PTB domain-peptide complex suggests a basis for diverse binding specificity. *Nat Struct Biol* **5**: 1075-1083.
- Li, W., and A. Godzik, 2006 Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**: 1658-1659.
- Little, N. A., and A. G. Jochemsen, 2001 Hdmx and Mdm2 can repress transcription activation by p53 but not by p63. *Oncogene* **20**: 4576-4580.
- Liu, B. A., E. Shah, K. Jablonowski, A. Stergachis, B. Engelmann *et al.*, 2011a The SH2 domain-containing proteins in 21 species establish the provenance and scope of phosphotyrosine signaling in eukaryotes. *Sci Signal* **4**: ra83.

- Liu, L., F. Lanner, U. Lendahl and D. Das, 2011b Numbl like and Numb differentially affect p53 and Sonic Hedgehog signaling. *Biochem Biophys Res Commun* **413**: 426-431.
- Manning, G., S. L. Young, W. T. Miller and Y. Zhai, 2008 The protist, *Monosiga brevicollis*, has a tyrosine kinase signaling network more elaborate and diverse than found in any known metazoan. *Proc Natl Acad Sci U S A* **105**: 9674-9679.
- Momand, J., A. Villegas and V. A. Belyi, 2011 The evolution of MDM2 family genes. *Gene* **486**: 23-30.
- Nishimura, T., and K. Kaibuchi, 2007 Numb controls integrin endocytosis for directional cell migration with aPKC and PAR-3. *Dev Cell* **13**: 15-28.
- Noda, T., 2011 The maternal genes Ci-p53/p73-a and Ci-p53/p73-b regulate zygotic ZicL expression and notochord differentiation in *Ciona intestinalis* embryos. *Dev Biol* **360**: 216-229.
- Owen, D. J., Y. Vallis, B. M. Pearse, H. T. McMahon and P. R. Evans, 2000 The structure and function of the beta 2-adaptin appendage domain. *EMBO J* **19**: 4216-4227.
- Pankow, S., and C. Bamberger, 2007 The p53 tumor suppressor-like protein nvp63 mediates selective germ cell death in the sea anemone *Nematostella vectensis*. *PLoS One* **2**: e782.
- Pece, S., S. Confalonieri, R. R. P and P. P. Di Fiore, 2011 NUMB-ing down cancer by more than just a NOTCH. *Biochim Biophys Acta* **1815**: 26-43.
- Pece, S., M. Serresi, E. Santolini, M. Capra, E. Hulleman *et al.*, 2004 Loss of negative regulation by Numb over Notch is relevant to human breast carcinogenesis. *J Cell Biol* **167**: 215-221.
- Petersen, P. H., K. Zou, J. K. Hwang, Y. N. Jan and W. Zhong, 2002 Progenitor cell maintenance requires numb and numbl like during mouse neurogenesis. *Nature* **419**: 929-934.
- Pincus, D., I. Letunic, P. Bork and W. A. Lim, 2008 Evolution of the phospho-tyrosine signaling machinery in premetazoan lineages. *Proc Natl Acad Sci U S A* **105**: 9680-9684.
- Rajendran, D., Y. Zhang, D. M. Berry and C. J. McGlade, 2016 Regulation of Numb isoform expression by activated ERK signaling. *Oncogene* **35**: 5202-5213.
- Riesgo, A., N. Farrar, P. J. Windsor, G. Giribet and S. P. Leys, 2014 The analysis of eight transcriptomes from all poriferan classes reveals surprising genetic complexity in sponges. *Mol Biol Evol* **31**: 1102-1120.
- Riley, M. F., and G. Lozano, 2012 The Many Faces of MDM2 Binding Partners. *Genes Cancer* **3**: 226-239.

- Saadatzadeh, M. R., A. N. Elmi, P. H. Pandya, K. Bijangi-Vishehsaraei, J. Ding *et al.*, 2017 The Role of MDM2 in Promoting Genome Stability versus Instability. *Int J Mol Sci* **18**.
- Salcini, A. E., S. Confalonieri, M. Doria, E. Santolini, E. Tassi *et al.*, 1997 Binding specificity and in vivo targets of the EH domain, a novel protein-protein interaction module. *Genes Dev* **11**: 2239-2249.
- Santolini, E., C. Puri, A. E. Salcini, M. C. Gagliani, P. G. Pelicci *et al.*, 2000 Numb is an endocytic protein. *J Cell Biol* **151**: 1345-1352.
- Sato, K., T. Watanabe, S. Wang, M. Kakeno, K. Matsuzawa *et al.*, 2011 Numb controls E-cadherin endocytosis through p120 catenin with aPKC. *Mol Biol Cell* **22**: 3103-3119.
- Satou, Y., L. Yamada, Y. Mochizuki, N. Takatori, T. Kawashima *et al.*, 2002 A cDNA resource from the basal chordate *Ciona intestinalis*. *Genesis* **33**: 153-154.
- Seo, H. C., M. Kube, R. B. Edvardsen, M. F. Jensen, A. Beck *et al.*, 2001 Miniature genome in the marine chordate *Oikopleura dioica*. *Science* **294**: 2506.
- Siau, J. W., C. R. Coffill, W. V. Zhang, Y. S. Tan, J. Hundt *et al.*, 2016 Functional characterization of p53 pathway components in the ancient metazoan *Trichoplax adhaerens*. *Sci Rep* **6**: 33972.
- Smith, C. A., S. E. Dho, J. Donaldson, U. Tepass and C. J. McGlade, 2004 The cell fate determinant numb interacts with EHD/Rme-1 family proteins and has a role in endocytic recycling. *Mol Biol Cell* **15**: 3698-3708.
- Song, Y., and B. Lu, 2012 Interaction of Notch signaling modulator Numb with alpha-Adaptin regulates endocytosis of Notch pathway components and cell fate determination of neural stem cells. *J Biol Chem* **287**: 17716-17728.
- Tan, B. X., H. P. Liew, J. S. Chua, F. J. Ghadessy, Y. S. Tan *et al.*, 2017 Anatomy of Mdm2 and Mdm4 in evolution. *J Mol Cell Biol* **9**: 3-15.
- Thompson, J. D., T. J. Gibson and D. G. Higgins, 2002 Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics* **Chapter 2**: Unit 2 3.
- Tosoni, D., S. Pambianco, B. Ekalle Soppo, S. Zecchini, G. Bertalot *et al.*, 2017 Pre-clinical validation of a selective anti-cancer stem cell therapy for Numb-deficient human breast cancers. *EMBO Mol Med* **9**: 655-671.

- Tosoni, D., S. Zecchini, M. Coazzoli, I. Colaluca, G. Mazzarol *et al.*, 2015 The Numb/p53 circuitry couples replicative self-renewal and tumor suppression in mammary epithelial cells. *J Cell Biol* **211**: 845-862.
- Uhlik, M. T., B. Temple, S. Bencharit, A. J. Kimple, D. P. Siderovski *et al.*, 2005 Structural and evolutionary division of phosphotyrosine binding (PTB) domains. *J Mol Biol* **345**: 1-20.
- Waterhouse, A. M., J. B. Procter, D. M. Martin, M. Clamp and G. J. Barton, 2009 Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**: 1189-1191.
- Wei, R., T. Kaneko, X. Liu, H. Liu, L. Li *et al.*, 2017 Interactome mapping uncovers a general role for Numb in protein kinase regulation. *Mol Cell Proteomics*.
- Zeng, X., L. Chen, C. A. Jost, R. Maya, D. Keller *et al.*, 1999 MDM2 suppresses p73 function without promoting p73 degradation. *Mol Cell Biol* **19**: 3257-3266.
- Zhong, W., M. M. Jiang, G. Weinmaster, L. Y. Jan and Y. N. Jan, 1997 Differential expression of mammalian Numb, Numbl like and Notch1 suggests distinct roles during mouse cortical neurogenesis. *Development* **124**: 1887-1897.
- Zobel, M., A. Disanza, F. Senic-Matuglia, M. Franco, I. N. Colaluca *et al.*, 2018 A NUMB-EFA6B-ARF6 recycling route controls apically restricted cell protrusions and mesenchymal motility. *J Cell Biol* **217**: 3161-3182.

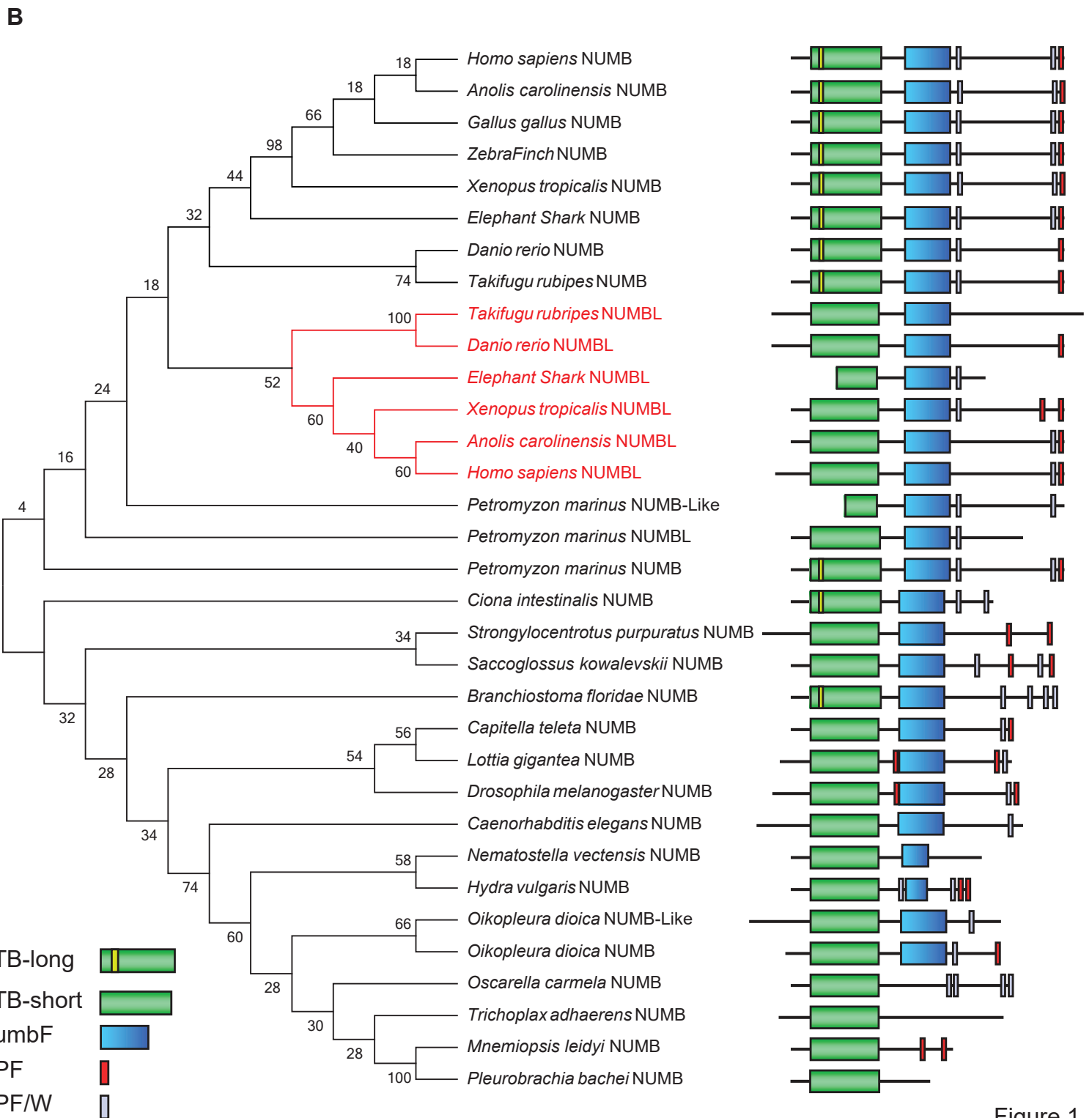
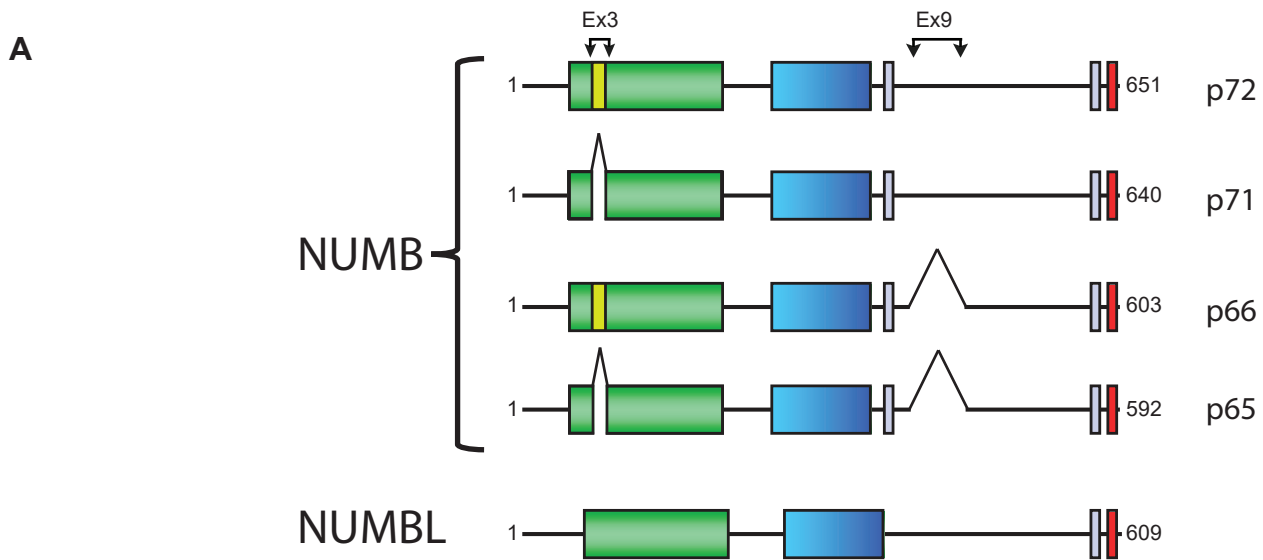


Figure 1

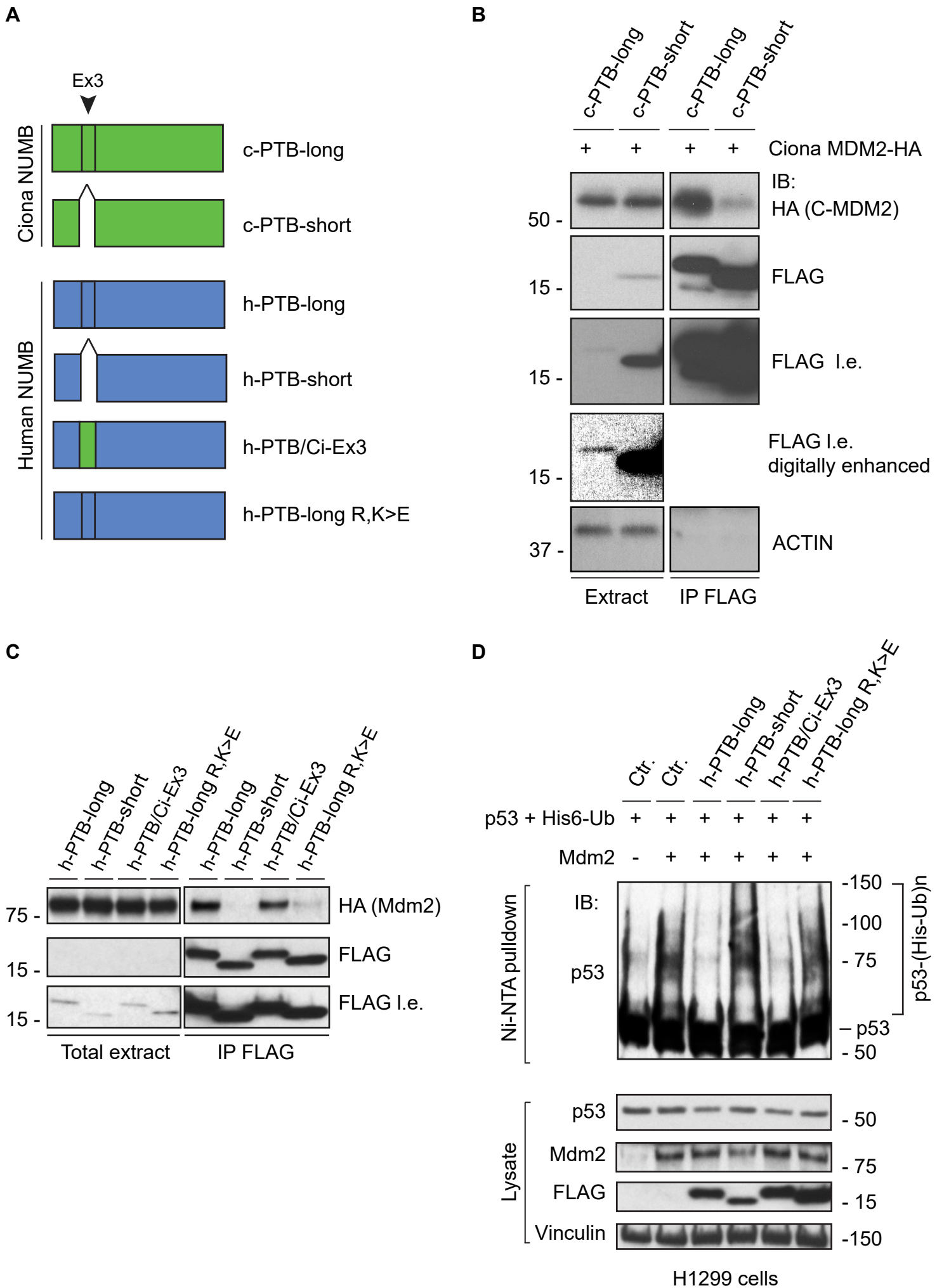


Figure 3