



UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE BIOLOGIA

JORDY ALEXANDER LARCO LASSO

ESTUDOS DAS RELAÇÕES EVOLUTIVAS E FUNCIONAIS DA PROTEINA VIRAL
HBX COM AS SUBUNIDADES DO COMPLEXO RAGULATOR

STUDIES OF THE EVOLUTIONARY AND FUNCTIONAL RELATIONSHIP OF VIRAL
PROTEIN HBX WITH RAGULATOR SUBUNITS

Campinas

2021

JORDY ALEXANDER LARCO LASSO

ESTUDOS DAS RELAÇÕES EVOLUTIVAS E FUNCIONAIS DA
PROTEINA VIRAL HBX COM AS SUBUNIDADES DO COMPLEXO
RAGULATOR

STUDIES OF THE EVOLUTIONARY AND FUNCTIONAL
RELATIONSHIP OF VIRAL PROTEIN HBX WITH RAGULATOR
SUBUNITS

*Dissertação apresentada ao Instituto de
Biologia da Universidade Estadual de
Campinas como parte dos requisitos
exigidos para a obtenção do título de
Mestre em genética e biologia molecular,
na Área de genética animal e evolução.*

*Dissertation presented to the Institute of
Biology of the University of Campinas in
partial fulfillment of the requirements for
the degree of Master in genetics and
molecular biology in the area of animal
genetics and evolution.*

Orientadora: Dra. Juliana Helena Costa Smetana

ESTE ARQUIVO DIGITAL CORRESPONDE À
VERSÃO FINAL DA DISSERTAÇÃO
DEFENDIDA PELO ALUNO JORDY
ALEXANDER LARCO LASSO E ORIENTADA
POR JULIANA HELENA COSTA SMETANA

Campinas
2021

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Biologia
Mara Janaina de Oliveira - CRB 8/6972

L321s Larco Lasso, Jordy Alexander, 1993-
Studies of the evolutionary and functional relationship of vital protein HBx with Ragulator subunits / Jordy Alexander Larco Lasso. – Campinas, SP : [s.n.], 2021.

Orientador: Juliana Helena Costa Smetana.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Biologia.

1. Complexo Ragulator. 2. Vírus da hepatite B. 3. Serina-treonina quinases TOR. 4. Proteínas virais. I. Smetana, Juliana Helena Costa. II. Universidade Estadual de Campinas. Instituto de Biologia. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Estudos das relações evolutivas e funcionais da proteína viral HBx com as subunidades do complexo Ragulator

Palavras-chave em inglês:

Ragulator complex
Hepatitis B virus
TOR serine threonine kinases
Viral Proteins

Área de concentração: Genética Animal e Evolução

Titulação: Mestre em Genética e Biologia Molecular

Banca examinadora:

Juliana Helena Costa Smetana [Orientador]
Gustavo Costa Bressan
Jorge Mauricio Costa Mondego

Data de defesa: 29-01-2021

Programa de Pós-Graduação: Genética e Biologia Molecular

Identificação e informações acadêmicas do(a) aluno(a)

- ORCID do autor: <https://orcid.org/0000-0002-7890-0987>

- Currículo Lattes do autor: <http://lattes.cnpq.br/3097630379949173>

Campinas, 29 de janeiro de 2021.

COMISSÃO EXAMINADORA

Dra. Juliana Helena Costa Smetana

Dr. Gustavo Costa Bressan

Dr. Jorge Mauricio Costa Mondego

Os membros da Comissão Examinadora acima assinaram a Ata de Defesa, que se encontra no processo de vida acadêmica do aluno.

A Ata da defesa com as respectivas assinaturas dos membros encontra-se no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria do Programa de Pós-Graduação em Genética e Biologia Molecular, do Instituto de Biologia da Universidade Estadual de Campinas

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior- Brasil (CAPES)- Código de Financiamento 001.

A minha orientadora Dra. Juliana Helena Costa Smetana por acreditar em mim desde o primeiro momento, pela oportunidade de crescimento e aprendizado única. Obrigado por sempre estar disponível para discutir resultados, e por sua imensa paciência que me ensinou a ser um melhor profissional e sobretudo suas opiniões e sugestões que preencheram este projeto de qualidade científica e sobretudo me levou pelo caminho para me converter em um melhor pesquisador.

Ao CNPEM (Centro Nacional de Pesquisa em Energia e Materiais), especialmente ao LNBio (Laboratório Nacional de Biociências) pela infraestrutura.

Ao Instituto de Biologia, especialmente ao Programa de Pós-Graduação em Genética e Biologia Molecular e a todos os docentes que ministraram as disciplinas que cursei durante este mestrado. A todos os membros da banca de qualificação e defesa, pela leitura atenciosa deste trabalho e valiosas contribuições.

A todos os amigos do laboratório, em especial aos que me ajudaram: Mariana Piccoli, Luiza Leme, Ana Iacia, Paloma Messias e Mayara Roquette. Mas também a todos os que passaram pelo grupo durante os anos de desenvolvimento deste trabalho e que com certeza contribuíram para sua realização. Um agradecimento especial para a Mariana pela colaboração neste projeto.

Aos pesquisadores colaboradores do projeto: a doutora Angela Saito pelas discussões sobre mitofagia e por ceder reagentes, ao doutor Matheus de Castro Fonseca pela ajuda na execução e interpretação de experimentos de microscopia confocal, e ao Professor Doutor Luiz Eduardo Del-Bem e seu aluno de doutorado Breno Lisboa pela ajuda com

as análises evolutivas e de bioinformática. Sem seus conhecimentos e experiência em distintas áreas, este projeto multifacetado não teria sido realizado.

Agradeço imensamente aos meus pais Tamara e Jaime, que sempre me apoiaram em todas minhas loucuras, mesmo você não concordando com o caminho escolhido sempre recebi apoio incondicional e suporte emocional e financeiro para a realização do meu sonho. Los quiero y los extraño muito. A meu irmão que quase nunca me liga, mas sempre pergunta por mim.

Resumo

O complexo mTORC1 integra sinais provenientes de nutrientes, energia e fatores de crescimento e coordena o crescimento e o metabolismo celular. A proteína HBXIP foi inicialmente identificada como um parceiro de interação para a proteína HBx do vírus da hepatite B que participa do processo de infecção viral e, mais recentemente, essa proteína também foi identificada como uma subunidade essencial do complexo Ragulator. Este complexo pentamérico, formado por p18, MP1, p14, C7orf59 e HBXIP, ancora as Rag GTPases na superfície do lisossomo e funciona como GEF (fator de troca de nucleotídeo) dessas GTPases, promovendo um estado ativado que recruta mTORC1 para a superfície do lisossomo e promove sua ativação. HBXIP é uma pequena proteína com um enovelamento do tipo Roadblock que possui duas isoformas descritas, HBXIP longa (18 kDa) e HBXIP curta (11 kDa). Apesar de sua importância como subunidade do Ragulator e parceira de interação de HBx, pouco se sabe sobre as diferenças entre essas isoformas. Neste projeto, investigamos a relação evolutiva do complexo Ragulator com a proteína viral HBx e a relação funcional entre eles. Usando ferramentas de bioinformática, verificamos sinais de seleção positiva na sequência da isoforma longa de HBXIP, o que caracteriza uma assinatura de conflito genético em relações evolutivas de parasita e hospedeiro. Análises de bioinformática também revelaram a existência de duas isoformas adicionais de HBXIP que se assemelham à isoforma curta exceto por uma pequena região N-terminal, e o padrão de expressão das quatro isoformas de HBXIP em diferentes tecidos humanos foi analisado em um banco de dados de RNA-seq revelando padrões tecido-específicos. Em células transfectadas, HBx colocaliza com as subunidades do Ragulator p14, C7orf59 e isoforma longa de HBXIP, mas não com p18, em estruturas que também são positivas para o corante mitocondrial Mitotracker Deep Red. O padrão de localização subcelular das subunidades do Ragulator é afetado pela expressão de HBx, que induz um padrão perinuclear indicativo do processo de mitofagia. Além disso, HBx interage com p14, C7orf59 e isoforma longa de HBXIP em células transfectadas. MP1 e a isoforma curta de HBXIP não foram testadas nesses experimentos. Esses resultados sugerem que HBx é capaz de interagir com todo o complexo Ragulator, possivelmente excluindo p18, em um complexo que provavelmente está envolvido no processo de mitofagia, e que a isoforma longa de HBXIP é uma aquisição evolutiva recente relacionada ao processo de infecção viral.

Abstract

The mTORC1 complex integrates signals arising from nutrients, energy, and growth factors and coordinates cell growth and metabolism. The HBXIP protein was initially identified as an interaction partner for the hepatitis B virus HBx protein that participates in the viral infection process, and more recently this protein has also been identified as an essential subunit of the Ragulator complex. This pentameric complex, formed by p18, MP1, p14, C7orf59, and HBXIP, anchors the Rag GTPases on the surface of the lysosome and functions as GEF (nucleotide exchange factor) of these GTPases, promoting an activated state that recruits mTORC1 to the surface of the lysosome and promotes its activation. HBXIP is a small protein with a Roadblock folding type that has two isoforms, HBXIP long (18 kDa) and HBXIP short (11 kDa), originating from alternative transcription initiation sites. Despite its importance as a Ragulator subunit and HBx interaction partner, little is known about the differences between these isoforms. In this project, we investigated the evolutionary relationship of the Ragulator complex with the viral protein HBx and the functional relationship between them. Using bioinformatics tools, we found signs of positive selection in the long isoform of HBXIP, which characterizes a signature of genetic conflict in evolutionary parasite-host relationships. Bioinformatics analyzes also revealed the existence of two additional HBXIP isoforms that resemble the short isoform except for a small N-terminal region, and the expression pattern of the four HBXIP isoforms in different human tissues was analyzed in an RNA-seq database revealing tissue-specific patterns. In transfected cells, HBx colocalizes with the subunits of Ragulator p14, C7orf59 and long isoform of HBXIP, but not with p18, in structures that are also positive for the mitochondrial dye Mitotracker Deep Red. The subcellular localization pattern of the Ragulator subunits is affected by the expression of HBx, which induces a perinuclear pattern indicative of the mitophagy process. In addition, HBx interacts with p14, C7orf59 and long isoform of HBXIP in transfected cells. MP1 and the short isoform of HBXIP were not tested in these experiments. These results suggest that HBx interacts with the entire Ragulator complex, possibly excluding p18, in a complex that is probably involved in the mitophagy process, and that the long isoform of HBXIP is a recent evolutionary acquisition related to viral infection.

FIGURE LIST

Figure 1: Crystallographic structure of pentameric Ragulator complex	17
Figure 2: Structure of the Rag-Ragulator complex	17
Figure 3: mTORC1 anchorage to lysosome	19
Figure 4: HBx motifs and domains graphic representation.	22
Figure 5: Graphic representation of Ubiquitin dependent or independent pathway	27
Figure 6: Graphic representation of LAMTOR5 multiple sequence alignment	44
Figure 7: Graphic representation of the median protein length of HBXIP	45
Figure 8: HBXIP protein phylogenetic tree	46
Figure 9: Co-occurrence of the LAMTOR5 long isoform and susceptibility to infection with Orthohepadnaviruses in mammalian species	47
Figure 10: BUSTED Result for selective pressure in <i>LAMTOR5</i>	49
Figure 11: BUSTED result for selective pressure for Lamtor5 protein in primates	51
Figure 12: BUSTED(S) Result for selective pressure for <i>LAMTOR5</i> gene in rodents	52
Figure 13: Likelihood ratio test statistic for episodic diversification MEME graphic results of HBXIP protein of primates sequences	54
Figure 14: Graphic representation of multiple sequence alignment X protein from Orthohepadnaviruses	55
Figure 15: X protein from Orthohepadnavirus family phylogenetic tree	56
Figure 16: Likelihood ratio test statistic for episodic diversification MEME graphic results of X protein of Orthohepadnaviruses family	57
Figure 17: Likelihood ratio test statistic for episodic diversification MEME graphic results of BCL-2 protein	59
Figure 18: Graphic cartoon representation of HBx BH3-like motif with BCL2 protein	61
Figure 19: BH3 like motif of X protein from Orthohepadnavirus family alignment	62
Figure 20: Graphic representation of HBXIP isoforms alignment	63
Figure 21: Comparison of the two short HBXIP isoforms: HBXIP short A and HBXIP 3B.	65
Figure 22: heatmap representation from HBXIP isoforms transcripts quantification	66
Figure 23: Boxplot graphic representation from HBXIP isoforms transcripts	67
Figure 24: Mitochondrial localization of HBx in HepG2 cells	69
Figure 25: Immunolocalization of transiently expressed Ragulator subunits in U2OS	70

cells

Figure 26: Ragulator subunits colocalization with HA- GFP-HBx WT in U2OS transfected cells	71
Figure 27: Ragulator subunits localization with HA- GFP-HBx NESM in U2OS transfected cells	72
Figure 28: HBXIP Long has a negative effect on GFP-HA-HBx expression	73
Figure 29: Colocalization of HBx with Ragulator subunits in HepG2 cells	74
Figure 30: Co-immunoprecipitation of HBx with Ragulator subunits in transiently transfected HEK293-T cells	76
Figure 31: Ragulator subunits co-transfected with Parkin GFP construction in Huh-7 Cells with CCCP treatment	78
Figure 32: Ragulator subunits co-transfected with Parkin GFP construction in Huh-7 cells without CCCP treatment	79
Figure 33: Ragulator subunits co-transfected with HBx-GFP construction in Huh-7 cells with CCCP treatment	79
Figure 35: Ragulator subunits co-transfected with HBx-GFP construction in Huh-7 cells without CCCP treatment	80

TABLE LIST

Table 1: Plasmid constructions and Operetta channels	38
Table 2: UniProt proteins codes	40
Table 3: Number of samples or FASTA files per tissue RNA Atlas	41
Table 4: Transcript codes of LAMTOR5 isoforms from ESEMBLE	43
Table 5: Summary of BUSTED analyses for Ragulator proteins	50
Table 6: Summary of BUSTED analyses (after update) for Lamtor5, DDB1, Bcl2-L, Ragulator subunits and X	53
Table 7: A summary of MEME results of HBXIP Long	54
Table 8: A summary of MEME results of X protein	57
Table 9: A summary of MEME results BCL-2	60
Table 10: HBXIP isoforms UniProt codes and amino acids length	63
Table 11: Deeploc subcellular distribution prediction for human HBXIP isoforms	64

Summary

1.	Introduction	14
1.1.	The PI(3)K / Akt / mTOR pathway	14
1.1.1.	Amino acids and the mTORC1 complex.	15
1.1.2.	Ragulator complex	16
1.2.	The hepatitis B virus (HBV)	19
1.3.	The HBx protein	20
1.3.1.	Multifunctional HBx	21
1.3.2.	Transcriptional regulation of HBx	22
1.3.3.	HBx and apoptosis	23
1.3.4.	HBx and Calcium-Signaling Pathways	23
1.3.5.	HBx and the Cell Cycle	24
1.4.	HBX and mitophagy	24
1.4.1.	Mitophagy a kind of Autophagy	24
1.4.1.1.	Mechanism and regulation of Mitophagy	25
1.4.2.	HBx and mitophagy induction	28
1.5.	HBXIP and HBx	29
1.6.	Evolutionary Concepts	30
1.7.	Methods of Evolutionary Analysis	33
2.	Objectives:	34
2.1.	Main objectives:	34
2.2.	Specific objectives:	34
3.	Materials and methods	35
3.1.	Plasmids	35
3.2.	Cell culture, transfection, and immunoprecipitation	35
3.2.1.	PEI transfection	35
3.2.2.	Transfection of HEK293-T cells in 96-well plate	35
3.2.3.	Transfection of HepG2 cells in 24 well plates with lipofectamine 2000	36
3.5.	Transfection of HeLa and U2OS cells in 24 well plates with lipofectamine 3000	36
3.2.5.	Transfection of Huh-7 in 96-well plates	36
3.2.6.	Immunocytochemistry	37
3.2.7.	Immunoprecipitation	37
3.2.8.	Mitophagy live cell experiments	38
3.3.	Evolutionary analysis and bioinformatic analyses	39
3.3.1.	Preliminary analyses	39
3.3.2.	Evolutionary analyses	39
3.3.3.	Subcellular localization prediction	40
3.3.4.	RNA-seq Analyses	41
3.3.5.	Structural model of HBXIP_3	42

4.	Results	44
4.1.	HBXIP Molecular Evolution Analyses	44
4.2.	HBx molecular evolution analyses	55
4.3.	HBXIP isoforms bioinformatic analysis	63
4.4.	Colocalization of HBx and Ragulator subunits in transfected cells	69
4.5.	Results of interaction experiments	75
4.6.	Live cell mitophagy experiments	78
5.	Discussion	81
5.1.	Phylogenetic and evolutionary analysis of HBXIP and its relationship with HBx	81
5.2.	Protein-protein interaction assays between HBx and the subunits of the Ragulator complex	83
5.3.	Role of HBx viral protein and Ragulator subunits in mitophagy	85
6.	Bibliography	87
7.	Anexos	99
Anexo 1	HBXIP protein phylogenetic tree	99
Anexo 2	Declaração de bioética e/ou biossegurança	100
Anexo 3	Documento do comite de bioética e/ou biossegurança	101
Anexo 4	Termo de direitos autorais	104

1. Introduction

1.1. The PI(3)K / Akt / mTOR pathway

The mTOR (mechanistic Target of Rapamycin) protein is a kinase considered as a critical control point in cell growth and receives stimulating signals from Ras and PI(3)K (phosphatidylinositol-3-OH kinase). Ras and PI(3)K are stimulated by growth factors and nutrients such as amino acids and glucose. The importance of Ras and PI(3)K is evident from the fact that mutations involving proteins in these pathways are present in many types of human cancer, approximately 30% of tumors, second only to p53 (Shaw & Cantley, 2006).

The Ras-PI(3)K-mTOR signaling pathway is part of a biochemical network, in which environmental stimuli control cell growth. When these proteins are altered, for example in tumors, cells grow relatively independent of environmental stimuli, proliferate, and may become insensitive to pro-apoptosis stimuli because it is a survival pathway. The Ras and PI(3)K proteins converge to activate the mTOR kinase in order to activate cell growth. The tumor suppressors TSC1 (hamartin), TSC2 (tuberin), and LKB11 (serine/threonine-protein kinase 11) attenuate the effects of mTOR signaling (Inoki et al., 2003).

For more than three decades, the PI(3)K/AKT pathway has been studied in the activation of the protein kinase mTOR. The tumor suppressor protein TSC2 belonging to the AKT pathway is a critical target for blocking the activity of the mTOR protein kinase. TSC2 has been identified as a gene for the human hamartoma/tuberoma tumor syndrome, and as a gene that affects cell size in *Drosophila* flies. TSC2 downregulates mTOR signaling due to its ability to act as GAP (GTPase Activating Factor) for a Ras-like GTPase, Rheb, and AKT activation also helps in this inhibition (Hay & Sonenberg, 2004).

The activation of the PI(3)K-AKT pathway plays an important role in maintaining cellular metabolism under conditions of limited growth factors. The AKT pathway stimulates cell growth and ATP production by regulating the activity and expression of key glycolytic enzymes, as well as the amount of glucose and other nutrients. New studies have shown that AKT and mTOR regulate the expression of HIF-1 α (hypoxia-inducible factor-1 α), HIF-

1 α increases the expression of GLUT-1 and glycolytic enzymes (Wolfson & Sabatini, 2017).

In mammals, there are two complexes of the mTOR kinase protein, both have a regulatory subunit called LST8 and at least a third substrate-specific subunit. The third specific subunit is Raptor in the mTORC1 complex and Rictor (rapamycin-insensitive companion of mTOR) in the mTORC2 complex. These two complexes have different substrates and perform different functions in the cell. mTORC1 is strongly inhibited by rapamycin at low concentrations (of the nanomolar order), while mTORC2 is considered insensitive to rapamycin but can respond to high concentrations or prolonged rapamycin treatments. Most of the work initially done on mammalian cells is focused on the sensitivity of the mTORC1 complex to treatment with rapamycin, but recently catalytic inhibitors have been identified that also affect mTORC1 and mTORC2, for example, Torin-1 and PP242.

The mTORC1 complex integrates nutrients and growth factors derived from signals for the regulation of cell growth, the cell process that accumulates mass and increases cell size. The mTORC2 complex modulates the phosphorylation of PKC (protein kinase C) and the actin cytoskeleton (Robert A & David M, 2017).

1.1.1. Amino acids and the mTORC1 complex

The mTORC1 complex integrates signals that arise from nutrients, energy, and growth factors. Recent discoveries have placed the lysosome, a key organelle in cellular catabolism, as the center of mTORC1 regulation by amino acids. The multiprotein complex that includes the Rag GTPases, the Ragulator complex, and the v-ATPase form a piece of machinery sensitive to amino acids on the lysosomal surface that affects the decision between anabolism and catabolism at multiple levels (Wolfson & Sabatini, 2017). The involvement of a catabolic organelle in signs of cell growth plays an important role in understanding pathologies associated with mTORC1.

The activation of mTORC1 in the presence of amino acids is mediated by Rag GTPases. These small GTPases belong to the Ras superfamily and are rather unusual because they exist as heterodimers while other GTPases are monomeric. In these heterodimers, RagA or RagB interact with RagC or RagD, which are structurally related, allowing four

possible combinations. The availability of amino acids indirectly regulates the activation state of Rags, causing them to change their conformation, finally activating mTORC1 through physical interaction. The key to mTORC1-dependent amino acid activation is the conversion of RagA or RagB from an inactive state linked to GDP to an active state linked to GTP, while its dimerization partner (RagC or RagD) responds in the opposite way. In their state activated by amino acids, Rag GTPases interact with Raptor, its effector, to bring the mTORC1 complex to the surface of the lysosome where it can be activated by Rheb. As in other GTPases, this conversion depends on the activity of a nucleotide exchange factor or GEF (Guanine nucleotide Exchange Factor), a role which was later found to be accomplished by the Ragulator complex (Bar-Peled et al., 2012).

1.1.2. Structure and function of the Rag-Ragulator complex

Unlike many GTPases that depend on a lipid fraction for their subcellular location, Rags use the Ragulator complex to stay attached to the lysosomal surface. This complex was initially identified as a trimer formed by the proteins p18, p14 and MP1, which are encoded by the genes *LAMTOR1*, *LAMTOR2* and *LAMTOR3* respectively. In cells where these proteins were silenced, Rags and mTORC1 were unable to locate on the surface of the lysosome, and consequently the mTORC1 pathway was inactive (Efeyan et al., 2012).

Later, two additional proteins were identified as integral parts of the Ragulator complex: C7orf59 and HBXIP, encoded by the genes *LAMTOR4*, and *LAMTOR5*, respectively (Figure 1). Both proteins are necessary for the lysosomal location of the Rags to activate mTORC1. Surprisingly, the pentameric Ragulator has GEF activity, but none of its isolated subunits or the trimeric Ragulator complex has GEF activity for Rag A and Rag B. This finding defined the role of the pentameric Ragulator as both a scaffold and an activator of Rag GTPases in the context of amino acid signaling (Bar-Peled et al., 2012), however, the elucidation of the mechanism involved in this regulation required structural insight.

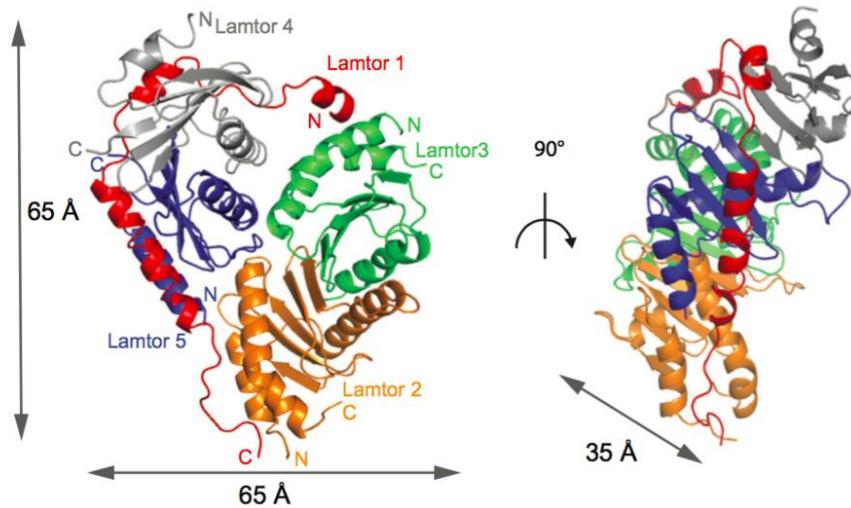


Figure 1: Crystallographic structure of pentameric Ragulator complex (Ming-Yuan Su et al., 2017).

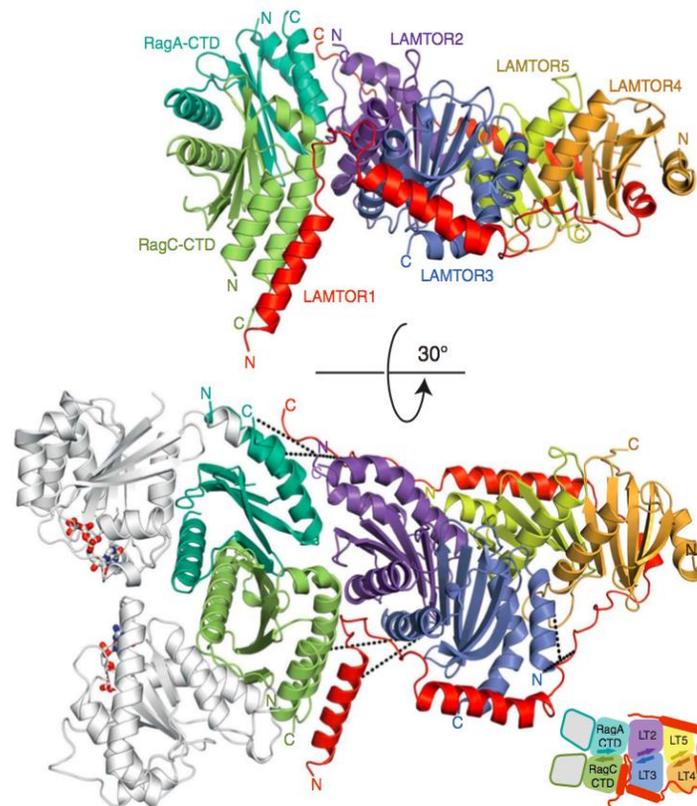
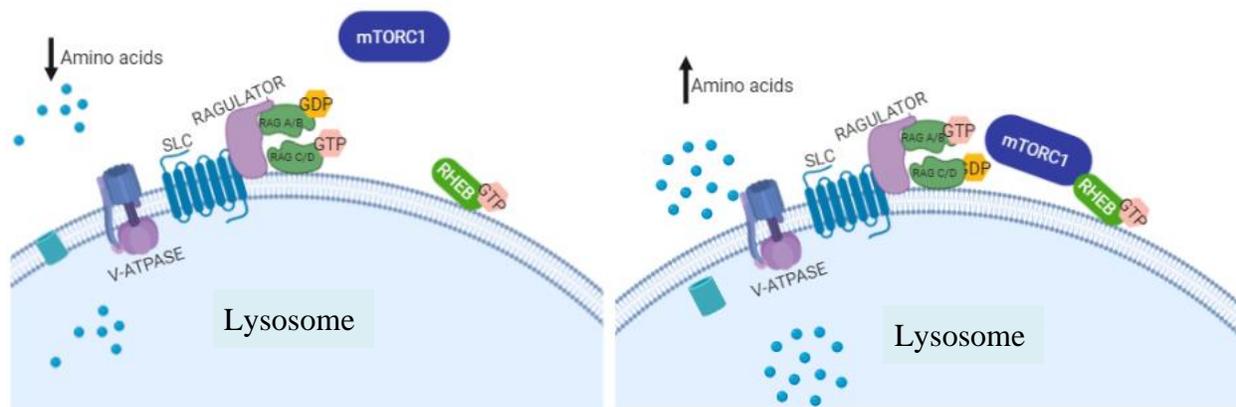


Figure 2: Structure of the Rag-Ragulator complex (De Araujo et al., 2017). The colored regions are present in the crystal structure while the white regions (N-terminal GTPase domains of Rags A and C) were modelled. The diagram at the bottom of the figure indicates the positions of the subunits in the complex.

Several groups were involved in elucidating the structure of pentameric Ragulator and Rag-Ragulator complexes, which were published five years after the identification of this complex (De Araujo et al., 2017; Ming-Yuan Su et al., 2017; Yonehara et al., 2017). The crystal structures revealed that Ragulator is composed by two Roadblock heterodimers (HBXIP-C7orf59 dimer and MP1-p14 dimer) wrapped by an unstructured p18 protein. It was believed that the formation of Rag-Ragulator complex is mediated by p18 subunit, p18 may be a prerequisite for mTORC1 amino acid regulation (Yonehara et al., 2017), but Our group contributed to understanding how the pentameric Ragulator is assembled by solving the structure of the HBXIP-c7orf59 dimer, showing that C7orf59 subunit possesses an unstructured N-terminal region which is necessary for its interaction with p18 and presumably nucleates the assembly of the pentameric Ragulator (Rasheed et al., 2019).

The N-terminus of p18 subunit can anchor to lysosome surface because it possesses myristoylation and palmitoylation sites. The second guanine in p18 protein was identified like a potential site for myristoylation, others sites identified as potential palmitoylation are the next two cysteines 3 that are the third and fourth amino acids in the sequences. Another important site in p18 is the sequence EERKLL that has the potential to have a lysosomal/vacuolar localization signal (Ratnayake et al., 2018). When Ragulator complex is anchored to the lysosome surface, it forms the Ragulator-Rag GTPases complex, with the interactions of p18 and Roadblock heterodimer MP1-p14 with Rag A/C (Yonehara et al., 2017).

Besides being a subunit of the Ragulator complex, the HBXIP protein has also been described as an anti-apoptotic factor and a cofactor of the Survivin protein, promoting the suppression of pro-caspase-9. There are two isoforms of HBXIP protein, a short isoform composed of 91 amino acid residues and an approximate molecular weight of 11 kDa (Fei et al., 2017), and a long isoform composed of 173 amino acid residues and an approximate molecular weight of 18 kDa (Wen et al., 2008).



Created in BioRender.com 

Figure 3: mTORC1 anchorage to lysosome mediated by Ragulator complex by the response to amino acids. Amino acids stimulated v-ATPase that is necessary for Ragulator GEF activity. Ragulator GEF activity promotes conversion of RAG A/B GDP complex to RAG A/B GTP complex, also promotes the contrary effect in RAG C/D. Once RAGs activation is complete mTORC1 is recruited to lysosome with an unknown mechanism. Created with BioRender.com

1.2. The hepatitis B virus (HBV)

The Hepatitis B virus (HBV) is a worldwide health problem, which in 2015 resulted in 887,000 deaths, most of them due to secondary complications including cirrhosis and hepatocellular carcinoma (HCC). Chronic infections of HBV are associated with an increased risk of developing HCC (Beck & Nassal, 2007).

HBV belongs to the hepadnaviral family. Hepadnaviral genome is a unique copy of a circular double stranded DNA around 3.0 – 3.5 kilobases which replicates through reverse transcription of an RNA intermediate, the pregenomic RNA (pgRNA). The transcription of the HBV genome is controlled by four transcription promoters and two enhancers (Beck & Nassal, 2007).

According to the international committee on taxonomy of viruses, Hepadnaviridae is a family composed by two genera, Orthohepadnaviruses, which infect mammals, and Avihepadnavirus, which infect birds. Orthohepadnaviruses include, in addition to HBV,

Ground squirrel hepatitis virus, Woodchuck hepatitis virus, Woolly monkey hepatitis B virus and the four recently discovered bat virus species. The most accepted theory on how these two groups evolved is that the ancestral host were the birds and later the barrier was crossed to mammals (Drexler et al., 2013; Kew, 2011; Lauber et al., 2017; Suh et al., 2013).

The orthohepadnaviral family genome contains seven ORFs (open reading frames) which express structural and non-structural proteins, one of them (ORF X) encodes HBx, a small multifunctional protein involved in carcinogenesis. The gene encoding HBx, or protein X, is an evolutionary acquisition of the genus Orthohepadnavirus. This genus infects mammals exclusively, while other Hepadnaviridae infect other vertebrate species but do not express the HBx protein and do not induce liver tumors, so the presence of ORF X in Orthohepadnaviruses could be associated to carcinogenesis events with the chronic infection in mammals (Netter et al., 2001).

1.3. The HBx protein

The HBx protein has an important role in the pathogenesis of hepatocellular carcinoma induced by HBV. It can transactivate different viral and cellular promoters. The gene encoding this protein is frequently integrated into the host DNA and expressed in the infected cell, promoting the progression of the cell cycle, inactivating negative growth regulators and inhibiting the expression of the tumor suppressor p53. Due to the lack of a DNA binding motif or cellular localization signals, HBx is believed to function through protein-protein interactions. These interactions are necessary for the cellular environment to be suitable for viral replication (Sir et al., 2010).

Because the HBx protein is necessary to ensure a productive HBV infection, the proteins that interact with it have been extensively studied. However, despite the vast number of publications, much remains to be understood about the mechanisms of these interactions in the life cycle of HBV, in particular because of the scarcity of structural data. Many techniques were applied to make the study of protein-protein interactions with HBx possible, which revealed that HBx interacts with many proteins including members of the general transcription machinery, the components of proteasomes, probably with DNA

repair proteins and with members of the CREB/ATF family, where it influences its specificity and its affinity for DNA. In one of these studies using a two-hybrid screening method in yeast, a protein that interacts with HBx was isolated and was called HBXIP (HBx-interacting protein). This protein is the same that was later identified as the LAMTOR5 subunit of the pentameric Ragulator (Bar-Peled et al., 2012; Wen et al., 2008).

1.3.1. Multifunctional HBx and its interactions

HBx is capable of interacting with many proteins because of its intrinsically unstructured nature. Multiple functions have been assigned in in vitro experimental cell system, like: transcription, viral replication, apoptosis, DNA repair, proteolysis, signaling pathways, cell cycle, regulation and tumor suppressor pathways (Slagle & Bouchard, 2016).

HBx is the only regulatory protein in HBV and the most important in the generation of an appropriate environment for HBV replication. The generation of this environment is possible because of the different localization in subcellular compartments. This ability gives HBx the possibility to interact with many cellular signaling pathways and components of the transcription machinery (Lucifora et al., 2011).

Despite being a small protein, HBx has many sequence motifs and domains with known functions (**Figure 3**). Residues 1-50 comprise an inhibitory domain for HBx activities (Murakami et al., 1994), residues 58-119 are involved in signal transduction activities, residues 120-140 are involved in nuclear transactivation mechanism (NT) and finally the last 20 residues of HBx protein are necessary for its stability (Slagle & Bouchard, 2016). In the HBx amino acid sequence, we can find sites that are involved in subcellular localization, like residues 54-70, 75-88 and 109-131 which are essential for HBx mitochondrial localization (ML) (Fessler et al., 2008).

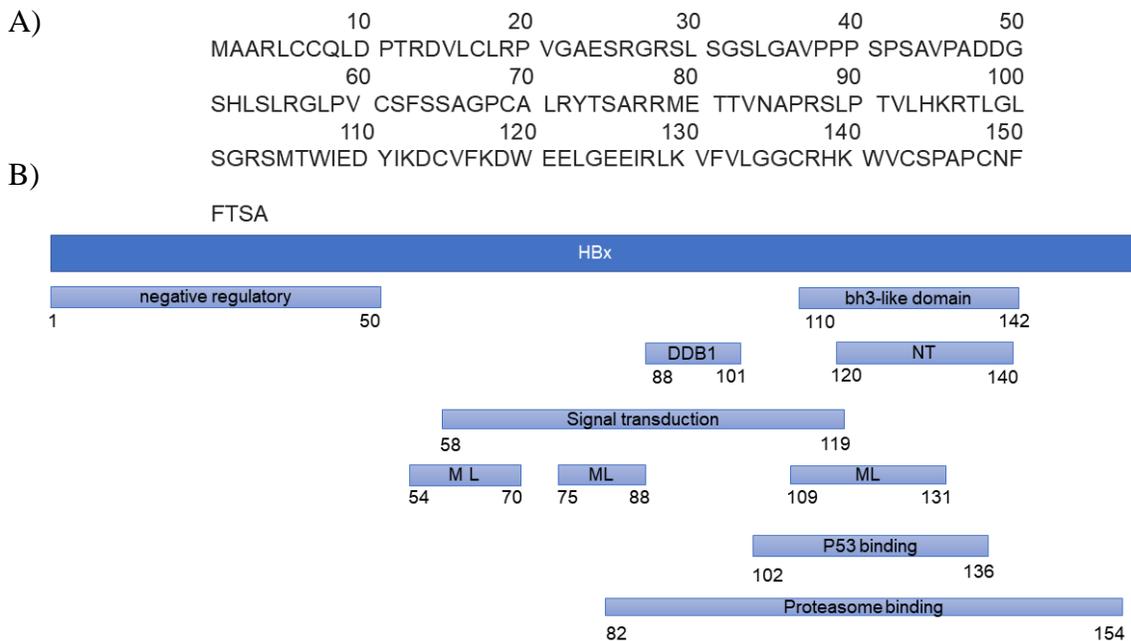


Figure 4: HBx motifs and domains graphic representation. A) HBx sequence from UniProtKB database (code Q156X2). B) Graphic representation of HBx motifs and domains published in different studies. This graphic helps in the understanding of the multiple functions and interactions of HBx. Negative regulatory domain is the only motif that does not share residues with another domain.

1.3.2. Transcriptional regulation by HBx

Hepatocytes are quiescent cells and this condition is not an ideal intracellular environment for DNA virus replication. Several viruses that infect quiescent cells rely on different proteins that transform the conditions of these cells to an appropriate environment for viral replication, several of them acting as transcriptional regulators like BALF2 protein which is a single stranded DNA binding protein from Epstein-Barr virus (Tsurumi et al., 1996). In the case of HBV, this protein is HBx, but HBx does not bind directly to DNA. Instead, HBx activates RNA polymerase I, II and III-dependent promoters which has been linked with some transcription factors and cellular signal transduction that regulates transcription (Aufiero & Schneider, 1990).

One of the most important interactions of HBx is with DDB1 (damaged DNA-binding protein 1). DDB1 is an adaptor protein for Cullin4- E3 ligase and it is possible with the

DDB1 Cullin adaptor factors (DCAFs). HBx shares 16 amino acid motif with DCAFs so it is possible that this interaction displays one or more DCAFS from the Cul4/DDB1 complex (Li et al., 2010). These interactions are so important that mutants of HBx that lose interaction with DDB1 are unable to function in the viral replication like the wild HBx type (Fessler, Michael B.; Rudel, Lawrence L.; Brown, 2008).

1.3.3. HBx and apoptosis

The most important function of HBx is to control the cell cycle and apoptosis in order to promote a better condition for HBV viral replication. Many studies have shown the control of HBx on cellular apoptosis. HBx is capable of blocking cell death by mediation of tumor necrosis factor (TNF), Fas, p53 or transforming growth factor beta (TGF-beta), and also can promote apoptosis with the same factor in other cell lines (Shirakata & Koike, 2003). Investigations of HBx effects in mitochondria reveal that HBx can cause mitochondrial aggregation and cytochrome C release, both processes are linked to apoptosis induction (S. Kim et al., 2007).

The crystal structure of alpha helix of BH3-like HBx motif interacting with Bcl-xL hydrophobic pocket was published in 2019 (Zhang et al., 2019). The structure showed that Trp120 and Leu123 of HBx are essential for its interaction with Bcl-xL. These discoveries opened a new chapter about HBx relation with apoptosis by providing a structural basis to understand how HBx targets Bcl-2 and Bcl-xL to elevate intracellular calcium and promote cell death (Geng et al., 2012).

1.3.4. HBx and Calcium-Signaling Pathways

HBx effects involve direct interactions between HBx and the proteins that induce signal transduction cascades for cellular modulation of calcium (Ca^{2+}) and reactive oxygen species (ROS) levels (Yang & Bouchard, 2012). These data were obtained by measured levels of Ca^{2+} and ROS in many liver cell lines (Bouchard et al., 2003). HBV replication in HepG2 cells and cultured primary rat hepatocytes are stimulated by HBx Ca^{2+} and ROS regulation. This Ca^{2+} regulation can be blocked by inhibition of the mitochondrial permeability transition pore (MPTP). The reported interaction of HBx with VDAC (voltage

dependent anion channel) suggests that HBx regulation of mitochondrial Ca^{2+} increases cytosolic Ca^{2+} levels (T. Gearhart & Bouchard, 2008).

1.3.5. HBx and the Cell Cycle

Many viruses need specific cell-cycle related conditions for their replication, such as a specific stage or stages of cell cycle, often requiring the action of a viral-encoded a protein that can stimulate the cell cycle progression. Cell cycle control by HBx is necessary for HBV viral replication in nondividing primary hepatocytes. Studies in cultured primary hepatocytes from rats and humans showed that these cells must exit G0 and enter G1 for replication of HBV, but then remain arrested in the G1 phase of the cell cycle. HBx can cause this by decreasing the expression of the cell cycle inhibitor proteins p15 and p16 which block entrance to G1, but it increased the levels of the cell cycle inhibitor proteins p21 and p27 (T. L. Gearhart & Bouchard, 2010).

1.4. HBx and mitophagy

1.4.1. Mithophagy, a kind of Autophagy

Autophagy is a broad term referring to a degradation system, which uses different pathways, to fusion and degrade cytoplasmatic components and organelles in the lysosome. New investigation points that autophagy is more than a system that just eliminates cytoplasmatic materials, it has recycling dynamics in order to produce building blocks and energy for homeostasis. There are three kinds or classes of autophagy: microautophagy, chaperone-mediated autophagy (CMA), and macroautophagy (Mizushima & Komatsu, 2011).

Microautophagy is a process in which the invaginations of the lysosomal membrane can engulf small components of the cytoplasm. During the process of microautophagy, the membrane dynamics will be quite similar to the late endosome endosomal sorting

complex required for transport (ESCRT)-dependent multivesicular body (MVB) formation. The endosomal lumen incorporates amounts of cytosolic proteins during MVB formation (Sahu et al., 2011).

The chaperone-mediated autophagy does not need membrane recognition. This process uses the translocation of heat shock cognate 70 (Hsc70) chaperone protein during its autophagy. This chaperone with its co-chaperone recognizes unfolded cytosolic proteins that contain the KFERQ-like motif. The substrate complex formed by Hsc 70 chaperone, co-chaperone, and the unfolded protein substrate is recognized by Lamp-2A transmembrane protein. The Lamp-2A transmembrane protein acts as a receptor on the lysosome and delivers the unfolded proteins to the lumen (Orenstein & Cuervo, 2010).

The third type of autophagy, macroautophagy is the major type and is conserved since yeasts. Macroautophagy uses an intermediated organelle called the autophagosome. The formation of the autophagosome occurs when an isolation membrane sequesters a small portion of the cytoplasm to form a double membranous structure that is the phagosome. Later the phagosome can fusion with lysosomes and finally its fusion with the lysosome becoming an autolysosome (Yu et al., 2010). Autophagy can be induced by starvation to provide essential nutrients to the cells or it can be specific to the degradation of an organelle. Mitochondria are removed by a special form of macroautophagy called mitophagy (Mizushima & Komatsu, 2011).

1.4.1.1. Mechanism and regulation of Mitophagy

Mitochondria are vital organelles in the life of eukaryotic cells because they are involved with critical cell functions like fatty acid oxidation, Krebs cycle, and oxidative phosphorylation, apoptosis, necrosis, autophagy, stress regulation, Ca^{2+} storage, and innate immunity. To cover all these functions, mitochondria frequently change their shape and regulate their networks by mitochondrial dynamics (fusion and fission). These various functions in which mitochondria are involved expose them to damages by intra and extramitochondrial events (Yoo & Jung, 2018).

To control all these damages that mitochondria suffer accomplishing their functions, mitochondria have their own quality control. Mitochondrial quality control involves biogenesis, fusion, and fission. Mitochondrial biogenesis is controlled by the delivery of NEMP and is induced by ROS, nitrogen oxide, carbon monoxide, and external signal. Mitochondrial fusion is controlled by MitoPLD to fusion the outer membrane and OPA1 that is anchored the inner membrane for the fusion of the inner membrane for mitochondria content exchange. In mitochondrial fission, Drp1 translocates to the outer membrane and binds multiple receptors (FIS1, MFF, MID49, and MID51) forms a ring structure around the fission site and produce a membrane scission (Ashrafi & Schwarz, 2013).

When the mitochondrial quality control or the repair systems fail in mitochondrial preservation, this organelle is targeted to mitophagy. The mitochondria can be degraded by vesicles that go directly to the lysosome (Lemasters, 2014). However, most of the studies are focused on the degradation by the autophagosome, especially on the importance of mitophagy interactors of LC3 protein. Mitophagy can be categorized through LC3 adapters in a ubiquitin-dependent pathway, a ubiquitin-independent pathway, and direct interaction with LC3 (Yoo & Jung, 2018).

The ubiquitination of mitochondrial substrates that are going to be recognized by the LC3 adapter is the first step in ubiquitin-dependent interaction. Two proteins have shown great importance in mitophagy: PINK1 (PTEN-induced putative kinase 1) and Parkin. In normal conditions, PINK1 is degraded in the inner membrane by the matrix processing peptidases (MPP), presenilin-associated rhomboid like (PARL), and the proteasome. Damages like mitochondrial membrane depolarization reduce the degradation of PINK1 and cause accumulation in the outer membrane through translocase of the outer membrane (TOM). Another factor that can cause PINK1 accumulation in the outer membrane is the increased pyruvate level (Park et al., 2015).

PINK1 accumulation causes its autophosphorylation and phosphorylates ubiquitin that gains interaction with Parkin. Later, Parkin is phosphorylated by PINK1, and

polyubiquitinates its substrate. The polyubiquitinated proteins are VDAC, MFN 1/2 and Miro1, these cause its degradation by the proteasome and at the same time induce mitochondrial fission and mitophagy. Finally, when the polyubiquitination is made the LC3 adaptors (p62, NPD52, OPTN, TAX1B1, and NBR1) interacts with the k63-linked polyubiquitinated substrates. When adaptors are linked to the mitochondria outer membrane they interact with LC3 through the LIR motif, initiating the formation of the isolation membrane (Nguyen et al., 2016).

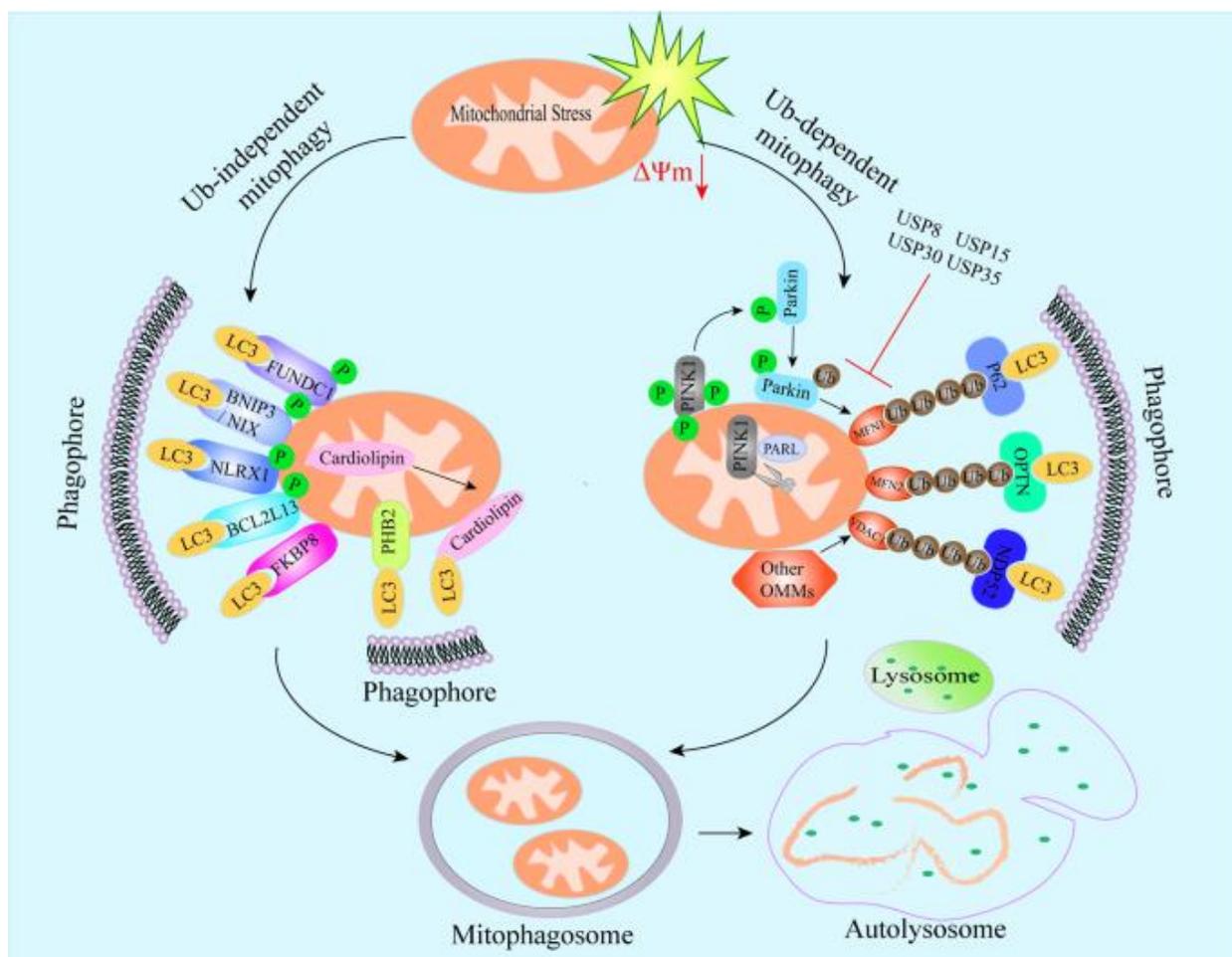


Figure 5: Graphic representation of mitophagy Ubiquitin-dependent or independent pathways. Following a mitophagy stimulus, like mitochondria depolarization ($\Delta\Psi_m \downarrow$), mitochondria can take two pathways: Ubiquitin-independent mitophagy or Ubiquitin-dependent mitophagy. Ubiquitin-dependent mitophagy begins with the accumulation of Pink1 protein in the outer mitochondrial membrane, Pink1 recruits Parkin to the outer mitochondrial membrane and ubiquitinates it. Ubiquitinated Parkin ubiquitinates some outer mitochondrial membrane proteins (OMM) like VDAC1 and MFN1/2. The polyubiquitination of the OMM proteins makes them recognizable to adaptors

protein like p62, and they recognize LIR motif of LC3 protein. LC3 promotes the mitophagy of damaged mitochondria by the phagophore formation. In the Ubiquitin-independent mitophagy, OMM proteins, like FUNDC1, BNIP3, NIX, BCL2L13, FKBP8 and NLRX1, and inner mitochondrial membrane (IMM) proteins like PHB2 and Cardiolipin interact with LIR motif of LC3 protein. Finally, the interaction with the LC3 protein promotes phagophore formation and mitophagy.

1.4.2. HBx and mitophagy induction

HBV replication in HepG2 cells and cultured primary rat hepatocytes is stimulated by HBx Ca^{2+} and ROS regulation. This Ca^{2+} regulation can be blocked by inhibition of the mitochondrial permeability transition pore (MPTP). The reported HBx-VDAC interaction suggests that HBx regulation of mitochondrial Ca^{2+} increases cytosolic Ca^{2+} levels (Rahmani et al., 2000),

In 2012, the interaction of HBx with anti-apoptotic proteins BCL-2 and Bcl-X was described. In this study, they propose a BH3-like motif in HBx protein that interacts with BCL-2 and Bcl-XL proteins. These interactions cause an elevation for cytosolic Ca^{2+} and apoptosis induction. This HBx-BCL2/BCLXL interaction and its consequent elevation of cytosolic Ca^{2+} is necessary for HBV replication (Geng et al., 2012).

HBV can modulate mitochondrial dynamics to promote mitochondrial fission and mitophagy. HBV induces the perinuclear clustering of mitochondria localization and promotes Drp1 leading to mitochondrial fission. HBx is the major protein of HBV genome to produce aberrant mitochondrial dynamics, even when it is expressed alone it can produce mitochondrial alteration. All of these findings, in addition to the fact that only Orthohepadnaviruses, but not other hepadnaviruses that lack HBx, cause HCC, contributed to the understanding that HBx plays a central role in HCC development (S. J. Kim et al., 2013).

HBx increases the PINK1/Parkin gene expression and accelerates Parkin partial recruitment to mitochondria. Additionally, it helps mitochondrial LONP1 to increase and reduce cytoplasmic LONP1 inducing the LONP1-Parkin pathway. HBx causes with these two pathways the enhancing mitophagy. These results were demonstrated in two conditions in two different cells. In the first one, they show that HBx can promote HBx in

Huh-7 cells but they can show it in HepG2 cells (Kim et al., 2013). Later, they induce not only the HBx expression but starving condition for HepG2 cells and found out the mechanism described before (Huang et al., 2018).

1.5. HBXIP and HBx

HBXIP, the product of LAMTOR5 gene, was initially identified in a yeast two-hybrid screening for HBx interactors. and named XIP (X Interacting Protein). Through this interaction, HBXIP was attributed a function of altering HBV replicative life cycle (Melegari et al., 1998). HBx was later shown to interact with the HBXIP-Survivin complex and suppress caspase activation, this suppression is Survivin-dependent (Marusawa et al., 2003). There is evidence of a relation between HBV infection and HBXIP levels. HBXIP levels were increased in chronic HBV cancerous and non-malignant liver tissues compared to healthy tissues (Lok, 2000). Finally, HBXIP seems to be negatively regulated by miR-501 in the stable HBV-producing cell line HepG2.2.15, and this suppression of HBXIP induces HBV replication, supporting a negative role for HBXIP in HBV life cycle (Jin et al., 2013).

HBXIP is conserved in mammalian species, with orthologous proteins in all vertebrate species. In preliminary analyses with the BLAST tool (www.ncbi.nlm.nih.gov/BLAST), we observed that the long isoform of HBXIP is found exclusively in mammals, which coincidentally are also the ones that suffer from Orthohepadnaviruses infections. In all other groups of vertebrates, only the short isoform is detected in the search for similarity with the long isoform of human HBXIP. This suggests that the short isoform is the canonical isoform of HBXIP, and the long has a more recent evolutionary origin and possibly co-evolved with HBx, indicating that this isoform provides some adaptive advantage against viral infection, which is consistent with its ability to suppress HBV replication.

1.6. Evolutionary concepts

In this study, one of the major questions involves a possible evolutionary relationship between HBx and HBXIP. Whether these two proteins coevolved and how this coevolution affected their sequences and functions is an important question that has never been explored in the literature.

Coevolution is a process of reciprocal evolutionary changes shaped by natural selection. Coevolution relationships like parasites-hosts, predators-prey, competitors and mutualists caused most evolutionary changes. The interaction between species derived in a natural selection on each other, continually shaping its evolutionary pathways. John N. Thompson; 2010 described four central points for understanding coevolution (Thompson, 2010).

One of these four points is that complex organisms require coevolved interactions to survive and reproduce. All complex organisms are the result of their coevolution relationships. We need interactions with other species to survive and these interactions caused both organisms to evolve together. A typical case of this coevolution dependency is mitochondria, a cellular organelle that originates from an ancient bacterium and coevolved with eukaryotic organisms (Cavalier-Smith, 2009).

The second point cited by John N Thompson is that species-rich ecosystems are built on coevolved interactions. These interactions form the base of all species-rich ecosystems and allow different species of an ecosystem to survive. Their absence could result in a big disbalance and collapse of the ecosystem. One example of this relationship is Rhizobia, which is the interaction between bacteria and plants, it fixes nitrogen and from the roots of legumes and plants (Piculell et al., 2008).

Coevolution takes multiple forms and generates a diversity of ecological outcomes. The arms race observed in evolution is generated from coevolutionary escalation. For example, parasite-host coevolution produces a selective pressure in rare genetic forms.

Over time, the selection fluctuates on the host and parasite populations (Nuismer & Thompson, 2006).

The coevolving interactions are changing geographic mosaics. We observe that different coevolution could be adaptable in a favorable way in different ecosystems. A good example is, in North America and Eurasia squirrels have coevolved with conifers, but where there are no squirrels the conifers have coevolved with crossbills (Benkman & Parchman, 2009).

In coevolution theory, parasite-host interaction is always changing and in order to maintain this interaction both need to change constantly and in the end, if there is an advantage for one of them it will suffer a selection for those that have these advantages. These will generate selective pressure on the host (Brockhurst & Koskella, 2013).

Much evolutionary information can be obtained from amino acid sequences of proteins and their respective coding sequences. It can give an insight into the protein's function and structure. In recent years, the context of these studies has changed to reveal more information about the sequences and gained an evolutionary focus. These sequences evolve under selective pressure which can be either negative or positive selection (Duggal & Emerman, 2012).

Extreme sequence conservation in some regions or amino acids in the interacting proteins usually means that this region is important to the function or structure of the protein, so they are maintained by a process called purifying or negative selection (Holmes, 2004). Most sequence changes in these highly conserved regions would lead to deleterious consequences, therefore these variants are expected to be eliminated by natural selection. In the host, most of the protein-coding genes are expected to evolve under purifying selection (Duggal & Emerman, 2012; Halabi et al., 2009).

Positive selection occurs when natural selection accelerates the fixation of advantageous characteristics in a species, this kind of selection is considered a signature of genetic conflict and is important in viral-host interactions. There are many studies that confirm that viral antagonist proteins have evolved under positive selection. These evolutionary conflicts of viral antagonists against their host protein targets are a typical case of the Red Queen Hypothesis (Duggal & Emerman, 2012).

The Red Queen Hypothesis is an evolutionary hypothesis that proposes an evolutionary system where the continuing development is needed to maintain the fitness relative to the coevolving system. Basically, for antagonistic species interactions, the relative fitness of each species does not increase over time because of its constant adaptation to the countering part. There is a reciprocal selection in the host-parasite interaction and an acceleration in the evolutionary rates in the clear need for an adaptation to the counteradaptation. In simple words, the host adaptation is often counteracted by the viral adaptation, which will be counteracted by the host again. These cyclical host-parasite interactions involve them in an evolutionary arms race (Strotz et al., 2018).

Positive selection leaves a weaker evolutionary signature in genes or amino acid residues than purifying selection. Using these signals, we can understand the impact of host-pathogen interactions on the evolution of protein interfaces (Daugherty & Malik, 2012). However, these evolutionary conflicts have some conditions that must be followed. It has to be a virus-host arms race where the reproductive fitness of either part is challenged. For example, host-virus interaction with non-pathogenic viruses are not likely to exert selective pressure on the host because there will not be any adaptive changes in the host viral antagonist during this infection (Emerman & Malik, 2010). The main idea is that positive selection in the viral antagonist protein can reveal footprints of ancient viruses' actions. It is possible to detect selective pressure through protein-coding DNA from different species (Grossman et al., 2010).

Positive and negative selection can be differentiated on the basis of dN/dS ratio, where dN is the rate of nonsynonymous mutations and dS is the rate of synonymous mutations.

$dN/dS < 1$ means the amino acid or protein is suffering negative pressure, if $dN/dS = 1$ indicates neutrality, and $dN/dS > 1$ indicates positive (diversifying) selection (Daugherty & Malik, 2012; Spielman & Wilke, 2015).

1.7. Methods of Evolutionary Analysis

The evolutionary forces that shaped the extant genetic diversity can be detected because of the advances in sequence generation and the increase of the statistical sophistication of relevant methods at an increased computational cost. Datamonkey is an open source framework webserver for analyzing evolutionary signatures in sequence data. The Datamonkey webserver provides a collection of methods for aligned coding-sequence analyses for imprints of natural selection, the results can be viewed in a fully interactive way in graphics (Weaver et al., 2018).

Datamonkey consists of a set of tools, like BUSTED which is a new approach to identify gene-wide evidence for positive selection, this method calculates if the nonsynonymous substitution is transiently greater than the synonymous rate. This open webserver offers the option to use an entire phylogeny data or a prespecified subset of foreground lineages. Selection is modeled as varying stochastically over branches and sites. BUSTED proposes an evidence metric for identifying sites subject to episodic selection on any foreground branches. An implementation is available on www.datamonkey.org/BUSTED, with a widget allowing the interactive specification of foreground branches (Murrell et al., 2015).

Another Datamonkey tool that helps in evolutionary analysis is MEME. Many questions in biology and genetics can be answered by identifying regions of protein genes that have undergone adaptive evolution. Natural selection is frequently transient or episodic and is often difficult to identify its imprint on protein genes. The webserver MEME (Mixed Effects Model of Evolution) can identify episodic and pervasive positive selection at an individual site. Different from other computational methods which detect the imprint of natural selection, MEME does not consider that natural selection is constant across all lineages, but episodic (Murrell et al., 2015).

2. Objectives:

2.1. Main objectives:

This study aims to investigate the possible functional and evolutionary relationship of the long isoform of HBXIP with the viral protein HBx using bioinformatics, biochemistry and cellular imaging tools, seeking to understand its functional relationship with the Ragulator complex.

2.2. Specific objectives:

- Phylogenetic and evolutionary analysis of HBXIP using bioinformatics tools, seeking to understand the evolutionary origin of the Long isoform and its relationship with HBx.
- Conduct protein-protein interaction assays between HBx and the subunits of the Ragulator complex, to verify whether HBx interacts with other subunits of this complex in addition to HBXIP and whether the presence of the N-terminal region in the long isoform affects the interaction.
- Conduct HBx colocalization studies with Ragulator subunits in different cells lines.
- Explore the role of HBx viral protein and Ragulator subunits in mitophagy.

3. Materials and methods

3.1. Plasmids

The plasmid pCDNA3-FLAG-c7orf59 was constructed by the former doctoral student of the group Nadia Rasheed (Rasheed et al., 2019). Parkin-GFP plasmid was a kind donation from Dr. Angela Saito (LNBio). The other plasmids were obtained from the Addgene repository (www.addgene.org), with the respective numbers in parentheses: pcDNA3.1-FLAG HBx WT (42596), pGFP-HA-HBx (24931), pGFP-HA-HBx NESM (24932), pRK5-FLAG-HBXIP (42326), pRK5-FLAG-p14 (42330), pRK5-p18-FLAG (42331). The plasmids were purified using a midiprep kit (Qiagen) and verified by sequencing.

3.2. Cell culture, transfection, and immunoprecipitation

HEK293-T, HepG2, Huh-7, U2OS and HeLa cells were cultured in DMEM High Glucose (Sigma) medium supplemented with 10% fetal bovine serum (FBS) with penicillin/streptomycin, at 37°C in a humidified atmosphere with 5% CO₂. For transfections using Lipofectamine 2000, the medium was changed one day before to antibiotic-free medium.

3.2.1. PEI transfection

HEK293-T cells were transfected in 10 cm plates using PEI (Polysciences) using a 1: 3 ratio (DNA: PEI). 10 µg of plasmid DNA were diluted in 300 µl of a 150 mM NaCl sterile solution. After diluting the DNA, we added 30 µL of PEI (Polyethyleneimine) 1 µg/µL solution. The final solution was incubated for 15 minutes at room temperature. During the incubation time, the medium was changed and then we gently added the final solution to the plate while agitating it in circular movements. After 24 hours the cells were collected.

3.2.2. Transfection of HEK293-T cells in 96-well plate

For preliminary immunolocalization experiments, HEK293-T cells were transfected with different plasmid constructs in 96-well plates (NEST) using PEI. The transfection was optimized with different conditions. The best condition was with 1:12 (DNA: PEI) and a

total of approximately 10,000 cells for each well, with incubation for 24 hours after transfection.

3.2.3. Transfection of HepG2 cells in 24 well plates with Lipofectamine 2000

For confocal imaging experiments, HepG2 cells were plated in 24 well plates with glass coverslips and later transfected in antibiotic-free DMEM using Lipofectamine 2000 (Life Technologies) according to the manufacturer's instructions. 2.5×10^5 cells were plated one day before transfection in 500 μ l of DMEM growth medium with 10% FBS and penicillin/streptomycin so that cells would be 70-90% confluent at the time of transfection. For transfection, a solution of 0.8 μ g plasmid DNA in 50 μ l of Opti-MEM I Reduced Serum Medium without serum and other solution of 2 μ l Lipofectamine in 50 μ l of Opti-MEM-I Medium were prepared and incubated for 5 minutes at room temperature. Later, these solutions were combined and incubated for additional 20 minutes at room temperature, and then the mixture was transferred to each well.

3.2.4. Transfection of HeLa and U2OS cells in 24 well plates with Lipofectamine 3000

HeLa and U2OS were transfected in 24-well plates using Lipofectamine 3000 (Life technologies) according to the manufacturer's instructions. The procedure was similar to HepG2 transfection, with the following changes. For each well to be transfected, 1 μ g of plasmid DNA and 2 μ l of P3000 reagent were combined in in 25 μ l of Opti-MEM I Reduced Serum Medium without serum and separately, 0.75 μ l Lipofectamine was diluted in 25 μ l of Opti-MEM I Medium. The two solutions were combined and incubated for 15 minutes at room temperature, and then the mixture was transferred to each well.

3.2.5. Transfection of Huh-7 in 96-well plates

Huh-7 cells were transfected in 96-well plates with glass bottom (Corning) using Lipofectamine 3000 (Life technologies) according to the manufacturer's instructions. The procedure was similar to U2OS transfection, with the following changes. For each well to be transfected, 100 ng of plasmid DNA and 1 μ l of P3000 reagent were combined in in 5 μ l of Opti-MEM I Reduced Serum Medium without serum and separately, 0.15 μ l Lipofectamine was diluted in 5 μ l of Opti-MEM I Medium. The two solutions were

combined and incubated for 15 minutes at room temperature, and then the mixture was transferred to each well.

3.2.6. Immunocytochemistry

After 24 hours of transfection, the medium was removed and the cells were fixed with a 20-minute incubation in fixative solution (3.7% formaldehyde, 0.2% Triton X-100 in 1x PBS) at room temperature. The cells were then washed twice with 1x PBS and blocked for 30 minutes (3% BSA + 0.8% Triton X-100 in 1x PBS). After two additional washes with PBS + 0.05% Tween-20 solution, the cells were incubated with primary anti-FLAG 1:1000 or anti-HA 1:1000 (Sigma) overnight at 4°C. The next day, the cells were washed three times with PBS + 0.05% Tween-20, then incubated for 3 hours at room temperature with secondary anti-mouse antibody conjugated to Alexa 546 (Life technologies) diluted 1:1000 in PBS + 0.05% Tween-20. After washing three more times, DAPI was added at the final concentration of 1 µg/mL in PBS. The images were acquired in a Confocal Leica Microscope with 63x objective at LNBio.

3.2.7. Immunoprecipitation

For immunoprecipitation experiments, after 24 hours of transfection, the cells were washed with PBS 1x and the pellets were frozen at -20°C. Lysates from these cells were prepared by resuspension in ice cold lysis buffer (30 mM Tris-HCl pH 8.0; 150 mM NaCl; 3 mM EDTA and 0.3% IGEPAL CA-650) supplemented with SIGMAFAST Protease Inhibitor cocktail, EDTA-Free (Sigma), and incubation on ice for 25 minutes followed by centrifugation at 20.000 x g at 4°C for 30 minutes. The cleared lysates were subjected to anti-FLAG immunoprecipitation using anti-FLAG antibody conjugated to agarose beads (Sigma). The immunoprecipitates were washed 3 to 5 times in lysis buffer and the beads were resuspended in SDS-PAGE loading buffer. Then, the samples were separated by SDS-PAGE on 15% acrylamide gels, followed by transfer to nitrocellulose membranes (GE) previously activated with transfer buffer, for 8 min at 30 V in Trans-Blot Bio-Rad equipment. Then the membranes were blocked with 3% powdered milk in TBST and incubated overnight with primary Anti-FLAG or Anti-HA antibody (Sigma). The next day, the membranes were washed and incubated with secondary antibody for 4 hours.

Western blot results were detected by chemiluminescence with ECL reagent (GE) and detected using ECL Hyperfilm or a GE Healthcare ImageQuant LAS 500 detector.

3.2.8. Mitophagy live cell experiments

The day before the experiment, Huh-7 cells were plated and transfected as described before. A solution of 1 $\mu\text{g}/\text{mL}$ Mitotracker Deep Red and 5 $\mu\text{g}/\text{mL}$ Hoescht was applied to each well; meanwhile, the live cell chamber of Operetta was installed and incubation conditions were set to 37°C and 5% CO₂. After stabilization of the conditions, four initial images of the cells were captured with 20X objective at 5 minutes intervals to set the baseline. Then, mitochondrial depolarization was stimulated by adding 17 μM CCCP (Carbonyl cyanide 3-chlorophenylhydrazone). Then, images were acquired every 20 minutes with 20x objective in GFP, Hoescht and Mitotracker Deep Red channels. After 6 hours of incubation, we collected the plate and fixed the cells, the immunocytochemistry was made following the described protocol in immunocytochemistry chapter. Finally, images of the fixed cells were acquired using the Operetta the same objective and channels of the live cell experiment, including the Alexa 546 channel to visualize FLAG-tagged transfected proteins. In the live cell experiment, the following constructions with Mitotracker Deep Red (647) and Hoescht channel were used (**Table 1**).

Table 1: Plasmid constructions and their respective preset channels for image capture in Operetta.

Plasmid construction	Secondary antibody	Operetta preset channel
HBXIP Long-FLAG	Alexa 546 conjugate	Alexa 546
c7orf59-FLAG	Alexa 546 conjugate	Alexa 546
p14-FLAG	Alexa 546 conjugate	Alexa 546
p18-FLAG	Alexa 546 conjugate	Alexa 546
Parkin-GFP	N. A.	GFP
HBx-GFP-HA	N. A.	GFP

3.3. Evolutionary and bioinformatic analyses

3.3.1. Preliminary analyses

To explore the possible evolutionary relationship of HBx and HBXIP, a detailed survey was carried out in the literature and in the NCBI (www.ncbi.nih.gov) and Uniprot (<http://www.uniprot.org/>) databases, searching for the following information in a sufficient number of vertebrate species: 1- presence / absence of the long isoform of HBXIP and, if present, number of amino acid residues from the orthologs of each species; 2- presence / absence of infection by HBV virus or other related viruses (Orthohepadnaviridae) in these same species. The information obtained was represented graphically in a phylogenetic tree of vertebrates. Then, we performed a survey of the amino acid sequences of HBXIP and HBx in several species of mammals and Orthohepadnavirus, respectively.

3.3.2. Evolutionary analyses

The bioinformatic analysis of nucleotide sequences to detect selective pressure was divided in three different pipelines. The first pipeline showed the pathway to obtain evolutionary information for proteins. Our first search was the human HBXIP amino acids sequence (UniProt code A0A0C4DGV4), which was BLASTed in NCBI Database using the BLASTp tool to find its orthologues, limiting our search to mammalian proteins. The sequences of HBXIP orthologues were downloaded and aligned in the T-Coffee server (<http://tcoffee.crg.cat/>) (Notredame et al., 2000). Finally, for graphic representation of the alignment, coevolutionary cluster and phylogeny from HBXIP orthologues we used BoxShade (Expasy), Pymol and ITOL (Letunic & Bork, 2019) (<https://itol.embl.de/>) respectively.

The second pipeline made was for evolutionary analysis of selective pressure from Ragulator complex proteins and HBx known structure interactors. Amino acid sequences of Ragulator subunits and selected HBx interactors were retrieved from UniProt database

(**Table 2**) and were BLASTed in NCBI database using TBLASTn tool. The search was limited to three conditions: primates, rodents and primates and rodents combined. The orthologues from the desired proteins were downloaded and aligned in T-Coffee server.

Table 2: UniProt codes for Ragulator subunits and known HBx interactors.

Protein Name	UniProt code
HBXIP	A0A0C4DGV4
C7orf59	Q0VGL1
MP1	Q9UHA4
p14	Q9Y2Q5
p18	Q6IAA8
BCL-2	Q07817
DDB1	Q16531

The third pipeline was used for the analyses of HBx protein orthologues. The information of HBV genotypes DNA sequences were retrieved from VIRALZONE database (<https://viralzone.expasy.org/>). These sequences were downloaded and aligned with HBx sequence (ENA|CAB41697) to select only the HBx protein sequences from each genotype.

The resulting alignments were uploaded in CleanStopCodons Hyphy tool (Kosakovsky Pond et al., 2020) to remove stop codons, then, the alignments were submitted to BUSTED (Murrell et al., 2015) and MEME (Murrell et al., 2012) Datamonkey tools at <https://www.datamonkey.org/> to detect selective pressure and generate graphics.

3.3.3. Subcellular localization prediction

For the prediction of subcellular location of the HBXIP isoforms using Deeploc (Almagro Armenteros et al., 2017), the amino acid sequences downloaded from UNIPROT database were inserted in the Deeploc search window (<http://www.cbs.dtu.dk/services/DeepLoc/>) and the results were displayed as a table. Human Protein Atlas public database of immunolocalization images (Thul et al., 2017),

was used to retrieve experimental data of subcellular localization for the endogenous Ragulator subunits and compare with our own results.

3.3.4. RNA-seq Analyses

The FASTA files of RNA-seq data were obtained from the “Atlas of RNA sequencing profiles for normal human tissues”, published in 2019 in Scientific Data journal (Suntsova et al., 2019). We used all the uploaded data in SRA (Sequence Read Archive) database, and the number of FASTA files and their respective tissues were summarized in **Table 3**.

Table 3: Number of samples or FASTA files per tissue from Atlas of RNA sequencing profiles for normal human tissues uploaded to SRA database obtained from “Atlas of RNA sequencing profiles for normal human tissues”

Tissues	Number of samples
Adrenal gland	6
Bladder	5
Brain (mixed white and gray matter)	8
CD138+ enriched bone marrow cells	11
Cervix	2
Cervix, endometrium	2
Colon	6
Colon, epithelium	6
Esophagus	11
kidney	8
Liver	11
Lung	8
Mammary gland	5
Ovary	4

Pancreas	8
Prostate	7
Skeletal muscle	6
Skin	6
Small intestine	9
Stomach	15
Thyroid gland	6
Tonsil	7
Uterus (myometrium)	2
Whole blood nuclear cells	6
Total	165

The RNA sequencing data were uploaded to the Galaxy web platform (Afgan et al., 2018) and the public server at usegalaxy.org was used to analyze these data. “Download and Extract Reads” in FASTA/Q format from NCBI SRA (Galaxy Version 2.10.8+Galaxy0) was used to extract FASTA information from SRA database by the SRR (SRA Run Accession) accession code for each sample. A FASTA file with all the transcript sequences of *Lamtor5* gene found in ENSEMBL database (<https://www.ensembl.org/>). Only protein-coding transcripts were selected (**Table 4**).

Salmon quant (Patro et al., 2017) was used to perform dual-phase, reads or mapping-based estimation of transcript abundance from RNA-seq reads (Galaxy Version 1.3.0+Galaxy1). Reads Salmon quantification mode was selected, and as reference transcriptome the FASTA file with the six *Lamtor5* transcript sequences found in ESEMBL database was uploaded. Finally, 31 Kmer length selection parameter recommended by the Salmon tool in Galaxy web server was used. As input, the 165 FASTA files were uploaded without specifying the strandedness of the reads. The type of index chosen was “quasi” as recommended in Salmon tool and we used all the “flags” default parameters to

run the job. From the Salmon results, the TPM (Transcript Per Kilobase Million) was registered in a table to calculate the average expression in each tissue for each Lamtor5 transcript. These calculations were made in Prism9 and projected in a heat map and in scale bars graphs.

Table 4: Transcript codes of LAMTOR5 isoforms from ESEMBL database number of base pairs (BP) and protein amino acids (aa) length.

Name	Transcript ID	bp	Protein
LAMTOR5-201	<u>ENST00000256644.8</u>	868	<u>173aa</u>
LAMTOR5-203	<u>ENST00000474861.6</u>	1154	<u>90aa</u>
LAMTOR5-204	<u>ENST00000483260.5</u>	694	<u>90aa</u>
LAMTOR5-206	<u>ENST00000602318.6</u>	620	<u>91aa</u>
LAMTOR5-207	<u>ENST00000602858.5</u>	553	<u>79aa</u>
LAMTOR5-208	<u>ENST00000614544.1</u>	846	<u>173aa</u>

3.3.5. Structural model of HBXIP_3

A structural model of HBXIP_3 isoform was obtained by homology modeling using Modeller 9.24 with an alignment of HBXIP_3 with HBXIP_1 amino acids sequence. The HBXIP_1 PDB was obtained from a Ragulator structure uploaded in PDB database with 6ehp code. Finally, the was refined with Foldt (standalone version) and visualized in Pymol.

4. Results

4.1. HBXIP Molecular Evolution Analyses

Bioinformatic analyses are helping to discover the function, expression, and origin of different proteins. There are many methods to make this kind of analysis, but the amino acid sequences comparative methods and the motif search are the most common.

Using NCBI public database, we obtained a reliable dataset of human HBXIP orthologues in vertebrates. Analyzing the 100-best scored orthologues for HBXIP protein, it was notorious that the C-terminal Roadblock domain is extremely conserved, but in the N-terminal regions from some orthologues showed variation in the protein length (**Figure 6**).

```

Homo_sapiens      1  MEPGAGHLDGHRAGSPSLRQALCDGSAVVFSSKERGRCTVINFPLEAPLRSTPRSROQVI
Pan_paniscus     1  MEPGAGHLDGHRAGSPSLRQALCGGSAVVFSSKERGRCTVINFPLEAPLRSTPRSROQVI
Mus_musculus     1  -----MFRKERGRGAALNFVPLETPSRRPFRSSHVI
Sus_scrofa       1  -----MTSSRRKRRDCATKFVPLEAPSRASCSHHVT
Rattus_norvegic  1  MPANAGSVV-----VAMSFKKECERGATMNLFPLETPSARFPFRSSHVT
Felis_catus      1  -----
Gallus_gallus   1  -----
Loxodonta_afric  1  -----
Pteropus_vampyr  1  -----

Homo_sapiens     61  EACGGEGRAVPLGSEPEWSV----GMEATLEQHLEDT-MKNPSIVGVLC TDSQGLNLC
Pan_paniscus     61  EACGGEGRAVPLGSEPEWSV----GMEATLEQHLEDT-MKNPSIVGVLC TDSQGLNLC
Mus_musculus     33  EVCGEGGSVLLRGPGLGRTA----CGMEATLEQHLEDT-MKNPSIVGVLC TDSQGLNLC
Sus_scrofa       33  DQAEPEFSF--LGEDPPWVVRPADGGMEATLEQHLEDT-MKNPSIVGVLC TDSQGLNLC
Rattus_norvegic  44  EACCEGGSVLRGLGLGRTV----YMEATLEQHLEDT-MKNPSIVGVLC TDSQGLNLC
Felis_catus      1  -----MEATLEQHLEDT-MKNPSIVGVLC TDSQGLNLC
Gallus_gallus   1  -----ME@TLEQHLEDT-MKSPAVVGVLC TDSQGLNLC
Loxodonta_afric  1  -----MEATLEQHLEDT-MKNPSIVGVLC TDSQGLNLC
Pteropus_vampyr  1  -----MEATLEQHLEDTRMKNPSIVGVLC TDSQGLNLC

Homo_sapiens     115 CRGTLSDHAGVISVLAQQAAKLTS DPTDIPVVCLES DNGNIMI QKHDGITVAVHKMAS
Pan_paniscus     115 CRGTLSDHAGVISVLAQQAAKLTS DPTDIPVVCLES DNGNIMI QKHDGITVAVHKMAS
Mus_musculus     87  CRGTLSDHAGVISVLAQQAARKLTS DPTDIPVVCLES DNGNIMI QKHDGITVAVHKMAS
Sus_scrofa       90  CRGTLSDHAGVISVLAQQAAKLTS DPTDIPVVCLES DNGNIMI QKHDGITVAVHKMAS
Rattus_norvegic  98  CRGTLSDHAGVISVLAQQAAKLTS DPTDIPVVCLES DNGNVMIQKHDGITVAVHKMAS
Felis_catus      33  CRGTLSDHAGVISVLAQQAAKLTS DPTDIPVVCLES DNGNIMI QKHD@ITVAVHKMAS
Gallus_gallus   33  CRGTLSDHAGVISVLAQQAAKLTS DPTDIPVVCLES DSGNIMI QKHD@ITVAVHKMAS
Loxodonta_afric  33  CRGTLSDHAGVISVLAQQAAKLTS DPTDIPVVCLES DSGNVMIQKHDGITVAVHKMAS
Pteropus_vampyr  34  CRGTLSDHAGVISVLAQQAAKLTS DPTDIPVVCLES DNGNIMI QKHDGITVAVHKMAS

```

Figure 6: Multiple sequence alignment of LAMTOR5 protein sequences from representative vertebrate species made using T-Coffee and Expsy Box Shade tools.

To confirm that HBXIP protein length was different among its orthologues in eukaryotes, a search in the OrthoDB database (Kriventseva et al., 2015) was made with the name of LAMTOR5. In this database, the median protein length information was verified for each group. In most groups, the median protein length for HBXIP is 91 amino acid residues, which corresponds to the length of the Roadblock domain. Interestingly, a different median length of 145 residues appears in the Euarchontoglires (Supraprimates) superorder. The major mammalian orders in Euarchontoglires are Rodentia and Primates, which showed median protein lengths of 145 and 155 for HBXIP, respectively. Another exception is Aves with a median length of 79 residues for HBXIP (**Figure 7**).

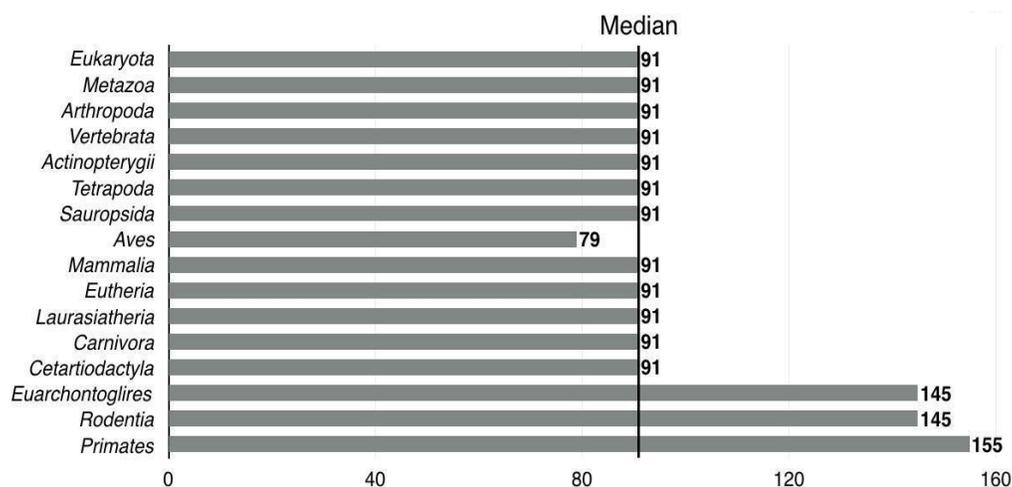


Figure 7: Graphic representation of the median protein length of HBXIP in different groups obtained in OrthoDB database (<https://www.orthodb.org/>).

This finding indicated that the short HBXIP isoform is ancestral in eukaryotes while its long isoform is a recent evolutionary acquisition of the Supraprimates superorder. With this possibility in mind, the next step was to make a phylogenetic tree of the Long isoform to understand how the LAMTOR5 gene evolved and why these different protein lengths are found only in some groups of mammals. In the BLASTp search, only mammalian protein sequences were selected, the best scored and not repeated species amino acid sequences for HBXIP were downloaded. Next, we analyzed these sequences in T-Coffee simple alignment so we could obtain a FASTA file to upload in Simple Phylogeny, a tool from EMBL which uses Clustal algorithm with a neighbor-joining matrix (Madeira et al.,

2019). Finally, we got our phylogenetic tree in Newick format and illustrate it using used ITOL (**Figure 8**).

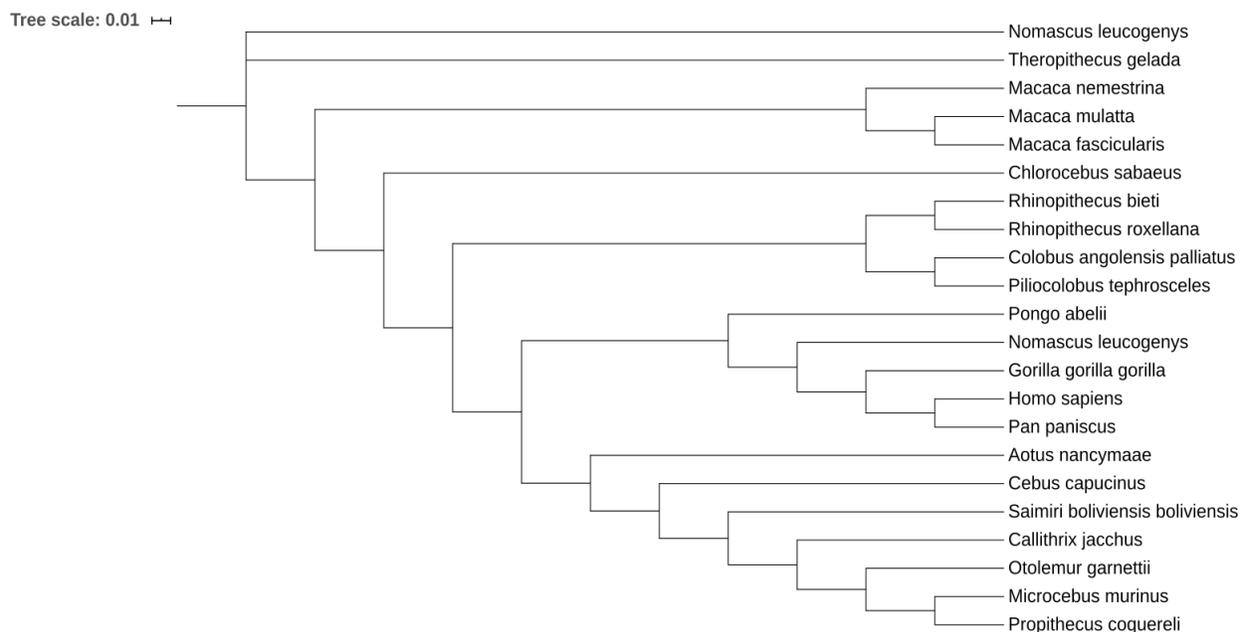


Figure 8: HBXIP Long protein phylogenetic tree in primates made with ClustalW2 algorithm in Phylogeny tool form EMBL sever.

This new information suggested that HBXIP Long isoform could have some evolutionary relation with its viral interactor HBx. The relation between the protein length and the interaction with HBx could be an evolutionary adaption to viral infection. To test this hypothesis, initially we made a graphic showing the HBXIP protein length obtained in the String database (Szklarczyk et al., 2019) for different mammalian species correlating it with their susceptibility to Orthohepadnaviruses infection. This co-occurrence graphic shows a positive relation between the occurrence of the Long isoform of LAMTOR5 and Orthohepadnaviruses infection (**Figure 9**).

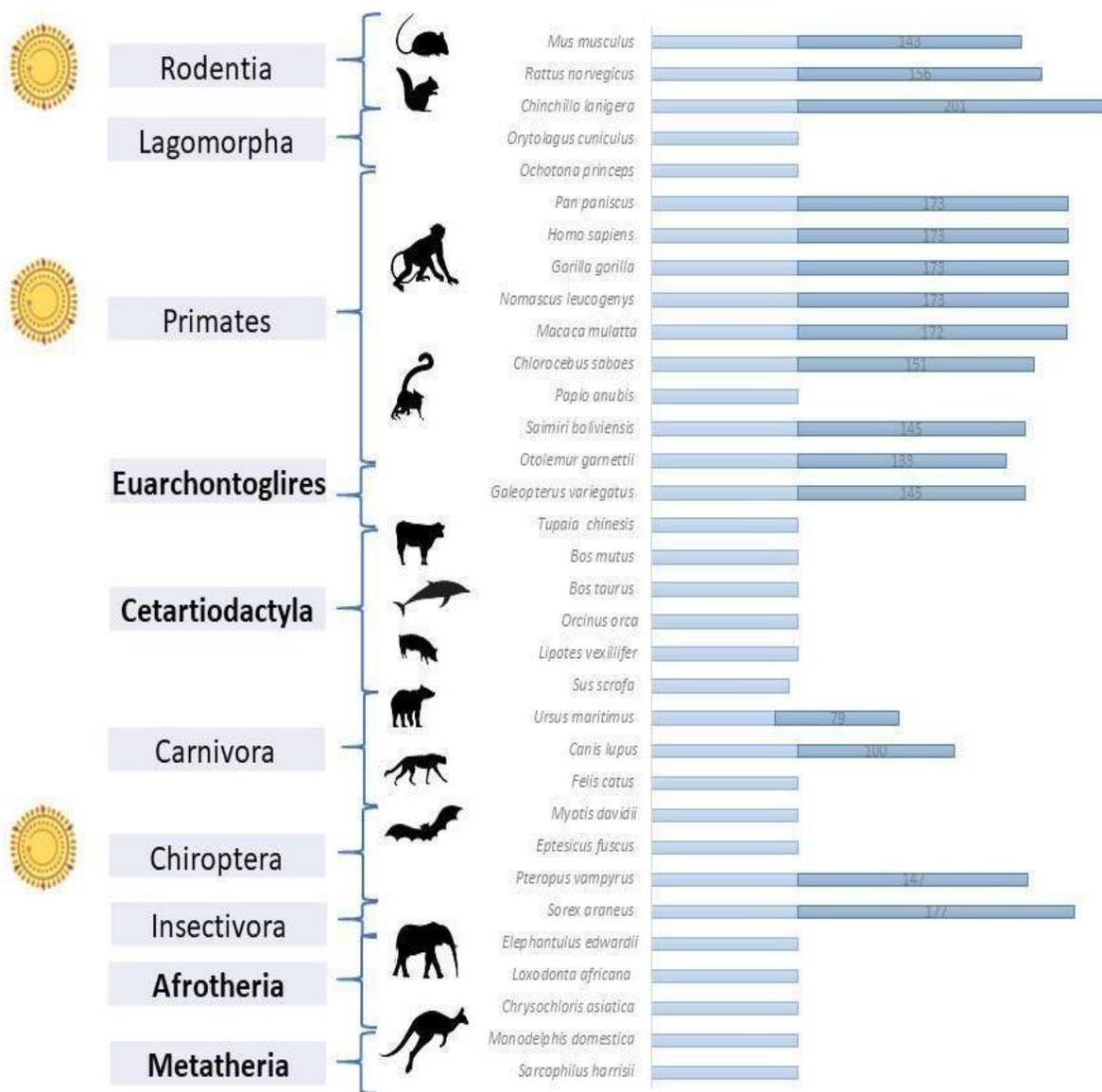


Figure 9: Co-occurrence of the LAMTOR5 long isoform and susceptibility to infection with Orthohepadnaviruses (indicated with an illustration of the viral particle) in mammalian species. A schematic representation of LAMTOR5 protein sequence is shown next to each species name. Light blue represents the canonic isoform of LAMTOR5 (short isoform). Darker blue represents additional residues corresponding to the long isoform.

In search for evolutionary answers to our investigation, Datamonkey was the tool of choice. Datamonkey is a collection of state-of-the-art statistical models and bioinformatics tools that identify selective pressure and other evolutionary events (Weaver et al., 2018). It is largely known that to maintain the host/parasite interaction, proteins must co-evolve. One way to search for these interactions sites is to find out if both interacting proteins have suffered selective pressure.

Euarchontoglires was the only clade that showed a constant presence of HBXIP long isoforms, so our search for evolutionary analyses was limited to *Lamtor5* genes from rodents and primates. The *Lamtor5* nucleotide sequences from rodents and primates were aligned in T-Coffee and submitted to BUSTED (Branch-site Unrestricted Statistical Test for Episodic Diversification) (Murrell et al., 2015), to find whether *Lamtor5* has experienced positive (diversifying) selection. BUSTED calculates three omega values ($\omega_1 \leq \omega_2 \leq 1 \leq \omega_3$). Then, it estimates the portion of sites belonging to each ω . The unconstrained model refers to the alternative model, that is, the model with selective pressure. BUSTED then tests for positive selection by comparing this model fit to a null model also known as constrained model where $\omega_3=1$. If the null hypothesis is rejected, then there is evidence that at least one site has, at least some of the time, experienced positive selection on the foreground branches. The sites of selective pressure found in primate *Lamtor5* sequences are mostly in the N-terminal region which is exclusive of the Long isoform (**Figure 10**).

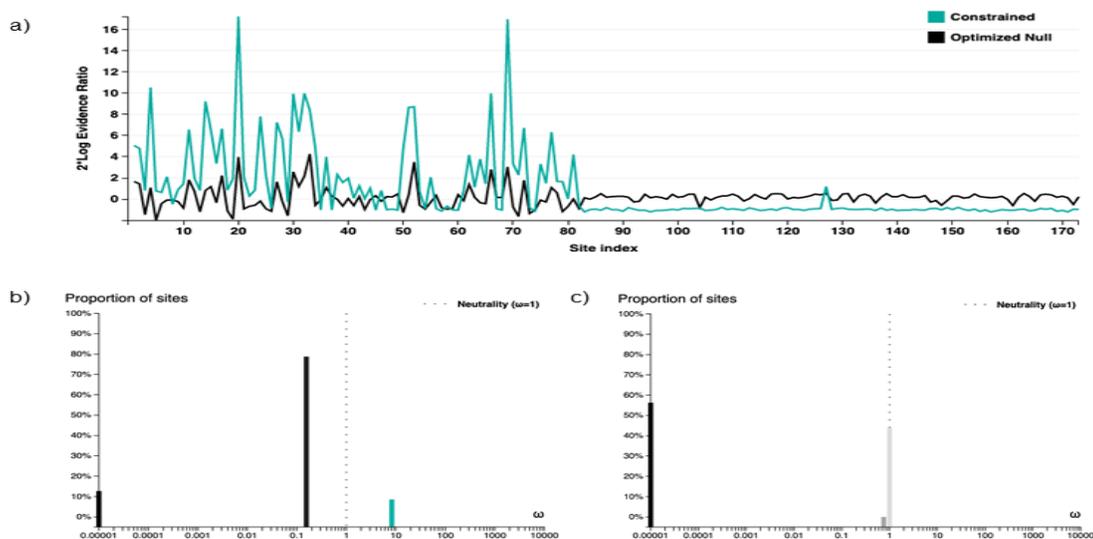


Figure 10: BUSTED results for selective pressure in *Lamtor5* from primates before BUSTED update. A) Model Test Statistics Per Site for *Lamtor5* nucleotide sequence. B) and C) Percentage of *Lamtor5* sites suffering selective pressure in the unconstrained and constrained models, respectively. The green bar shows the proportion of sites that suffer selective pressure. The dotted grey line indicates the neutrality threshold ($\omega=1$). The constrained model refers to the null model with $\omega=3=1$.

To verify whether other Ragulator subunits in addition to HBXIP also experienced diversifying selective pressures, their sequences were also analyzed using BUSTED. The other Ragulator subunits did not show any evidence of selective pressure in BUSTED analyses, as evidenced by low LRT (Likelihood Ratio Test). In comparison, DDB1 and BCL-2, two HBx known interactors, showed evidence of selective pressure providing further evidence that the selective pressure on HBXIP could be attributed to its interaction and evolutionary relationship with HBx. The LRT and p-values for each one of these sequences are listed in **Table 5**.

Table 5: Summary of BUSTED analyses before update for Ragulator proteins, rodents and primates Lamtor5, DDB1, Bcl2-L and X protein genes.

Protein	Family	LRT	P-VALUE
LAMTOR5	PRIMATES	29.096594	4.81E-07
	RODENTS	4.8990748	0.0863335
	RODENTS AND PRIMATES	0	1
LAMTOR1	PRIMATES	0	1
LAMTOR2	PRIMATES	0	1
LAMTOR3	PRIMATES	0	1
LAMTOR4	PRIMATES	3.5724075	0.1675952
BCL2-L	PRIMATES	88.623798	0
DDB1	PRIMATES	30.081266	2.94E-07
X	ORTHOHEPADNAVIRUS	31,17	5.94E-15

The latest version of BUSTED implemented the calculation of the coefficient of variation of synonymous rates (CV(SRV)). Our table of values was also updated (**Table 6**) and we found that the selective pressure on *Lamtor5* gene in primates does not reach the p-value to be significant, getting close to neutrality. We hypothesized with more sequences will get a p-value ≤ 0.05 , but as we can see on the graphic for the distribution of the model statistics per site of just *LAMTOR5* long isoform gene present the selective pressure (**Figure 11**).

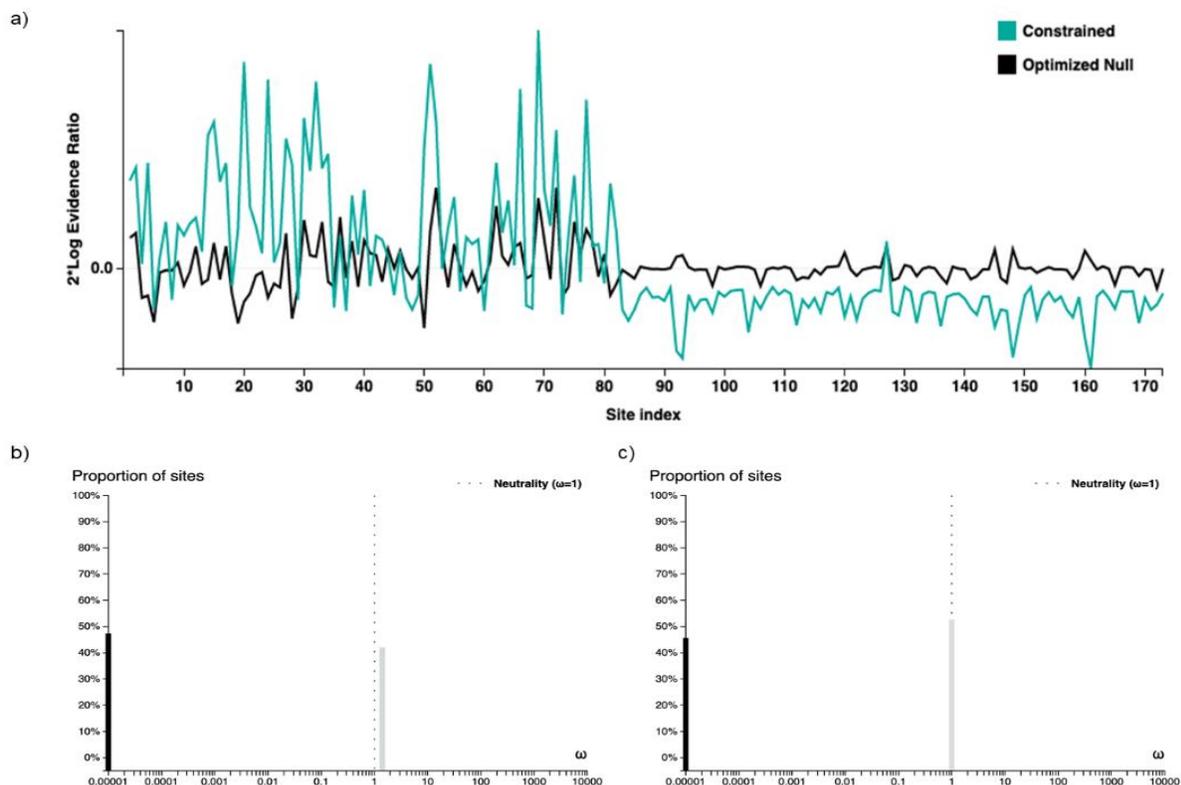


Figure 11: BUSTED results for selective pressure in *Lamtor5* from primates after BUSTED update. A) Model Test Statistics Per Site for *Lamtor5* nucleotide sequence. B) and C) Percentage of *Lamtor5* sites suffering selective pressure in the unconstrained and constrained models, respectively. The green bar shows the proportion of sites that suffer selective pressure. The dotted grey line indicates the neutrality threshold ($\omega=1$). The constrained model refers to the null model with $\omega = 3=1$.

With the CV(SRV) new update of BUSTED, rodents also showed selective pressure. The distribution of the (Table 6) Model Test Statics Per Site from rodents *Lamtor5* sequence shows better signals for selective pressure in the Long isoform, repeating the results with primates without CV (SRV) (Figure 12), showing another piece of evidence for the relationship of the Long isoform of HBXIP with HBx.

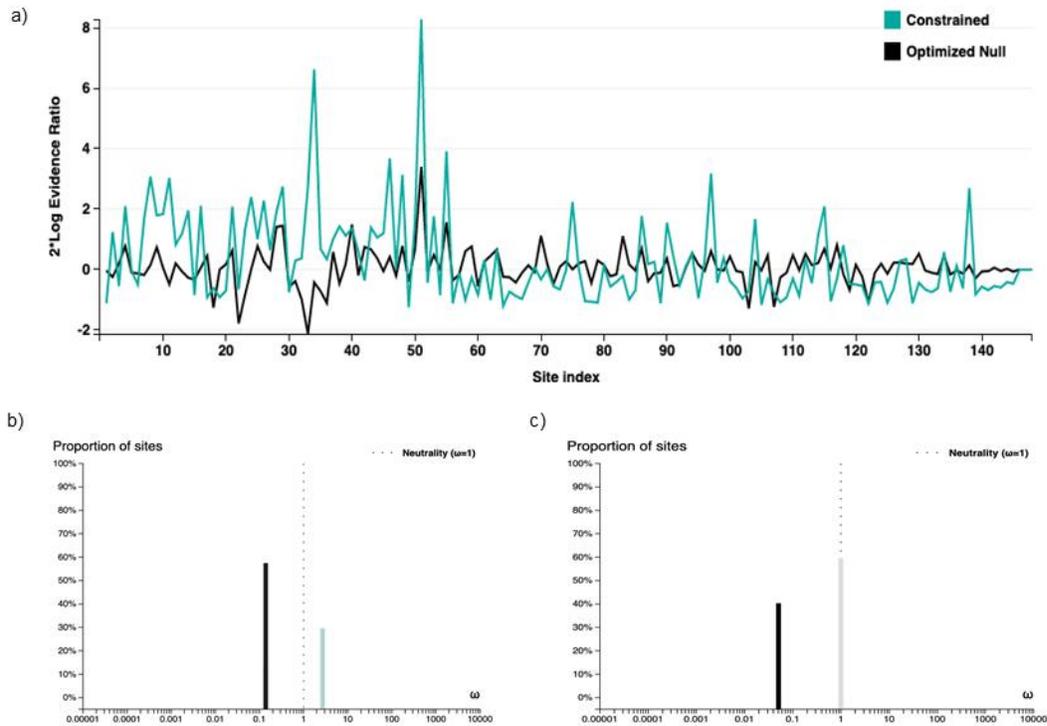


Figure 12: BUSTED results for selective pressure in *Lamtor5* from rodents. A) Model Test Statistics Per Site for *Lamtor5* nucleotide sequence. B) and C) Percentage of *Lamtor5* sites suffering selective pressure in the unconstrained and constrained models, respectively. The green bar shows the proportion of sites that suffer selective pressure. The dotted grey line indicates the neutrality threshold ($\omega=1$). The constrained model refers to the null model with $\omega=3$.

Table 6: Summary of BUSTED analyses (after update) for Lamtor5, DDB1, Bcl2-L, Ragulator subunits and X protein genes from rodents and primates.

Protein	Family	LRT	P-VALUE
	PRIMATES	0.99	3.05E-01
LAMTOR 5	RODENTS	8.87	5.93E-03
	RODENTS AND PRIMATES	0.00	5.00E-01
LAMTOR 1	PRIMATES	0.00	5.00E-01
LAMTOR 2	PRIMATES	0.00	5.00E-01
LAMTOR 3	PRIMATES	0.55	3.89E-01
LAMTOR 4	PRIMATES	2.79	1.24E-01
BCL 2L	PRIMATES	73.08	5.55E-17
DDB1	PRIMATES	18.66	4.44E-05
X	ORTHOHEPADNAVIRUS	31.17	8.52E-09

BUSTED calculated if there exists any evidence of selective pressure in the nucleotide sequence, but if it does not find this evidence we cannot conclude that there is no selective pressure in the sequence. BUSTED shows that the region that could be suffering selective pressure is in the Long isoform region of *Lamtor5* sequence. MEME (Mixed Effects Model of Evolution) aims to detect sites evolving under positive selection in a proportion of branches of the phylogenetic tree, considering that the same sequences of primate *Lamtor5* analyzed in BUSTED were also analyzed in MEME. MEME identified five amino acid residues suffering selective pressure, all of them located in the N-terminal region exclusive of the HBXIP Long isoform, as shown in the LRT graphic (**Figure 13**), it agrees with the BUSTED results (**Table 7**).

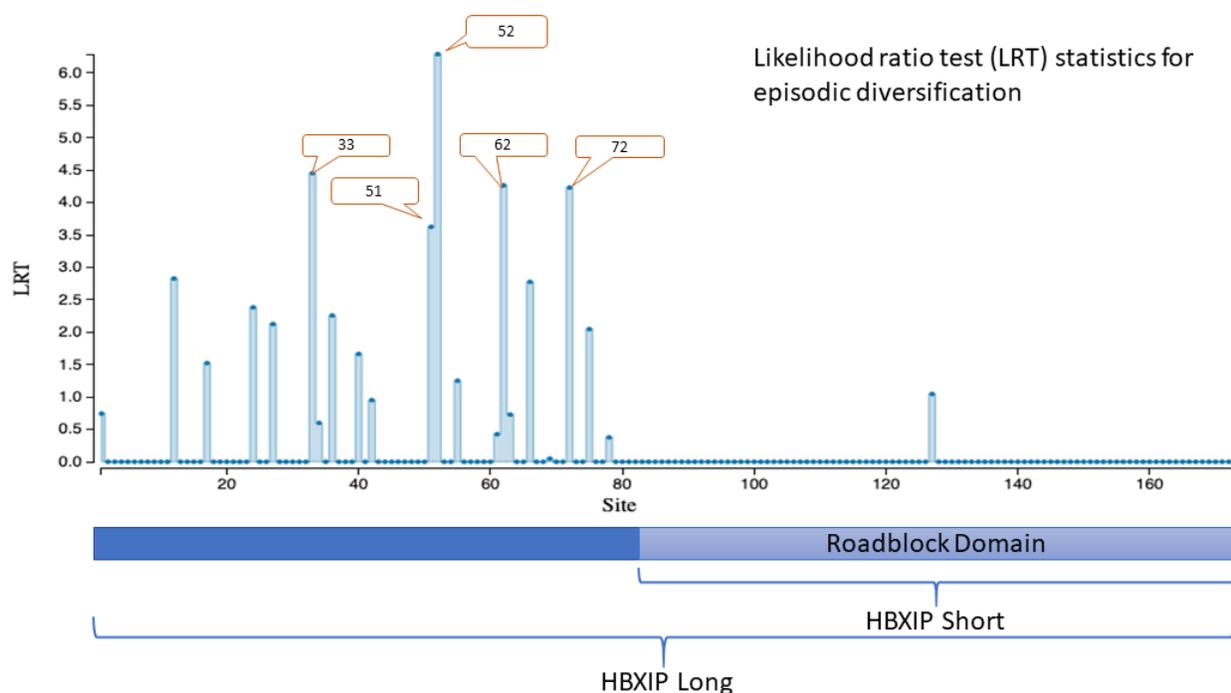


Figure 13: Likelihood Ratio Test (LRT) statistics graphic obtained with MEME for episodic diversification of HBXIP protein sequences from primates. The numbers inside the red boxes indicate the positions of amino acid residues which showed evidence of positive selection with p -value < 0.1 . Below the LRT graph is a representation of HBXIP primary sequence and its isoforms and domains.

Table 7: A summary of MEME results with 0.1 p -value cutoff showing the five amino acid residues of HBXIP Long sequence (highlighted in Figure 13) that suffer positive selective pressure.

Site	Partition	LRT	p-value	MEME LogL	FEL LogL
33	1	4.45	0.05	-18.01	-16.34
51	1	3.62	0.08	-24.69	-24.34
52	1	6.15	0.02	-20.89	-17.94
62	1	4.27	0.06	-16.86	-15.63
72	1	4.23	0.06	-15.70	-15.33

4.2. HBx molecular evolution analyses

Viral Zone database have the X Orthohepadnaviruses protein orthologues sequences search. Viral Zone is a comprehensive resource bringing together textbook knowledge with genomic and proteomic sequence for virus. As expected, the alignment shows many variations in the amino acids sequences from different X proteins of Orthohepadnavirus family that infects different hosts as shown in **Figure 14**, These amino acid sequences give us a strong clue that the X proteins have suffered selective pressure from their interaction with the host proteins.

```

X_GSHV      1  MAARLCCQLDSSRDVLLLRPLRGQPSGSPVSGT SAGSPSSAASAFSSGHQADIPVGR LPA
X_WHV1      1  MAARLCCQLDPARDVLLLRPFGSQSSGPPFPRPSAGS AASPASSLSASDES DPLGRLPA
X_ASHV      1  MAARLCCQLDSSRDVLLLRPFGSESGGPAVSRP SAGSASRADSP LPSAAESHLP LGR LPA
X_HBVOR     1  MAARLCCQLDTARDVLC LRPVGAESRGRPFSGS V GALPPSSPPAVPADHGAHLSLRGLPV
X_WMBHV     1  MAARLCCYLDPERDVLCLRPLOAEP SGRPF SGLSRPAETA AAAAVPAFHGAHLSLRGLPS
X_HBVD3     1  MAARLCCQLDPARDVLC LRPVGAESRGRPFSGSLGTLSSPSPSAVPTDHGAHLSLRGLPV

X_GSHV      61  CFYSSAGPCCLGFTCADLRTMDSTVNF----VPW HAKRQLGMMQKDF--WTAYIRDOLLT
X_WHV1      61  CFASASGPCCLVVTCAELRTMDSTVNF----VSWHANRQLGMPSKDL--WTPYIRDOLLT
X_ASHV      61  CFASPSGPCCLGFTCAEF GAVSTMNF----VTW HAKRQLGMPTKDL--WTPYVRNOLLT
X_HBVOR     61  CAFSSAGPCALRFTSA--RCMETTVNAPRNLPKVLHKRTLGLSTMSTTG IETVFKDCVFK
X_WMBHV     61  CAFSSAGPCALRFTSATWRCMETPMNS----VTCLRKR TLGLRTAPPTVMEQYIKDC LFE
X_HBVD3     61  CAFSSAGPCALRFTSA--RRMETTVNAHQILPKVLHKRTLGLSAMSTTDLEAVFKDC LFK

X_GSHV      115  LWEEGI IDPRLKLFV LGGCRHKYM-----
X_WHV1      115  KWEEGSIDPRLSIFV LGGCRHKCMRLP-----
X_ASHV      115  KWEEGTIDSRLPLFV LGGCRHKYM-----
X_HBVOR     119  DWEELGEEIRLKV FVLGGCRHKLVCSPAPCNFF TSA
X_WMBHV     117  QWEEOGEEPR LKV FVLGGCRHKLVTASPCIFF TSA
X_HBVD3     119  DWEELGEEIRLKV FVLGGCRHKLVCAPAPCNFF TSA

```

Figure 14: Graphic representation of multiple sequence alignment made in shade box (Expasy) of X protein from Orthohepadnaviruses using the T-Coffee tool (<http://tcoffee.crg.cat/>). GSHV: Ground Squirrel Hepatitis Virus, WHV1: Woodchuck Hepatitis Virus, ASHV: Artic Squirrel Hepatitis Virus, HBVOR: Orangutan Hepatitis B Virus, WMBV: Wolly Monkey Hepatitis B virus ,HBVD3: Hepatitis B Virus genotype D subtype ayw.

For X proteins, the amino acid sequences were downloaded from the ViralZone server. Next, we analyzed these sequences in T-Coffee simple alignment so we could obtain a FASTA file to upload in simple phylogeny, a tool from EMBL which uses Clustal algorithm with a neighbor-joining matrix (Madeira et al., 2019). Finally, we got our

phylogenetic tree in Newick format and used ITOL (Letunic & Bork, 2019) to illustrate (Figure 15).

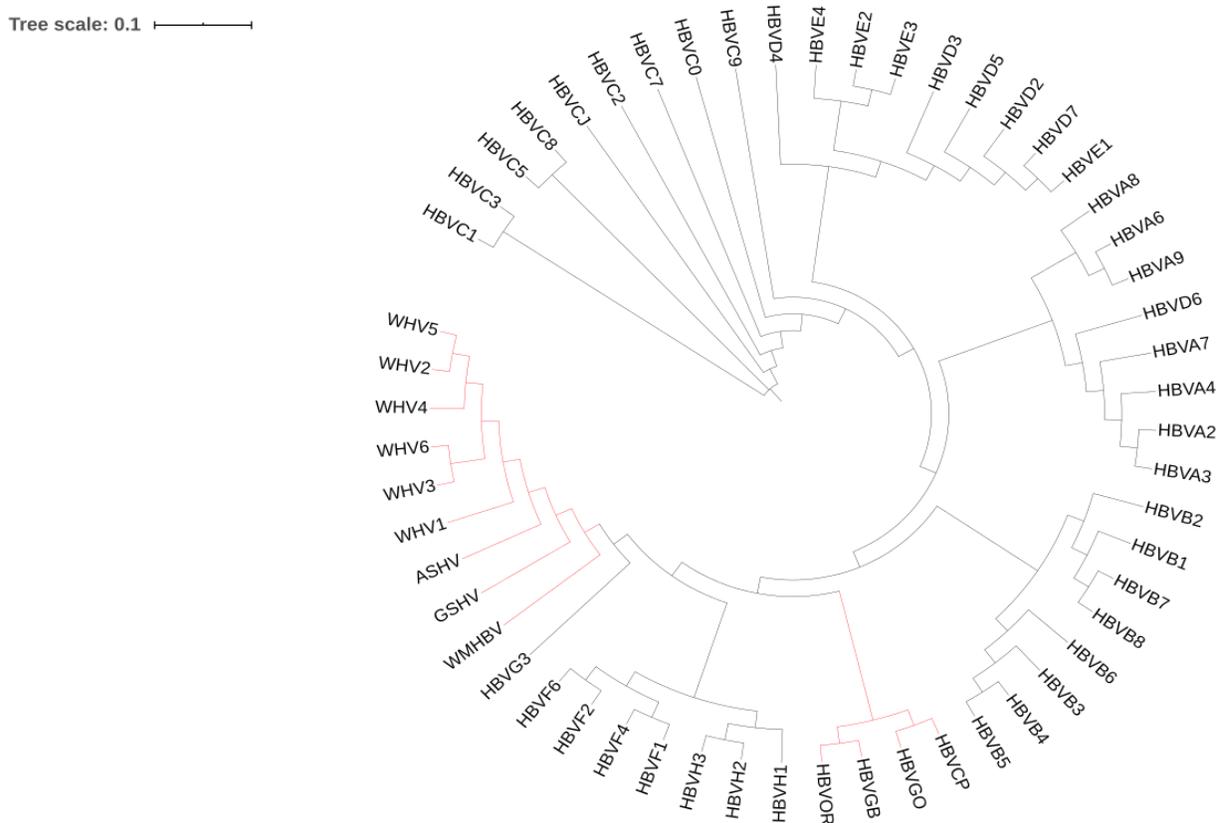


Figure 15: Unrooted phylogenetic tree of X protein from Orthohepadnavirus family made with phylogeny tool from EMBL sever and ITOL.

The BUSTED analysis confirmed that HBx gene in HBV genome is suffering selective pressure in different functional domains. The MEME analyses of HBx showed 38 sites that are suffering selective pressure, 22 of them are localized in the negative regulatory domain from HBx protein as shown in the graphic representation of LRT values (Figure 16) and the other 16 are in different functional domains (Table 8).

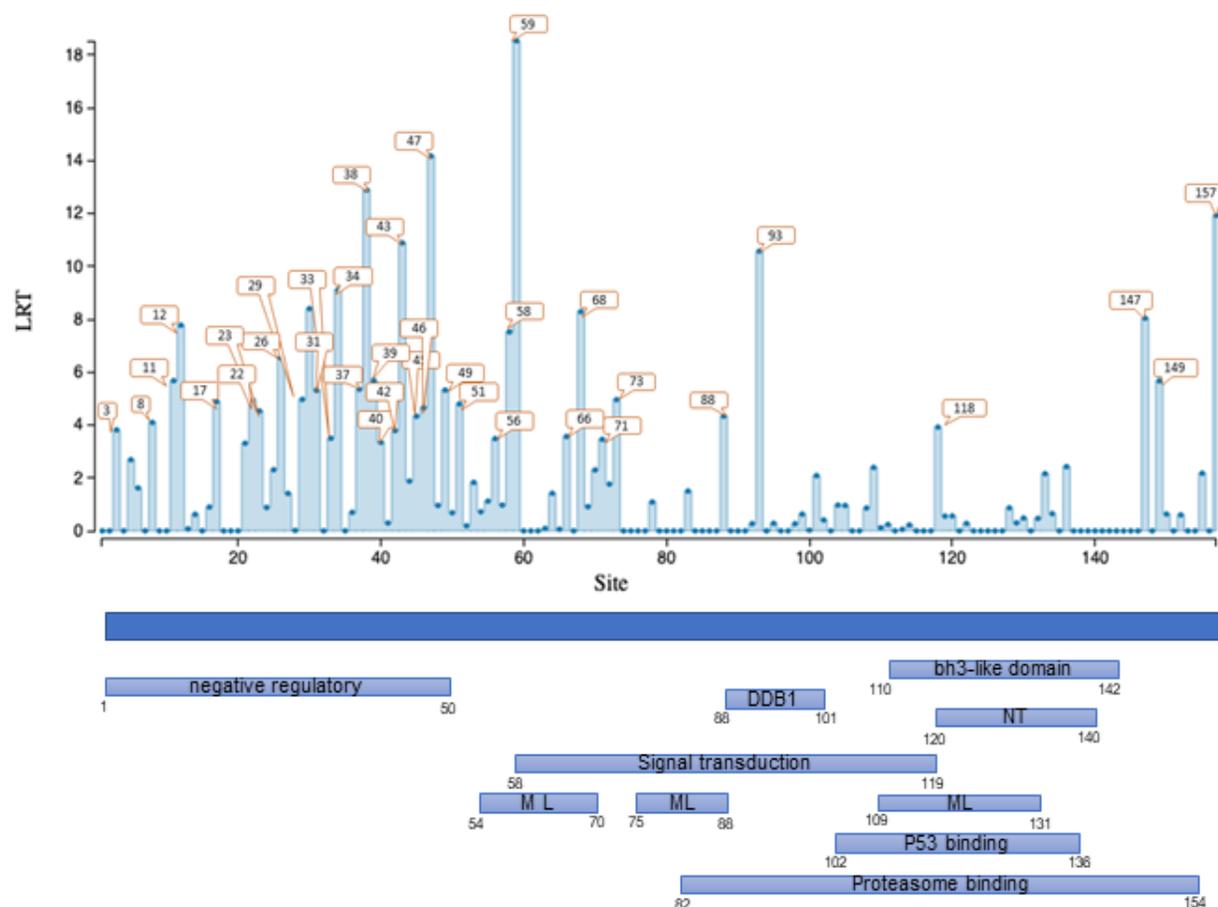


Figure 16: Likelihood Ratio Test (LRT) statistics graphic obtained with MEME for episodic diversification of X protein sequences from Orthohepadnavirus family. The numbers inside the red boxes indicate the positions of amino acid residues which showed evidence of positive selection with p -value < 0.1 . Below the LRT graph is a representation of HBx primary sequence and its domains.

Table 8: A summary of MEME results with 0.1 p -value cutoff showing the amino acid residues of X protein sequence (highlighted in Figure 16) that suffer positive selective pressure.

Site	Partition	LRT	p -value	MEME LogL	FEL LogL
3	1	3.82	0.07	-13.39	-11.96
8	1	4.10	0.06	-27.01	-25.79
11	1	5.68	0.03	-35.11	-35.11
12	1	7.78	0.01	-41.45	-41.45
17	1	4.89	0.04	-15.63	-12.67

21	1	3.31	0.09	-26.87	-26.87
22	1	4.96	0.04	-29.10	-28.49
23	1	4.54	0.05	-33.26	-31.21
26	1	6.54	0.02	-42.43	-41.37
29	1	4.97	0.04	-33.80	-33.68
30	1	8.40	0.01	-57.36	-57.36
31	1	5.31	0.03	-31.99	-31.89
33	1	3.50	0.08	-47.09	-46.97
34	1	9.12	0.00	-45.42	-44.90
37	1	5.36	0.03	-42.12	-40.34
38	1	12.90	0.00	-61.50	-61.50
39	1	5.69	0.03	-29.13	-29.13
40	1	3.35	0.09	-49.11	-49.11
42	1	3.79	0.07	-54.24	-52.81
43	1	10.89	0.00	-54.96	-54.79
45	1	4.34	0.05	-40.26	-39.95
46	1	4.65	0.05	-38.28	-38.28
47	1	14.18	0.00	-84.80	-84.78
49	1	5.33	0.03	-49.60	-49.60
51	1	4.80	0.04	-30.98	-30.83
56	1	3.49	0.08	-12.02	-10.87
58	1	7.53	0.01	-45.98	-45.98
59	1	18.53	0.00	-23.46	-13.23
66	1	3.57	0.08	-33.15	-33.02
68	1	8.29	0.01	-54.88	-54.88
71	1	3.46	0.08	-17.36	-17.36
73	1	4.96	0.04	-24.41	-23.28
88	1	4.34	0.05	-49.55	-48.57
93	1	10.58	0.00	-36.02	-33.47
118	1	3.92	0.07	-21.78	-20.09
147	1	08.04	0.01	-48.47	-48.44
149	1	5.67	0.03	-24.37	-23.21
157	1	11.93	0.00	-47.97	-41.84

The complex of Bcl-2 with HBx BH3-like motif is one of the few available structures for HBx complexes, making Bcl-2 a good control for bioinformatics analyses using MEME because of the structural information provided. BUSTED results show that Bcl-2 is suffering selective pressure and MEME analyses found 17 different sites that are suffering selective pressure, some of them in the Bcl-2-HBx interface (**Figure 17** and **Table 9**). To help in the interpretation of these results, the crystal structure of Bcl-2 complex with HBx BH3-like was used to find the amino acid residues that were identified in the Datamonkey analysis of selective pressure and compared to the interface residues (**Figure 18**). Additionally, the positions of these residues were also analyzed in a multiple sequence alignment focused on the BH3-like region of multiple X protein sequences, showing conservation in the interface residues and variation in the positively selected site (**Figure 19**). These findings could indicate that HBx is exerting selective pressure on Bcl-2, and provide a framework to analyze its interaction with HBXIP, however, this is not possible at the moment because of the lack of structural information for the HBx-HBXIP complex.

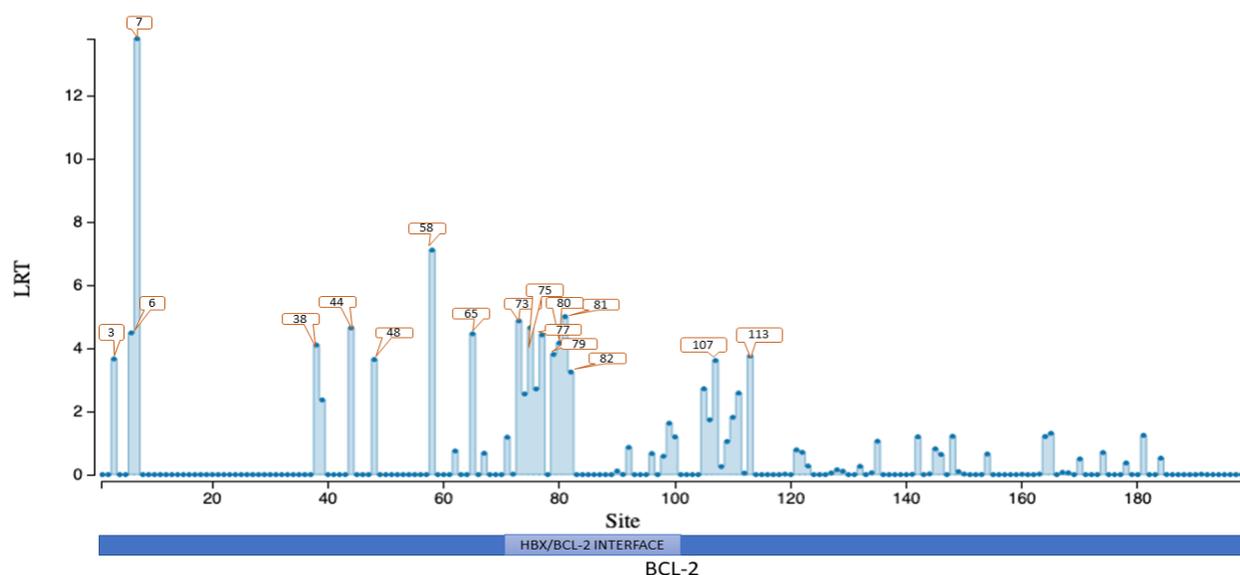


Figure 17: Likelihood Ratio Test (LRT) statistics graphic obtained with MEME for episodic diversification of BCL-2L protein sequences from primates. The numbers inside the red boxes indicate the positions of amino acid residues which showed evidence of positive selection with p -value < 0.1. Below the LRT graph is a representation of BCL2-L primary sequence and its HBx/BCL-2 interaction interface.

Table 9: A summary of MEME results with 0.1 p-value cutoff showing the amino acid residues of BCL-2 sequence (highlighted in Figure 17) that suffer positive selective pressure.

Site	Partition	LRT	p-value	MEME LogL	FEL LogL
3	1	3.67	0.08	-15.44	-12.16
6	1	4.49	0.05	-14.10	-12.18
7	1	13.83	0.00	-21.67	-14.94
38	1	4.11	0.06	-16.14	-13.88
44	1	4.66	0.05	-11.69	-9.70
48	1	3.65	0.08	-12.12	-10.36
58	1	7.12	0.01	-14.39	-11.55
65	1	4.47	0.05	-16.06	-13.66
73	1	4.88	0.04	-22.53	-19.86
75	1	4.66	0.04	-13.10	-10.00
77	1	4.43	0.05	-13.03	-9.99
79	1	3.81	0.07	-17.73	-17.47
80	1	4.18	0.06	-10.49	-9.02
81	1	05.01	0.04	-13.17	-9.87
82	1	3.25	0.09	-13.70	-11.86
107	1	3.62	0.08	-17.25	-17.25
113	1	3.76	0.07	-15.49	-15.49

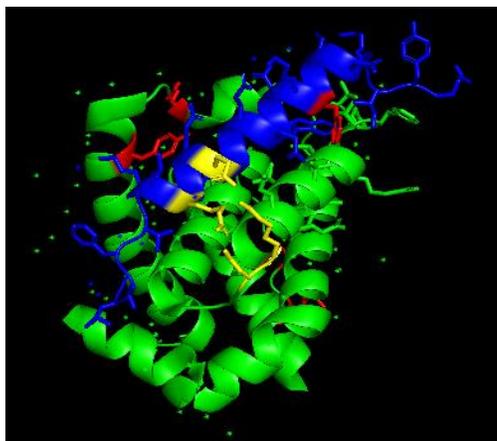


Figure 18: Graphic cartoon representation of HBx BH3-like motif (blue) in complex with Bcl-2 protein (green). The amino acid residues that are suffering selective pressure are colored in red in both proteins and the interacting amino acid residues are colored in yellow. The coordinate file of the complex was retrieved from PDB (accession number: 5fcg) and the figure was elaborated using Pymol.

```

BH3LIKE 110 -EYIKDCVFKDWEELGEEIRLKVFLGGCRHKLV 142
P03168 TAYIKDCVLLTKWEECSIDFRLSIFVLGGCRHKY-
P20975 EAYFKDCVFTWEELGEEIRLKVFLGGCRHKLV
Q67923 EAYFKDCVFNWEELGEEIRLKVFLGGCRHKLV
Q9PX75 EAYFKDCVFTWEELGEEIRLKVFLGGCRHKLV
Q69027 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
P24026 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
P03165 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
P03167 TPYIKDCVLLTKWEECSIDFRLSIFVLGGCRHKCM
Q64902 TPYVVRNQLLTKWEECSIDFRLSIFVLGGCRHKY-
P69714 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
P69713 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
P17102 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
Q91C38 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
O91531 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
Q4R189 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
Q4R181 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
P20976 EAYFKDCVFNWEELGEEIRLKVFLGGCRHKLV
P0C685 EAYFKDCVFTWEELGEEIRLKVFLGGCRHKLV
P20977 EAYFKDCVFNWEELGEEIRLKVFLGGCRHKLV
Q9PXA2 EAYFKDCVFTWEELGEEIRLKVFLGGCRHKLV
P0C678 EAYFKDCVFTWEELGEEIRLKVFLGGCRHKLV
Q9E6S8 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
P0C686 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
Q9YZR6 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
P12936 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
P0C689 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
Q913A9 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
Q81163 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
P0C687 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
P12912 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
Q9QM13 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
P0C681 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
Q67877 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
O93195 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
Q69604 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
Q80IU8 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
Q9QAX0 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
Q80IU5 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
Q05499 EAYIKDCVFKDWEELGEEIRLKVFLGGCRHKLV
Q8JMY3 EDYIKDCVFKDWEELGEEIRLKVFLGGCRHKLV
Q99HR6 EAYIKDCVFKDWEELGEEIRLKVFLGGCRHKLV
Q69607 EDYIKDCVFKDWEELGEEIRLKVFLGGCRHKLV
Q9IBI5 EKYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
P87743 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
Q9YJT2 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
Q8JMY5 EAYIKDCVFKDWEELGEEIRLKVFLGGCRHKLV
Q8JN06 EAYIKDCVFKDWEELGEEIRLKVFLGGCRHKLV
Q8JME5 EAYIKDCVFKDWEELGEEIRLKVFLGGCRHKLV
Q9J583 EAYFKDCVFKDWEELGEEIRLKVFLGGCRHKLV
P12937 TPYIKDCVLLTKWEECSIDFRLSIFVLGGCRHKCM
P12914 TPYIKDCVLLTKWEECSIDFRLSIFVLGGCRHKCM
P12913 TPYIKDCVLLTKWEECSIDFRLSIFVLGGCRHKCM
P17401 TPYIKDCVLLTKWEECSIDFRLSIFVLGGCRHKCM
P11294 TPYIKDCVLLTKWEECSIDFRLSIFVLGGCRHKCM
O71302 EAYIKDCVFKDWEELGEEIRLKVFLGGCRHKLV

```

Figure 19: Multiple sequence alignment of BH3-like motif of X protein from Orthohepadnavirus. The blue box shows Lys118, a site that suffers positive selective pressure according to MEME analysis, and the red boxes highlight interacting sites with Bcl-2 human protein.

4.3. Bioinformatic analyses of HBXIP isoforms

In spite of the numerous publications about HBXIP structure and function, the current knowledge about HBXIP isoforms is very limited and all these publications are limited to only two isoforms (Long and Short), sometimes even ignoring the existence of the other isoform. To better understand the differences between these isoforms, we made a search in UniProt and Ensembl databases which revealed that there are four protein isoforms of HBXIP (**Table 10**). A multiple sequence alignment showed that the four HBXIP isoforms differ from each other in the N-terminal regions (**Figure 20**).

Table 10: HBXIP isoforms UniProt codes and amino acids length

ISOFORM	UNIPROT NAME	# AMINO ACIDS
HBXIP LONG (HBXIP_2)	A0A0C4DGV4	173
HBXIP SHORT (HBXIP_1)	O43504	91
HBXIP NEW ISOFORM 1 (HBXIP_4)	R4GMU8	73
HBXIP NEW ISOFORM 2 (HBXIP_3)	E9PLX3	90

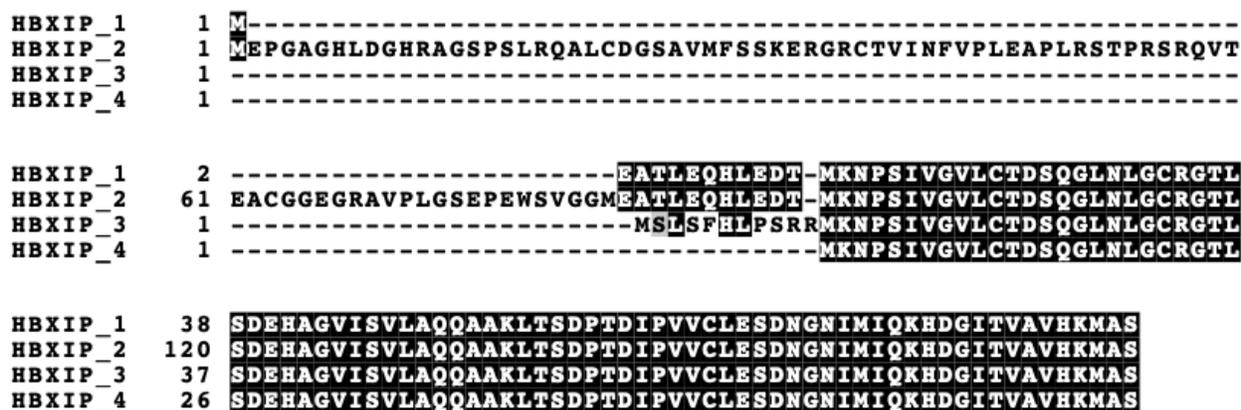


Figure 20: Multiple sequence alignment of the four HBXIP protein isoforms made in Box Shade (Expasy) from the T-Coffee tool alignment.

Next, we tried to understand the possible functional differences that these four isoforms might have. An initial attempt was to use Deeploc to predict their subcellular localizations (**Table 11**). Like many other proteins, the N-terminal sequences could harbor target sequences that specify a subcellular localization.

Table 11: Deeploc subcellular distribution prediction for human HBXIP isoforms.

HBXIP ISOFORM	HBXIP 173	HBXIP 91	HBXIP 73	HBXIP 90
N- terminal sequences prediction	0.992	0.992	0.994	0.971
PTS signal prediction	0.849	0.993	0.793	0.234
Mitochondrion/Chloroplast prediction	0.143	0.005	0.201	0.737
Perioxosome prediction	0.008	0.0	0.001	0.0
NLS prediction	0.84	0.993	0.792	0.234
Nucleus prediction	0.509	0.294	0.07	0.195
Cytoplasm prediction	0.331	0.699	0.723	0.039
Mitochondrion prediction	0.082	0.004	0.151	0.261
Chloroplast prediction	0.061	0.001	0.05	0.476
Final localization	Nucleus	Cytoplasm	Cytoplasm	Mitochondrion

The striking conservation of the Roadblock domain in all HBXIP isoforms indicated that a structural analysis could provide important insights, however, only the structure of the Short isoform (HBXIP_1) was available. With the collaboration of Breno Lisboa and Dr. Luiz Eduardo Del-Bem (UFMG), we made a structural model of the 90 aa isoform (HBXIP_3) using Modeller 9.24. This model suggested a structural conservation of HBXIP divergent N-terminus in these two short isoforms, in particular the presence of an N-terminal alpha helix in which some amino acid side chains are in similar positions in both isoforms (**Figure 21**).

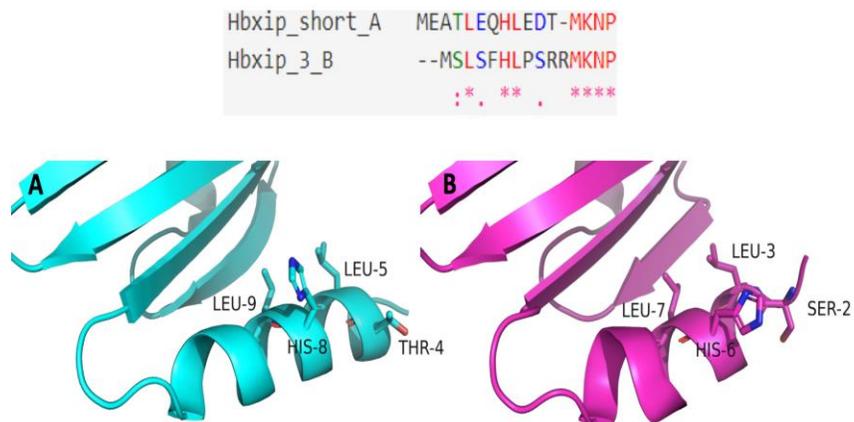


Figure 21: Comparison of the two short HBXIP isoforms: HBXIP short A (91 aa, HBXIP_1) and HBXIP 3B (90 aa, HBXIP_3). Upper panel: sequence alignment of the N-terminus of the two isoforms. Lower panel: graphic representation of the crystal structure of the 91 aa isoform in cyan (PDB: 6ehp) and a structural model of the 90 aa isoform identified in this study in magenta, highlighting the predicted N-terminal alpha-helix and the similar amino acid residues.

To understand the expression patterns of these four HBXIP isoforms in different human tissues, we used RNA-seq data available in a public database. The expression of the six transcripts was quantified using Salmon software (Patro et al., 2017), and graphed in GraphPad Prism 9 (**Figure 22**). The data was represented both as a heatmap and as bar plots. This analysis was made to find a canonic isoform of HBXIP; however, the results showed a striking variation in the different human tissues, indicating that no single isoform can be considered as canonic.

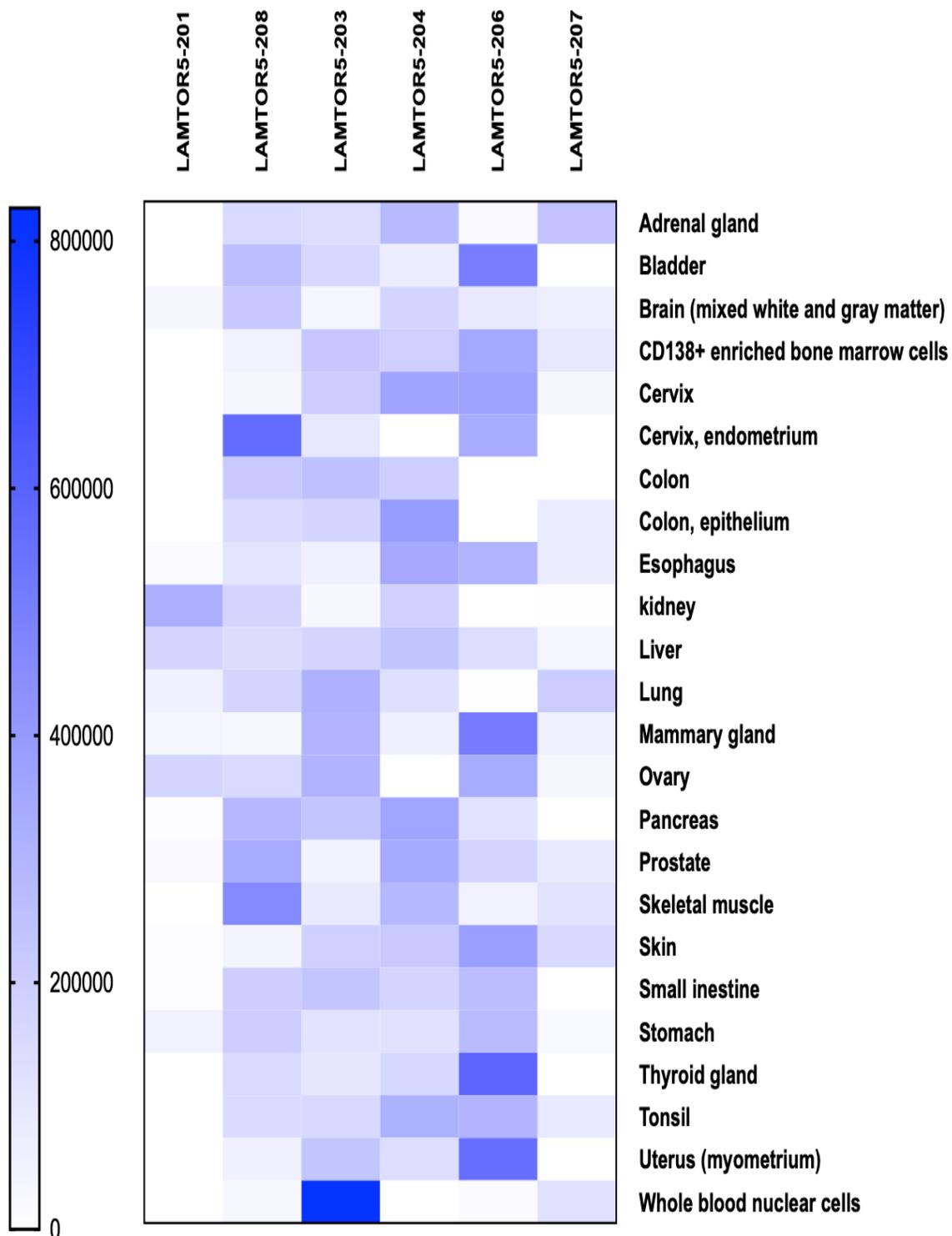
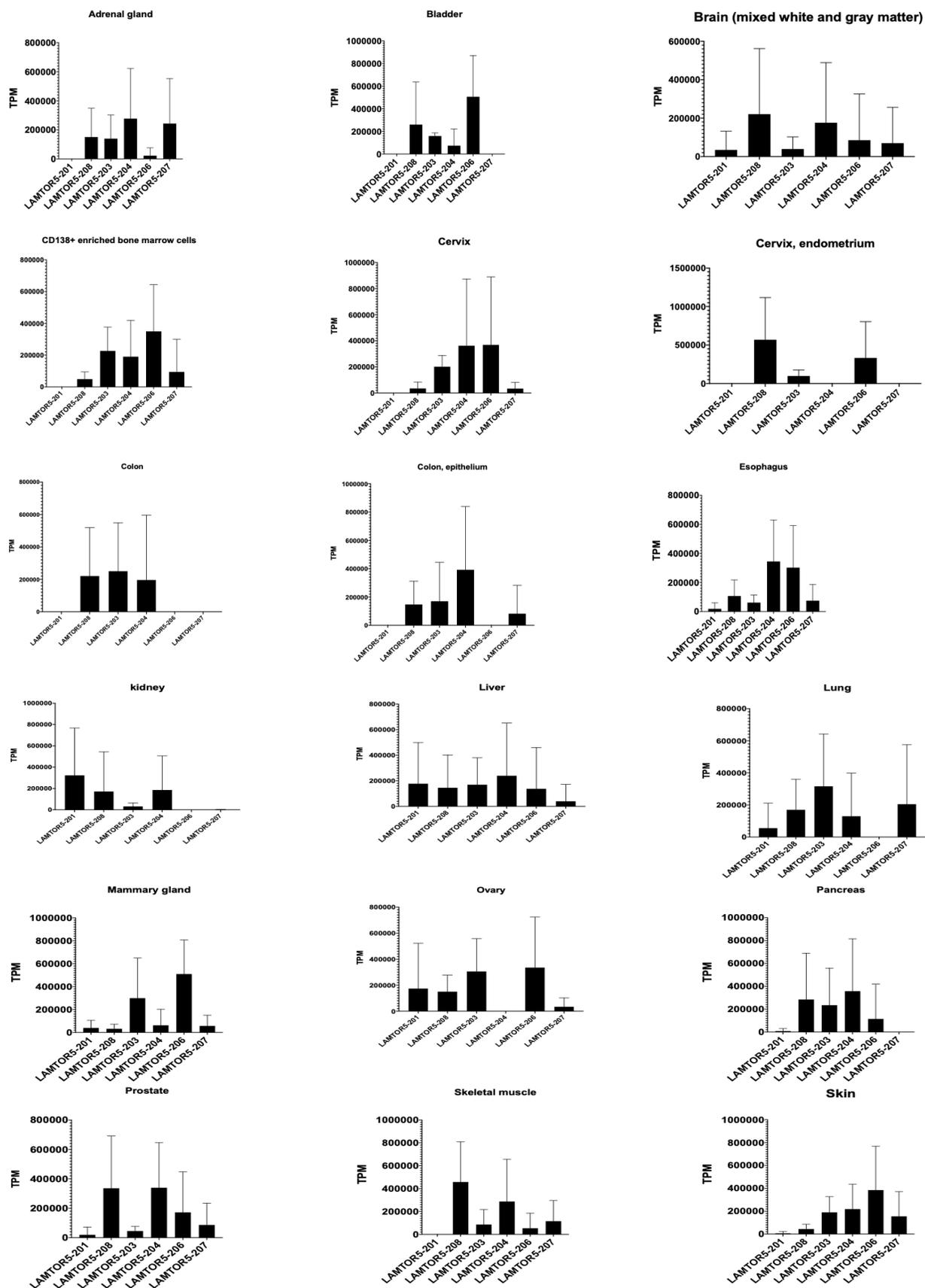


Figure 22: Heatmap graphic representation of the expression levels of HBXIP isoforms transcripts from SRA FASTA files from “Atlas of RNA sequencing profiles for normal human tissues”.



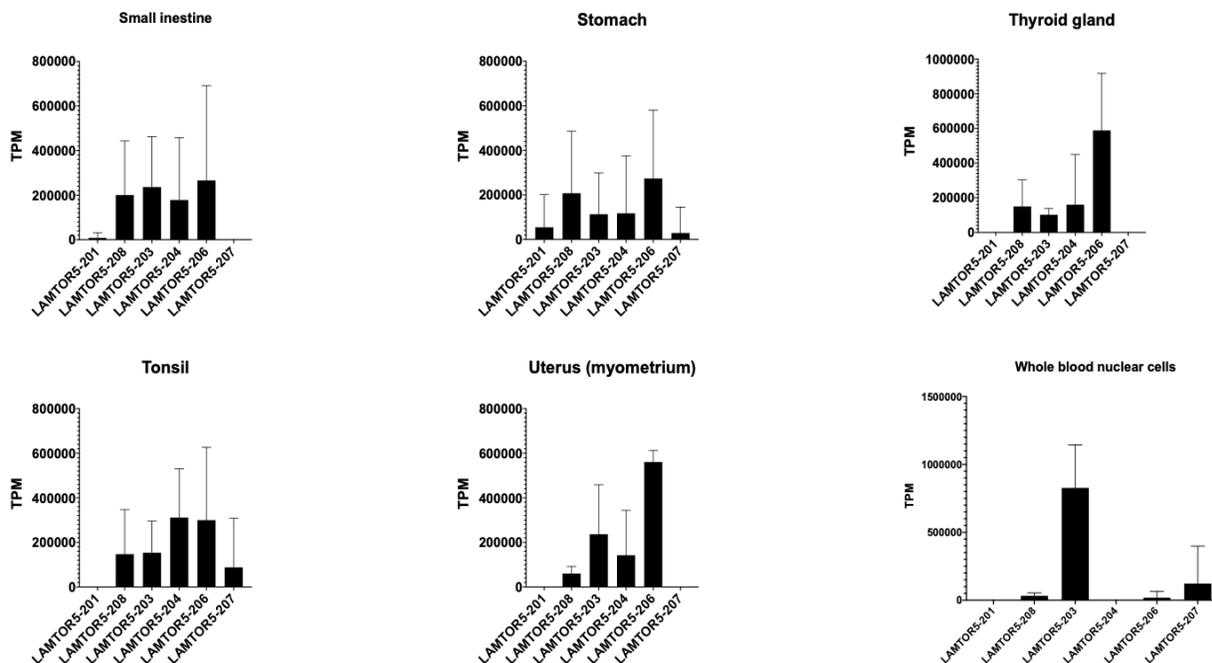


Figure 23: Quantification of HBXIP isoforms transcripts in different human tissues from SRA FASTA files from “Atlas of RNA sequencing profiles for normal human tissues”.

Finally, a complete evolutionary analysis of the four isoforms was performed with the help of Breno Lisboa and Dr. Luiz Eduardo Del-Bem. An advanced research with Blast + Hmmer found 158 hits for HBXIP short isoform and 158 hits for HBXIP long isoform. An extended evolutionary analysis identified these isoforms in different mammalian species (**Figure 1 appendix**). All these analyses indicate that the four HBXIP isoforms are expressed in human tissues, show a distinct evolutionary pattern and are likely to be functionally relevant.

4.4. Colocalization of HBx and Ragulator subunits in transfected cells

After observing HBx-GFP-HA subcellular distribution in cytoplasmic spots, the next step was to use confocal microscopy to better understand this localization. Hepatoma cell line HepG2 was transiently transfected with this construct and stained with Hoescht and Mitotracker Deep Red. HBx-GFP-HA showed a colocalization with cytosolic Mitotracker positive structures in HepG2 (**Figure 24**), which is consistent with previous studies of HBx subcellular localization (Clippinger & Bouchard, 2008).

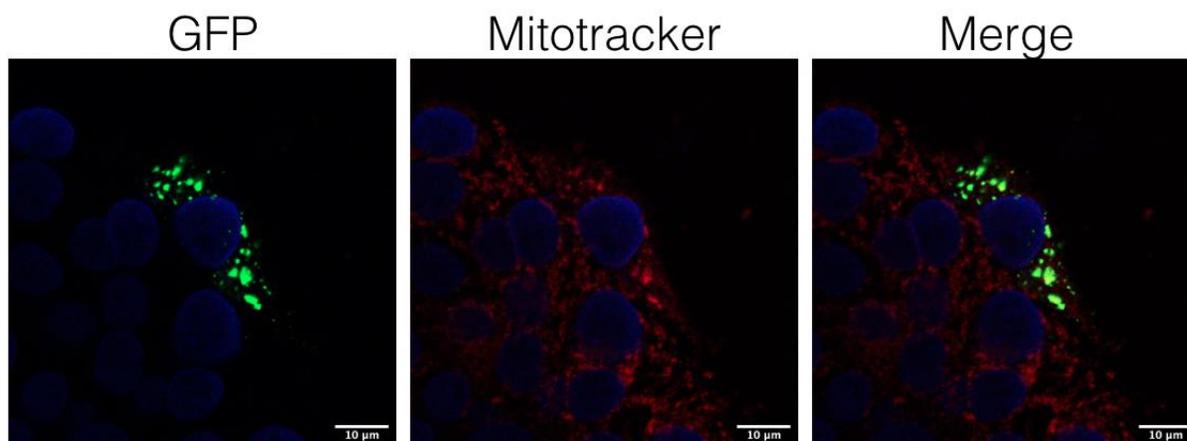


Figure 24: Mitochondrial localization of HBx in HepG2 cells. The cells were transfected with the HA-GFP-HBx WT plasmid, Mitotracker Deep Red and DAPI. The images were captured in Leica TCS SP8 confocal microscopy with 63x objective.

After confirming the mitochondrial localization of HBx-GFP-HA in HepG2, we proceeded to analyze the subcellular localization of the transfected FLAG-tagged Ragulator subunits and their possible colocalization with HBx. U2OS and HeLa cell lines were used in these experiments because of a better transfection efficiency compared to HepG2. These two cell lines are widely used in cellular imaging and in addition to being easily transfected, they have good adherence and morphology. In U2OS, all the Ragulator subunits tested (p18, p14, c7orf59 and HBXIP) showed subcellular localization patterns compatible with published literature and also with information available in the Human Protein Atlas. FLAG-p18 and FLAG-c7orf59 were found in cytoplasmic spots (probably lysosomes), while FLAG-p14 and FLAG-HBXIP Long were found both in the nucleus and cytoplasm. Importantly, neither of them colocalized with Mitotracker positive structures (**Figure 24**).

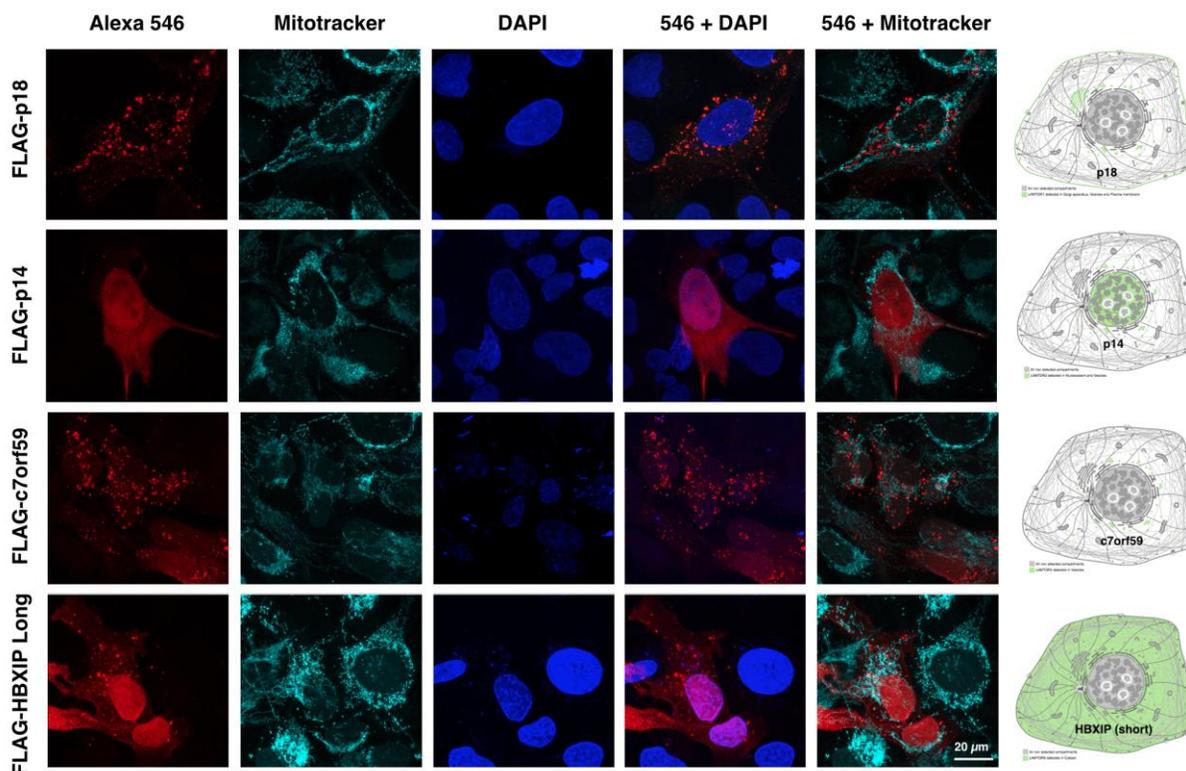


Figure 25: Immunolocalization (anti-FLAG) of transiently expressed Ragulator subunits (p18, p14, C7orf59 and HBXIP Long) in U2OS cells. The images were captured in Leica TCS SP8 confocal microscopy with 63x objective. The expected location of each subunit, based on the Human Protein Atlas database (<https://www.proteinatlas.org/>), is shown on the right.

However, when the FLAG-tagged Ragulator subunits were coexpressed with HBx-GFP-HA construct in U2OS, p14, c7orf59 and HBXIP Long showed a striking redistribution, colocalizing with HBx-GFP-HA in the Mitotracker positive cytoplasmic structures. While p14 and HBXIP Long showed only partial colocalization, c7orf59 was mostly redistributed and completely colocalized with HBx-GFP-HA in a perinuclear pattern which is characteristic of mitophagy (Vives-bauza et al., 2009). Only p18 didn't colocalize with HBx-GFP-HA. The redistribution of these subunits can be attributed to the presence of HBx by comparing the cotransfected cells with the single transfection pattern shown in **Figure 25** as well as with those cells in the cotransfection experiment which only expressed the FLAG-tagged construct. For example, comparing two adjacent cells expressing FLAG-p14, only the cell which also expresses HBx-GFP-HA presents FLAG-p14 in bright Mitotracker-positive cytoplasmic spots (**Figure 26**).

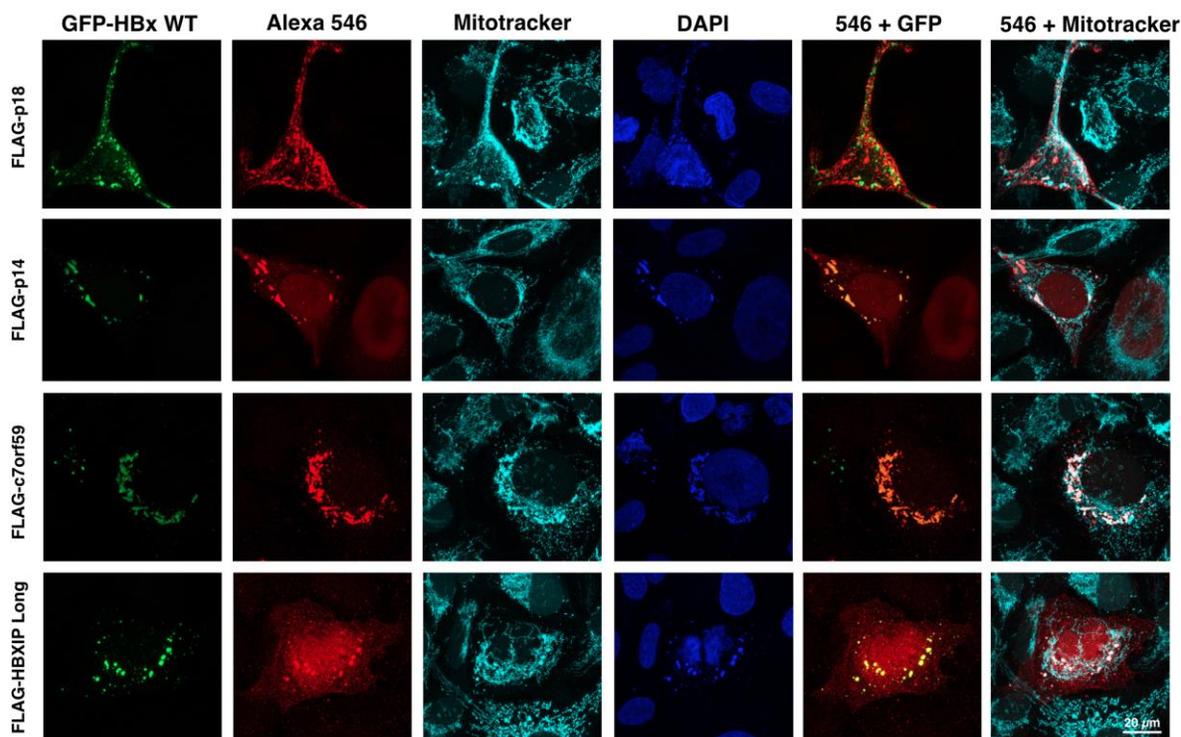


Figure 26: Ragulator subunits (p18, p14, C7orf59 e HBXIP Long) colocalization with HA- GFP-HBx WT in U2OS transfected cells. The Ragulator subunits were detected by immunolocalization (anti-FLAG, Alexa 546 secondary antibody). The images were acquired in a Leica TCS SP8 confocal microscope with 63x objective.

To further confirm that HBx is capable of changing the subcellular distribution of Ragulator subunits (p14, c7orf59 and HBXIP Long), the next panel shows the results of cotransfection of Ragulator subunits with HBx NESM-GFP-HA (Nuclear Export Sequence Motif mutation) instead of the WT HBx construct. In this mutant, the nuclear export sequence is defective, which results in partial accumulation of HBx in the nucleus (Forgues et al., 2001). Strikingly, p14, c7orf59 and HBXIP Long all changed their subcellular localizations to mirror that of the HBx mutant. Again, p18 did not colocalize with HBx and Mitotracker positive structures (**Figure 27**).

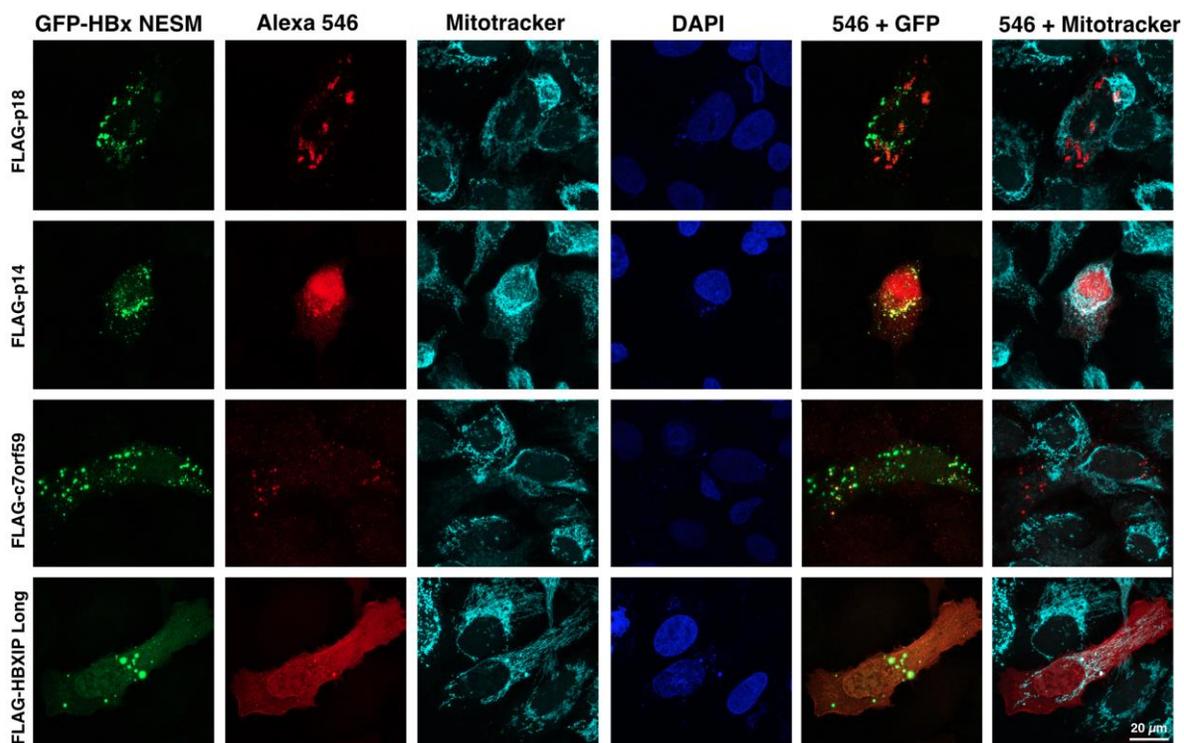


Figure 27: Ragulator subunits (p18, p14, C7orf59 and HBXIP Long) localization with HA- GFP-HBx NESM in U2OS transfected cells. The Ragulator subunits were detected by immunolocalization (anti-FLAG, Alexa 546 secondary antibody). The images were acquired in a Leica TCS SP8 confocal microscope with 63x objective.

The same experiments shown in **Figures 25 – 27** were also observed in a Leica DM6 microscope using a 20x objective to allow the visualization of more transfected cells in a single field. In this microscope, we found that the expression of HBXIP Long apparently resulted in a reduction in the expression of HBx-GFP-HA. This phenotype is not observed with the other Ragulator subunits (**Figure 28**).

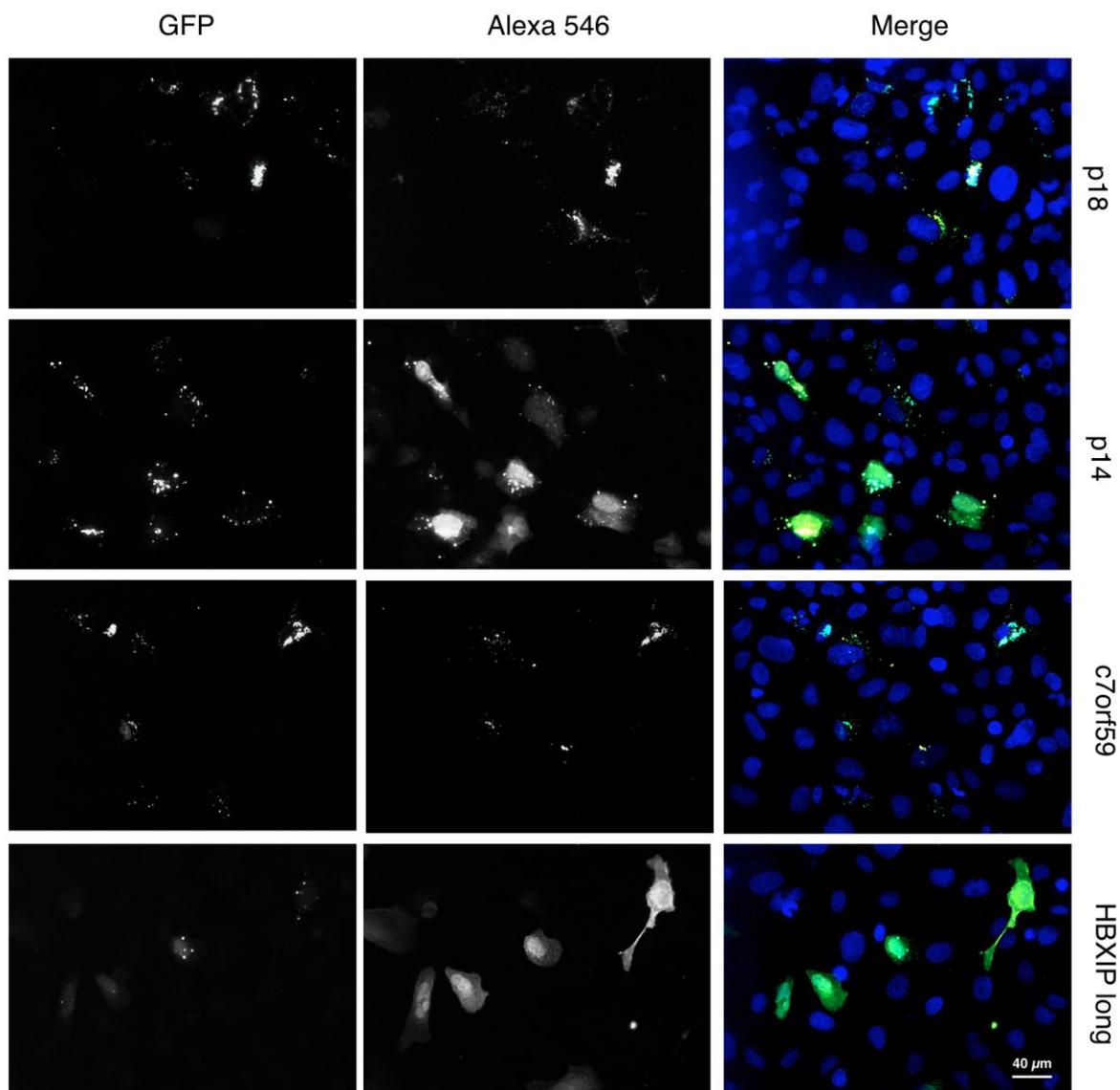


Figure 28: HBXIP Long has a negative effect on GFP-HA-HBx expression. U2OS cells were cotransfected with GFP-HA-HBx and Ragulator subunits with FLAG tag (p18, p14, c7orf59 and HBXIP Long). The Ragulator subunits were detected by immunofluorescence using primary anti-FLAG and secondary antibody conjugated to Alexa 546. The images were obtained using the Leica DM6 (non-confocal) fluorescence microscope. There is a decrease in the intensity of GFP-HA-HBx (identified as “GFP”, lower left image) and a change in its location pattern in cotransfection with FLAG-HBXIP Long, but not with the other Ragulator subunits.

We also cotransfected the Regulator subunits with HBx in HepG2 cell line, a widely used HBV infection model. In these cells, p14, c7orf59 and HBXIP also colocalized with HBx-GFP-HA construct and Mitotracker positive structures (**Figure 29**).

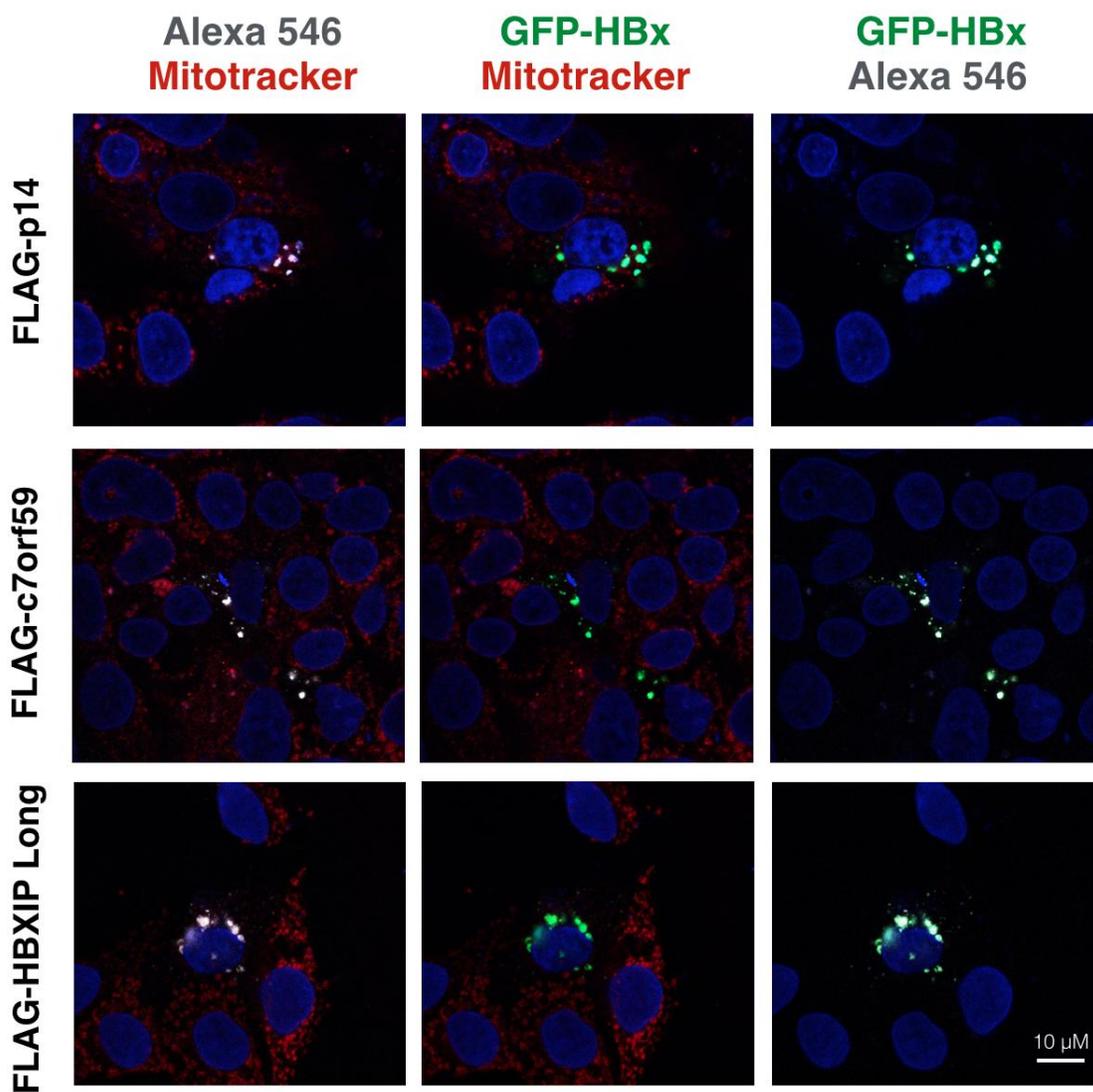


Figure 29: Colocalization of HBx with Regulator subunits in HepG2 cells. The images were obtained in confocal microscope Leica TCS SP8 with 63x objective.

4.5. Results of interaction experiments

A transient transfection of HBx-GFP-HA and Ragulator subunits (FLAG-p18, FLAG-p14, FLAG-c7orf59 and FLAG-HBXIP) was performed in HEK293-T cells to verify the expression of these proteins and their interaction. FLAG-tagged Ragulator subunits were immunoprecipitated with beads of FLAG-agarose. FLAG-p14, FLAG-C7orf59 and FLAG-HBXIP Long isoform were detected in cell extracts and in the immunoprecipitation process, but FLAG-p18 was not detected neither in immunoprecipitation or cell extract. HBx-GFP-HA was strongly detected in all the lysates but only in the immunoprecipitates of FLAG-p14, FLAG-c7orf59 and FLAG-HBXIP Long, indicating that HBx interacts with these three Ragulator subunits. No interaction was detected in the negative control (empty vector). This experiment was performed twice with identical results (**Figure 30**).

The absence of FLAG-p18 detection in this experiment might indicate either a problem of expression level or solubility of this construct. The same construct is expressed in high levels in other cell lines, as shown in the immunolocalization experiments. This observation, in addition to the fact that p18 is known to anchor on the lysosomal surface, indicates that FLAG-p18 is probably insoluble in our lysis buffer. Even with no detectable expression of FLAG-p18, a weak band of HBx-GFP-HA was detected in FLAG-p18 immunoprecipitate.

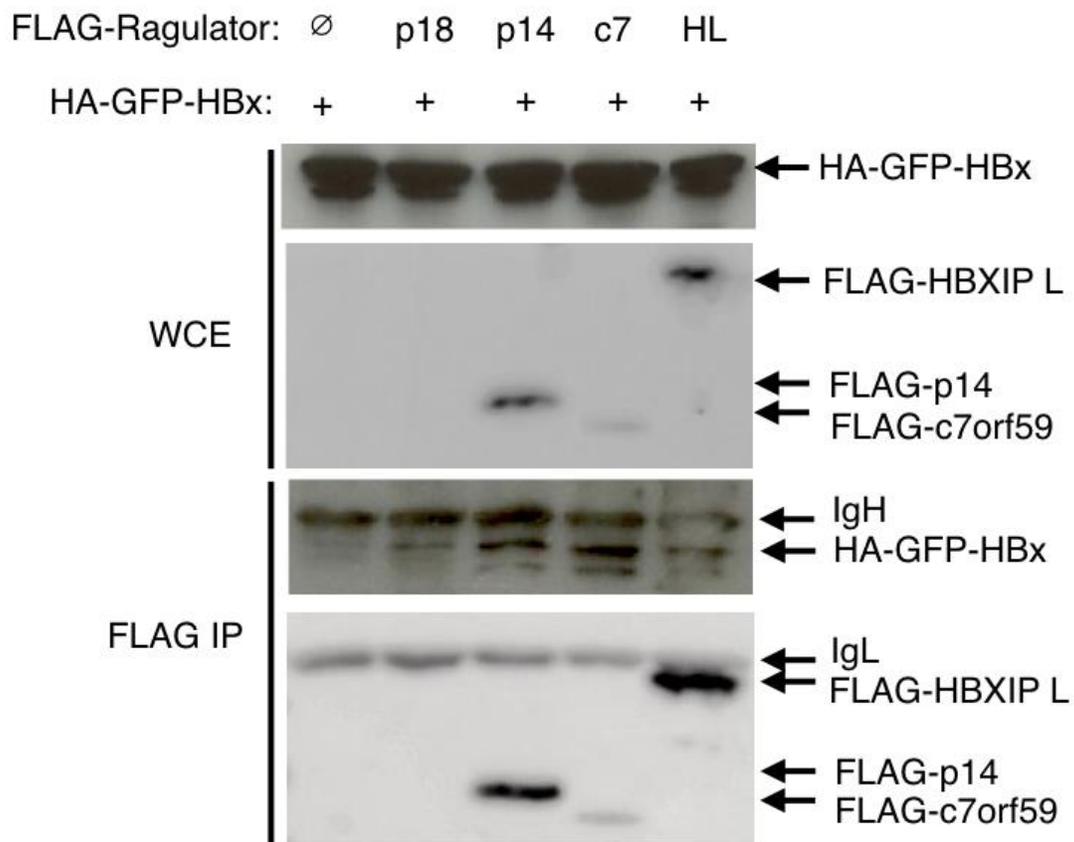


Figure 30: Co-immunoprecipitation of HBx with Ragulator subunits in transiently transfected HEK293-T cells. The cells were collected after 24 h of transfection, and the total extracts were incubated with FLAG-agarose beads to immunoprecipitate the Ragulator subunits, which contain the FLAG tag at the N-terminus. The extracts (WCE) and immunoprecipitation (FLAG IP) were analyzed by Western blot with anti-HA antibodies, to detect the HA-GFP-HBx protein, and anti-FLAG, to detect the Ragulator subunits. The presence of HA-GFP-HBx in samples immunoprecipitated with FLAG indicates interaction with the respective Ragulator subunits. C7: c7orf59, HL: HBXIP Long, ∅: empty vector (negative control).

4.6. Live cell mitophagy experiments

The interaction of HBx with Ragulator subunits and their colocalization in mitochondria, particularly when showing a perinuclear pattern characteristic of mitophagy (Vives-bauza et al., 2009), indicate that these proteins might function together in the regulation of mitophagy. To further investigate this, we used a live cell mitophagy assay. Parameters like time, concentration of mitophagy initiating compound and cell type were defined in an exploratory experiment, in which Huh-7, Vero and HEK293-T cells were tested. FLAG-tagged Ragulator subunits (p18, p14, c7orf59, HBXIP Long isoform) were cotransfected either with HA-GFP-HBx or GFP-Parkin. Huh-7 cells were selected for the final experiment, in which the cells were exposed to CCCP 24 hours after transfection to stimulate mitochondrial depolarization and the mitophagic process was followed in real time for six hours using the Operetta microscope. After that, the cells were fixed and the FLAG tag was detected with immunofluorescence to identify the cells transfected with each Ragulator subunit. Finally, the images were exported to the Columbus server (Perkin Elmer) and analyzed (**Figures 31, 32, 33 and 34**).

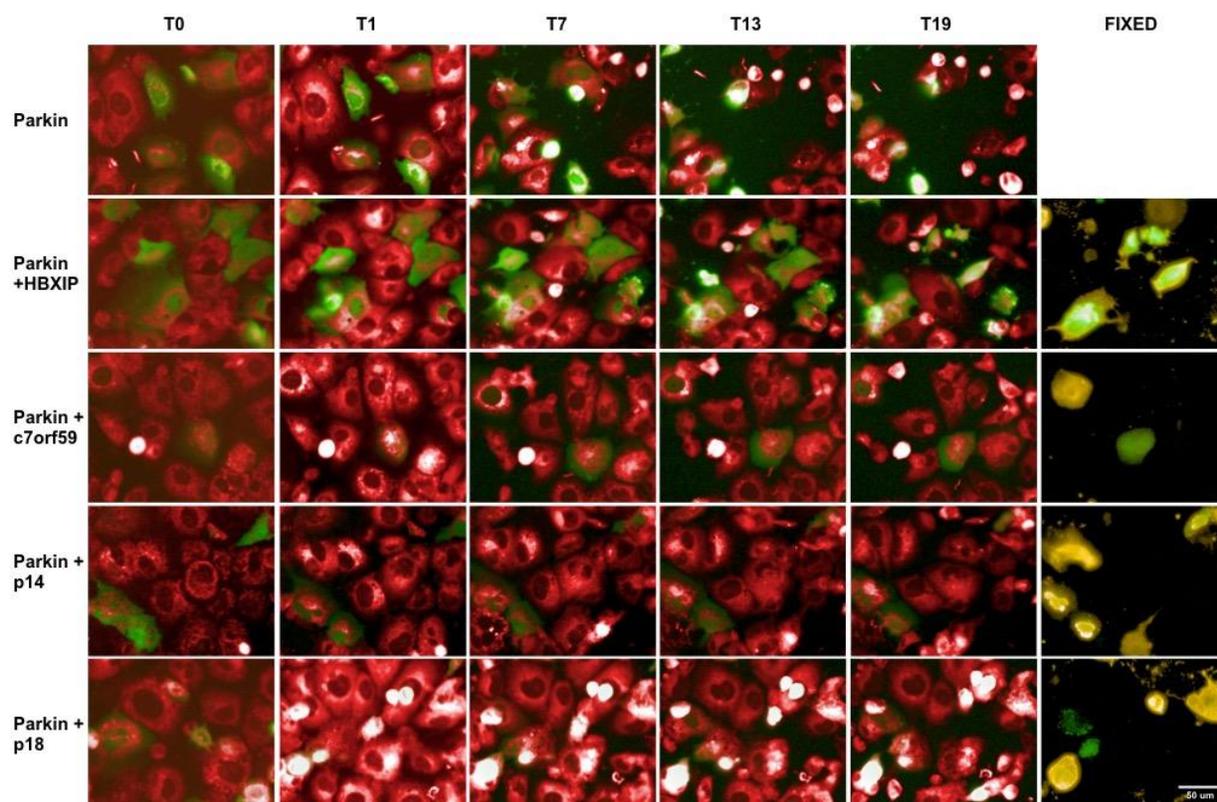


Figure 31: Regulator subunits (p18, p14, C7orf59 and HBXIP Long) co-transfected with Parkin GFP construction in Huh-7 cells. The Regulator subunits were detected by immunolocalization (anti-FLAG, Alexa 546 secondary antibody). The images were acquired in Operetta microscope with 20x objective for 6 hours after the CCCP treatment.

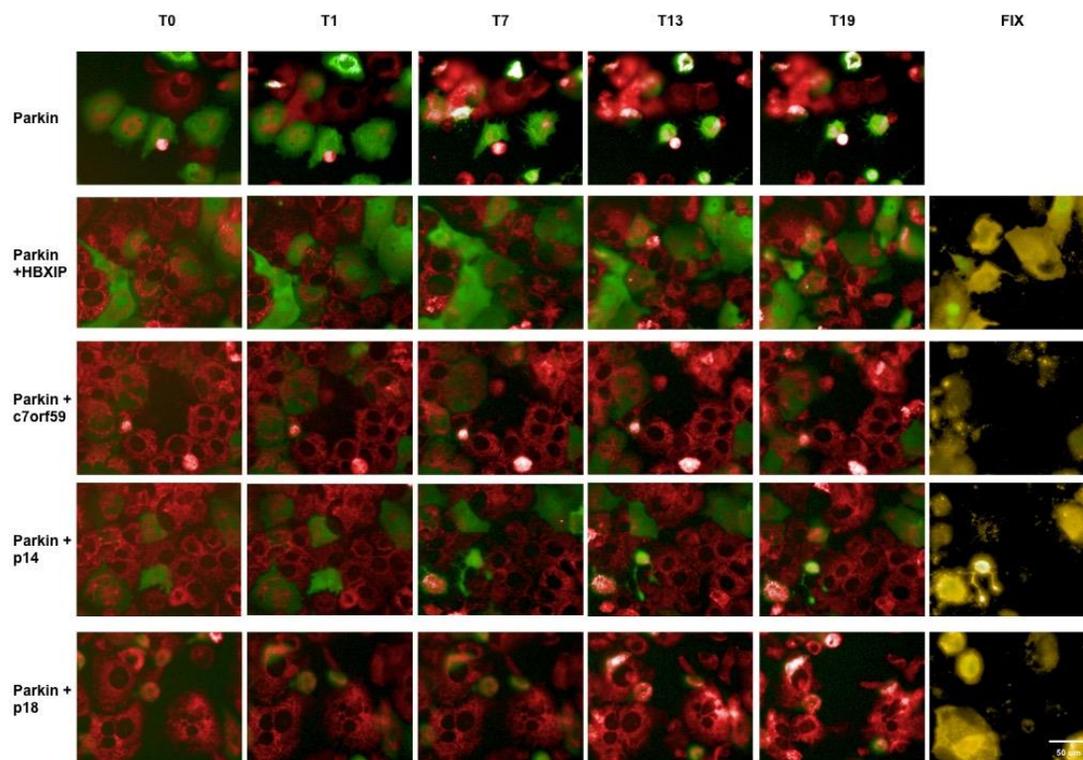


Figure 32: Regulator subunits (p18, p14, C7orf59 and HBXIP Long) co-transfected with Parkin GFP construction in Huh-7 cells. The Regulator subunits were detected by immunolocalization (anti-FLAG, Alexa 546 secondary antibody). The images were acquired in Operetta microscope with 20x objective for 6 hours without CCCP treatment.

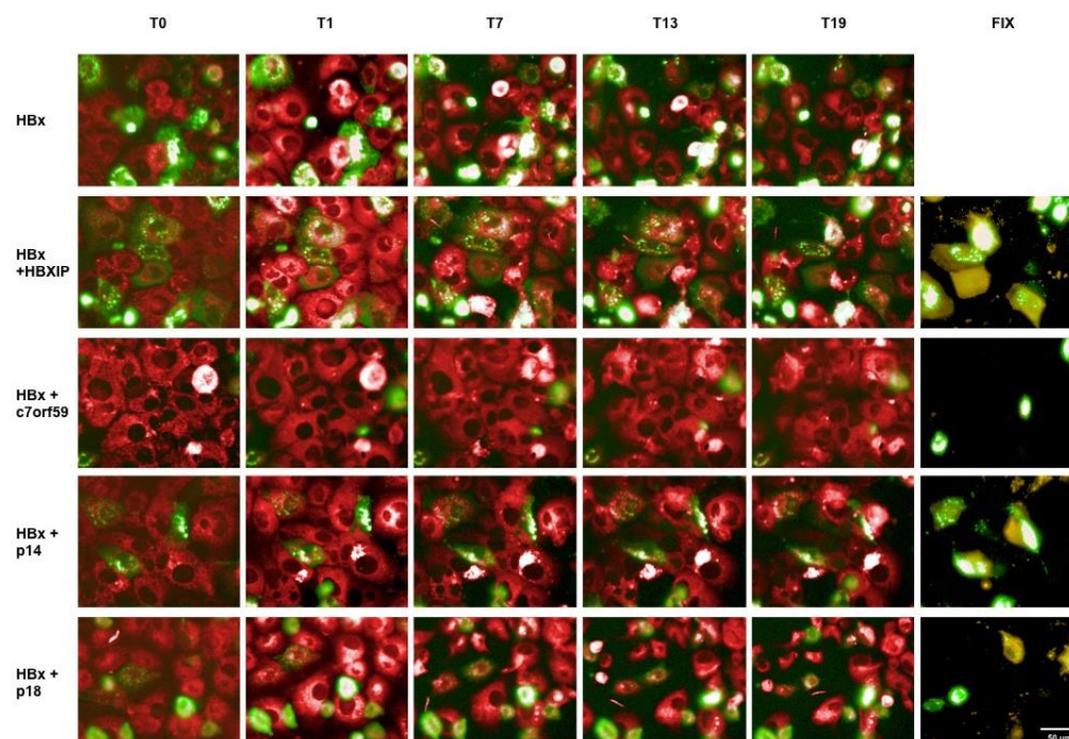


Figure 33: Ragulator subunits (p18, p14, c7orf59 and HBXIP Long) co-transfected with HBx-GFP construction in Huh-7 cells. The Ragulator subunits were detected by immunolocalization (anti-FLAG, Alexa 546 secondary antibody). The images were acquired in Operetta microscope with 20x objective for 6 hours after the CCCP treatment.

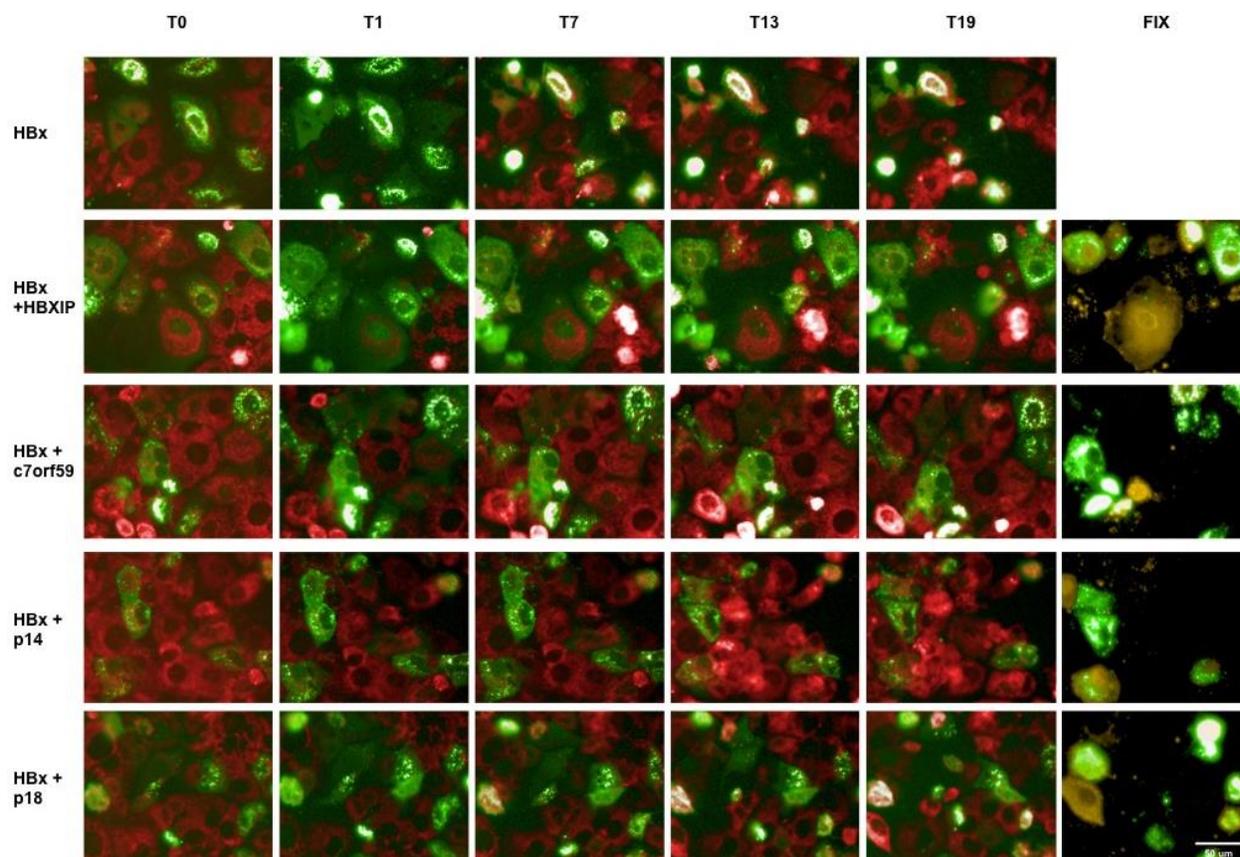


Figure 34: Ragulator subunits (p18, p14, C7orf59 and HBXIP Long) co-transfected with HBx-GFP construction in Huh-7 cells. The Ragulator subunits were detected by immunolocalization (anti-FLAG, Alexa 546 secondary antibody). The images were acquired in Operetta microscope with 20x objective for 6 hours without CCCP treatment.

The results of this experiment showed that HBx-GFP-HA transfected cells do not respond to the CCCP treatment. Also, cells that super expressed Parkin-GFP construction show an apoptosis phenotype. The cells cotransfected with Parkin-GFP and c7orf59 or p14 do not show this apoptosis phenotype, showing a protective effect to Parkin super expression and CCCP treatment. HBXIP Long isoform also showed a protective effect to the CCCP treatment or Parkin super expression, but it was different from p14 and c7orf59 protection because cells cotransfected with Parkin-GFP and HBXIP Long isoform have a colocalization of Mitotracker with Parkin-GFP construction..

5. Discussion

5.1. Phylogenetic and evolutionary analysis of HBXIP and its relationship with HBx

Although HBx was the first identified interactor of HBXIP (Melegari et al., 1998), there is not much information about this interaction in the literature and little progress was made in recent years. The objective of this investigation was to clarify the evolutionary and functional relationship of these two proteins and their effects in the cell, to help in the understanding of the multifunctional viral protein HBx.

Our primary clue about HBx and HBXIP evolutionary relationship was the finding that the Long HBXIP isoform is present mainly in mammals and rodents; coincidentally, these are also the groups of mammals affected by the HBV viral infection (Bonvicino et al., 2014). This evidence suggested that the appearance of the Long isoform in primates and rodents is connected to a virus-host relationship. The Long HBXIP isoform also appears sporadically in other species of mammals, indicating that it might have appeared several times during mammalian evolution.

The Long isoform of HBXIP originates from an alternative transcription initiation site. In selected mammalian species predicted to have only the HBXIP short isoform in our analysis, such as dogs, cats, whales, and cow, we found several in-frame stop codons in the genomic region that would encode the N-terminus of HBXIP Long confirming the impossibility to express this HBXIP isoforms from these genomes.

The traced evolutionary history of the genes encoding Regulator complex subunits (*Lamtor 1 – 4*) indicates that these subunits appear to be highly conserved in primates and didn't experience diversifying selection. The only exception is the *Lamtor5* gene, which shows some signatures of adaptive evolution in primates and rodents in the N-terminal regions which are exclusive of the HBXIP Long isoform. These include several sites under positive selection, which could be reminiscent of a virus-host evolutionary arms race with its interactor partner and viral protein HBx (Grossman et al., 2010).

Bcl-2 and DDB1 were used as positive controls in our evolutionary analysis because they are known HBx interactors and the crystal structures of these complexes are available

(Geng et al., 2012; Li et al., 2010; Zhang et al., 2019). The BUSTED analysis of Bcl-2 and DDB1 also detected signs of selective pressure, these results confirm that our approach is valid to detect signs of selective pressure in human proteins that interact with HBx or its orthologues.

The selective pressure signs detected in the *Lamtor5* gene were further investigated using MEME, which found significant signs of positive selection in five amino acid residues of HBXIP, all of which were located in the N-terminal region exclusive of HBXIP Long isoform. This indicates that these amino acids might be important for the interaction between HBx and HBXIP Long, but it doesn't mean that these sites are the direct interaction sites between HBXIP with HBx (Daugherty & Malik, 2012).

X protein nucleotide sequences from Orthohepadnaviruses were analyzed with MEME, which found many sites with signs of positive selective pressure, this may be attributed to the multiple interactors of HBx and its orthologues. The negative regulatory motif of HBx located in its N-terminal region harbors most of these sites. Mutations in this region are known to repress transactivation of X protein (Murakami et al., 1994), indicating that genetic variation in this motif could affect viral replication by controlling HBx transcriptional activity. Another interesting site which was found to suffer positive selection in HBx is lysine 118, which is located in the BH3-like motif, close to other residues (Gly124 and Ile127) that were shown to abolish HBx interaction with Bcl-2 when mutated (Geng et al., 2012)

The analysis of HBXIP isoforms culminated in the discovery of two additional isoforms in sequence databases (Uniprot and Ensembl). To find experimental evidence of their expression, we used evolutionary analysis and searched for their respective transcripts in RNA-seq database. A phylogenetic tree made by Breno Lisboa and Dr Luiz Del-Bem (**appendix**) showed the conservation these four isoforms in mammals. Two of these isoforms (HBXIP_1 and HBXIP_3) show a very similar size with 91 and 90 amino acid residues, respectively. This makes it difficult to distinguish them with the trivial protein analysis techniques like western blot or SDS-PAGE. The only technique that has the capacity to differentiate them is mass spectrometry, such evidence is found for HBXIP_3 isoform in the proteomics DB database (Mertins et al., 2016; Phanstiel et al., 2011).

The four HBXIP isoforms have different subcellular localization predictions. We were able to confirm the nuclear localization of HBXIP Long isoform (173 aa) with immunofluorescence techniques. Like many other proteins, the different N-termini of each isoform could give different functions for each protein but the essential function is associated to the Roadblock domain that is very conserved in all four isoforms and is involved in interactions with other Ragulator subunits. The RNA-seq analyses did not find a canonic isoform but revealed some important information. For example, kidney, liver and skeletal muscle tissues have comparable amounts of each short isoform (HBXIP_1 and HBXIP_3) and in nuclear blood cells there is a predominance of HBXIP_3 isoform. Endometrium expresses predominantly the Long isoform. These findings indicate that HBXIP regulation in different tissues might be more complex than anticipated.

5.2. Interaction and colocalization between HBx and the Ragulator subunits

We used confocal microscopy to study the subcellular distributions of HBx and Ragulator subunits and how they could affect each other. In the absence of HBx expression, HBXIP Long isoform shows a predominantly nuclear but also cytoplasmic localization in transfected U2OS, HeLa and Huh-7 cells, while endogenous HBXIP in Human Protein Atlas localized in the cytoplasm. Other Ragulator subunits also showed subcellular distributions in these cell lines compatible with literature and Human Protein Atlas information. HBx-GFP-HA also behaved as expected, localizing in the mitochondria.

When the Ragulator subunits were coexpressed with HBx-GFP-HA, their subcellular localizations changed drastically, with the exception of p18, which does not seem to colocalized with HBx in the mitochondria. It was clear that c7orf59 has a better colocalization with HBx than the other Ragulator subunits and localizes to mitochondria in perinuclear pattern which is a sign of mitophagy. This process is essential to viral replication and is controlled by HBx protein.

The expression of HBx-GFP-HA showed a decrease in the presence of HBXIP Long isoform but not with other Ragulator subunits. Thus, our data support the idea that HBXIP Long isoform acts as a HBx antagonist, which agrees with observations such as the negative regulation of HBx activity and the replication cycle of HBV by HBXIP Long

(Melegari et al., 1998) and the fact that miR-501 induces HBV replication partly through negative regulation of HBXIP Long (Jin et al., 2013).

HBXIP covers many of the conditions for the antagonism of the HBx protein: the short isoform is very conserved in the evolution of the vertebrates, but only mammals, specifically rodents and primates, have the Long isoform. In the HBXIP evolutionary analyses with BUSTED and MEME, only the Long isoform showed signs of positive selective pressure. These pieces of evidence can be a perfect example of the Red Queen hypothesis for a viral antagonism of host-pathogen interaction.

We do not have the HBXIP short isoform to see if there is a difference in the phenotypic effects between both isoforms, but we hypothesized it will have a phenotypic behavior like C7orf59, because both of them have a conserved Roadblock domain and form a heterodimer in the Ragulator complex. Considering these the behavior from both isoforms will have a significantly phenotypic difference for their relationship with HBx. The colocalization results repeat in HepG2 cells for a model more near to the physiological condition of HBV infection. The DAPI bodies that are colocalized with Mitotracker, HBx, and Ragulator subunits could be mitochondrial DNA, but further studies need to be made to clarify this information.

The immunoprecipitation experiment confirmed the interaction of HBx-GFP-HA with p14, C7orf59 and HBXIP Long isoform. There was a clear preference of HBx for C7orf59 in both colocalization and in immunoprecipitation results. The c7orf59 protein has an important role in the nucleation of the Ragulator complex (Rasheed et al., 2019). No colocalization or interaction of HBx was detected with p18, however, in the immunoprecipitation experiment we could not detect FLAG-p18 expression in the lysate. There are some interesting similarities between p18 and HBx, like the fact that both are unstructured proteins, their lengths are similar (161 and 154 amino acids residues, respectively), and both of them have palmytoilation sites which allow them to anchor in membranes (Lee et al., 2012; Yonehara et al., 2017). MP1 was not used in these experiments, but because it is a dimerization partner of p14, the interaction of p14 with HBx indicates that HBx might also interact with MP1. In conclusion, these results indicate

that a viral version of the Ragulator might exist, in which the p14-MP1 and c7orf59-HBXIP heterodimers are surrounded by HBx protein and targeted to mitochondria, where they might function as scaffolds for HBx interactions with other proteins. However, future structural studies and HBx-p18 interaction experiments are needed to confirm this hypothesis.

5.3. Role of HBx viral protein and Ragulator subunits in mitophagy

Autophagy and HBx are important for HBV viral DNA replication (Sir et al., 2010). It is largely known that HBx affects mitochondria, especially the process of mitophagy activation (S. J. Kim et al., 2013). The same results of perinuclear localization and fragmentation of mitochondria were observed in our confocal experiment, more visible effects were seen when HBx was transfected with c7orf59 (S. Kim et al., 2007). In the live cells experiment, we used Parkin-GFP as a marker of mitophagy, but the overexpression of Parkin caused an apoptotic phenotype. Parkin overexpression is known to rescue a defective mitochondrial phenotype and to decrease apoptosis in Cockayne syndrome A cells (Pascucci et al., 2017). Also, in Parkinson disease Parkin seems to have a dual regulation of mitophagy and apoptosis via VDAC1 ubiquitination (Ham et al., 2020).

Expression of c7orf59 and p14 subunits caused a protective effect to Parkin super expression and to treatment with CCCP, a compound that stimulates mitophagy. Other cells that presented a protective phenotype where those cotransfected with Parkin and HBXIP Long isoform, but in contrast to cells transfected with c7orf59 and p14, these cells present a colocalization of Parkin-GFP with Mitotracker positive structures. Different to the other Ragulator subunits investigated, p18 had no effects on these transfected cells. There is not much investigation on the effects of CCCP on the Ragulator complex, so this new information brings a new context of Ragulator subunits function involved in mitophagy. In the cells transfected with HBx protein and treated with CCCP, there was no effect at all. This indicates that HBx might be a very strong regulator of mitophagy or that CCCP uses a different pathway to induce mitophagy.

Finally, we propose a mechanism of HBx function in which, it regulates mTORC1 pathway by sequestering Ragulator subunits and this might be necessary in autophagy and

mitophagy and to help in viral replication. Guo et al have shown that mTOR activation inhibits HBV RNA transcription and reduced HBV DNA replication via PI3K signaling pathway (Guo et al., 2007). Another study that supports our hypothesis is Teng et al 2011, which concluded that the activation of mTOR signal in GGHs may suppress HBsAg synthesis during HBV tumorigenesis (Teng et al., 2011). To validate this hypothesis and clarify the role of HBXIP Long isoform in HBx regulation, more structural and functional studies are needed.

6. Bibliography

- Afgan, E., Baker, D., Batut, B., Van Den Beek, M., Bouvier, D., Ech, M., Chilton, J., Clements, D., Coraor, N., Grüning, B. A., Guerler, A., Hillman-Jackson, J., Hiltmann, S., Jalili, V., Rasche, H., Soranzo, N., Goecks, J., Taylor, J., Nekrutenko, A., & Blankenberg, D. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Research*, *46*(W1), W537–W544. <https://doi.org/10.1093/nar/gky379>
- Ashrafi, G., & Schwarz, T. L. (2013). The pathways of mitophagy for quality control and clearance of mitochondria. *Cell Death and Differentiation*, *20*(1), 31–42. <https://doi.org/10.1038/cdd.2012.81>
- Aufiero, B., & Schneider, R. J. (1990). The hepatitis B virus X-gene product trans-activates both RNA polymerase II and III promoters. *The EMBO Journal*, *9*(2), 497–504. <https://doi.org/10.1002/j.1460-2075.1990.tb08136.x>
- Bar-Peled, L., Schweitzer, L. D., Zoncu, R., & Sabatini, D. M. (2012). Ragulator is a GEF for the rag GTPases that signal amino acid levels to mTORC1. *Cell*, *150*(6), 1196–1208. <https://doi.org/10.1016/j.cell.2012.07.032>
- Beck, J., & Nassal, M. (2007). Hepatitis B virus replication. *World Journal of Gastroenterology*, *13*(1), 48–64. <https://doi.org/10.3748/wjg.v13.i1.48>
- Benkman, C. W., & Parchman, T. L. (2009). Coevolution between crossbills and black pine: The importance of competitors, forest area and resource stability. *Journal of Evolutionary Biology*, *22*(5), 942–953. <https://doi.org/10.1111/j.1420-9101.2009.01703.x>
- Bonvicino, C. R., Moreira, M. A., & Soares, M. A. (2014). Hepatitis B virus lineages in mammalian hosts: Potential for bidirectional cross-species transmission. *World Journal of Gastroenterology*, *20*(24), 7665–7674. <https://doi.org/10.3748/wjg.v20.i24.7665>
- Bouchard, M. J., Puro, R. J., Wang, L., & Schneider, R. J. (2003). Activation and Inhibition of Cellular Calcium and Tyrosine Kinase Signaling Pathways Identify

- Targets of the HBx Protein Involved in Hepatitis B Virus Replication. *Journal of Virology*, 77(14), 7713–7719. <https://doi.org/10.1128/jvi.77.14.7713-7719.2003>
- Brockhurst, M. A., & Koskella, B. (2013). Experimental coevolution of species interactions. *Trends in Ecology and Evolution*, 28(6), 367–375. <https://doi.org/10.1016/j.tree.2013.02.009>
- Cavalier-Smith, T. (2009). Predation and eukaryote cell origins: A coevolutionary perspective. *International Journal of Biochemistry and Cell Biology*, 41(2), 307–322. <https://doi.org/10.1016/j.biocel.2008.10.002>
- Clippinger, A. J., & Bouchard, M. J. (2008). *Hepatitis B Virus HBx Protein Localizes to Mitochondria in Primary Rat Hepatocytes and Modulates Mitochondrial Membrane Potential*. 82(14), 6798–6811. <https://doi.org/10.1128/JVI.00154-08>
- Daugherty, M. D., & Malik, H. S. (2012). Rules of Engagement: Molecular Insights from Host-Virus Arms Races. *Annual Review of Genetics*, 46(1), 677–700. <https://doi.org/10.1146/annurev-genet-110711-155522>
- De Araujo, M. E. G., Naschberger, A., Fürnrohr, B. G., Stasyk, T., Dunzendorfer-Matt, T., Lechner, S., Welti, S., Kremser, L., Shivalingaiah, G., Offterdinger, M., Lindner, H. H., Huber, L. A., & Scheffzek, K. (2017). Crystal structure of the human lysosomal mTORC1 scaffold complex and its impact on signaling. *Science*, 358(6361), 377–381. <https://doi.org/10.1126/science.aao1583>
- Drexler, J. F., Geipel, A., König, A., Corman, V. M., Van Riel, D., Leijten, L. M., Bremer, C. M., Rasche, A., Cottontail, V. M., Maganga, G. D., Schlegel, M., Müller, M. A., Adam, A., Klose, S. M., Carneiro, A. J. B., Stöcker, A., Franke, C. R., Gloza-Rausch, F., Geyer, J., ... Drosten, C. (2013). Bats carry pathogenic hepadnaviruses antigenically related to hepatitis B virus and capable of infecting human hepatocytes. *Proceedings of the National Academy of Sciences of the United States of America*, 110(40), 16151–16156. <https://doi.org/10.1073/pnas.1308049110>
- Duggal, N. K., & Emerman, M. (2012). Evolutionary conflicts between viruses and

- restriction factors shape immunity. *Nature Reviews Immunology*, 12(10), 687–695. <https://doi.org/10.1038/nri3295>
- Efeyan, A., Zoncu, R., & Sabatini, D. M. (2012). Amino acids and mTORC1: From lysosomes to disease. In *Trends in Molecular Medicine* (Vol. 18, Issue 9, pp. 524–533). <https://doi.org/10.1016/j.molmed.2012.05.007>
- Emerman, M., & Malik, H. S. (2010). Paleovirology - Modern consequences of ancient viruses. *PLoS Biology*, 8(2), 1–5. <https://doi.org/10.1371/journal.pbio.1000301>
- Fei, H., Zhou, Y., Li, R., Yang, M., Ma, J., & Wang, F. (2017). HBXIP, a binding protein of HBx, regulates maintenance of the G2/M phase checkpoint induced by DNA damage and enhances sensitivity to doxorubicin-induced cytotoxicity. *Cell Cycle*, 16(5), 468–476. <https://doi.org/10.1080/15384101.2017.1281482>
- Fessler, Michael B.; Rudel, Lawrence L.; Brown, M. (2008). Premature Cell Cycle Entry Induced by Hepatitis B Virus Regulatory HBx Protein During Compensatory Liver Regeneration. *Bone*, 23(1), 1–7. <https://doi.org/10.1038/jid.2014.371>
- Fessler, M. B., Rudel, L. L., & Brown, M. (2008). The C-terminal region of the hepatitis B virus X protein is essential for its stability and function. *Bone*, 23(1), 1–7. <https://doi.org/10.1038/jid.2014.371>
- Forgues, M., Marrogi, A. J., Spillare, E. A., Wu, C., Yang, Q., Yoshida, M., & Wang, X. W. (2001). Interaction of the Hepatitis B Virus X Protein with the Crm1-dependent Nuclear Export Pathway. *Journal of Biological Chemistry*, 276(25), 22797–22803. <https://doi.org/10.1074/jbc.M101259200>
- Gearhart, T., & Bouchard, M. J. (2008). Replication of the Hepatitis B Virus Requires a Calcium- Dependent HBx-Induced G1 Phase Arrest of Hepatocytes. *Bone*, 23(1), 1–7. <https://doi.org/10.1038/jid.2014.371>
- Gearhart, T. L., & Bouchard, M. J. (2010). The Hepatitis B Virus X Protein Modulates Hepatocyte Proliferation Pathways To Stimulate Viral Replication. *Journal of Virology*, 84(6), 2675–2686. <https://doi.org/10.1128/jvi.02196-09>
- Geng, X., Harry, B. L., Zhou, Q., Skeen-Gaar, R. R., Ge, X., Lee, E. S., Mitani, S., &

- Xue, D. (2012). Hepatitis B virus X protein targets the Bcl-2 protein CED-9 to induce intracellular Ca²⁺ increase and cell death in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(45), 18465–18470. <https://doi.org/10.1073/pnas.1204652109>
- Grossman, S. R., Shylakhter, I., Karlsson, E. K., Byrne, E. H., Morales, S., Frieden, G., Hostetter, E., Angelino, E., Garber, M., Zuk, O., Lander, E. S., Schaffner, S. F., & Sabeti, P. C. (2010). A composite of multiple signals distinguishes causal variants in regions of positive selection. *Science*, *327*(5967), 883–886. <https://doi.org/10.1126/science.1183863>
- Guo, H., Zhou, T., Jiang, D., Cuconati, A., Xiao, G.-H., Block, T. M., & Guo, J.-T. (2007). Regulation of Hepatitis B Virus Replication by the Phosphatidylinositol 3-Kinase-Akt Signal Transduction Pathway. *Journal of Virology*, *81*(18), 10072–10080. <https://doi.org/10.1128/JVI.00541-07>
- Halabi, N., Rivoire, O., Leibler, S., & Ranganathan, R. (2009). Protein sectors: evolutionary units of three-dimensional structure. *Cell*, *23*(1), 1–7. <https://doi.org/10.1161/CIRCULATIONAHA.110.956839>
- Ham, S. J., Lee, D., Yoo, H., Jun, K., Shin, H., & Chung, J. (2020). Decision between mitophagy and apoptosis by Parkin via VDAC1 ubiquitination. *Proceedings of the National Academy of Sciences*, *117*(8), 4281–4291. <https://doi.org/10.1073/pnas.1909814117>
- Hay, N., & Sonenberg, N. (2004). Upstream and downstream of mTOR. *Genes Dev.* 2004, *18*, 1926–1945. *Genes Dev*, *18*, 1926–1945. <https://doi.org/10.1101/gad.1212704.hibiting>
- Holmes, E. C. (2004). Adaptation and immunity. In *PLoS Biology* (Vol. 2, Issue 9, pp. 2–4). <https://doi.org/10.1371/journal.pbio.0020307>
- Huang, X. Y., Li, D., Chen, Z. X., Huang, Y. H., Gao, W. Y., Zheng, B. Y., & Wang, X. Z. (2018). Hepatitis B Virus X protein elevates Parkin-mediated mitophagy through Lon Peptidase in starvation. *Experimental Cell Research*, *368*(1), 75–83.

<https://doi.org/10.1016/j.yexcr.2018.04.016>

- Inoki, K., Li, Y., Xu, T., & Guan, K. L. (2003). Rheb GTPase is a direct target of TSC2 GAP activity and regulates mTOR signaling. *Genes and Development*, *17*(15), 1829–1834. <https://doi.org/10.1101/gad.1110003>
- Jin, J., Tang, S., Xia, L., Du, R., Xie, H., Song, J., Fan, R., Bi, Q., Chen, Z., Yang, G., Liu, J., Shi, Y., & Fan, D. (2013). MicroRNA-501 promotes HBV replication by targeting HBXIP. *Biochemical and Biophysical Research Communications*, *430*(4), 1228–1233. <https://doi.org/10.1016/j.bbrc.2012.12.071>
- Kew, M. C. (2011). Hepatitis B virus x protein in the pathogenesis of hepatitis B virus-induced hepatocellular carcinoma. *Journal of Gastroenterology and Hepatology (Australia)*, *26*(SUPPL. 1), 144–152. <https://doi.org/10.1111/j.1440-1746.2010.06546.x>
- Kim, S. J., Khan, M., Quan, J., Till, A., Subramani, S., & Siddiqui, A. (2013). Hepatitis B Virus Disrupts Mitochondrial Dynamics: Induces Fission and Mitophagy to Attenuate Apoptosis. *PLoS Pathogens*, *9*(12), 1–12. <https://doi.org/10.1371/journal.ppat.1003722>
- Kim, S., Kim, H.-Y., Lee, S., Kim, S. W., Sohn, S., Kim, K., & Cho, H. (2007). Hepatitis B Virus X Protein Induces Perinuclear Mitochondrial Clustering in Microtubule- and Dynein-Dependent Manners. *Journal of Virology*, *81*(4), 1714–1726. <https://doi.org/10.1128/jvi.01863-06>
- Kosakovsky Pond, S. L., Poon, A. F. Y., Velazquez, R., Weaver, S., Hepler, N. L., Murrell, B., Shank, S. D., Magalis, B. R., Bouvier, D., Nekrutenko, A., Wisotsky, S., Spielman, S. J., Frost, S. D. W., & Muse, S. V. (2020). HyPhy 2.5—A Customizable Platform for Evolutionary Hypothesis Testing Using Phylogenies. *Molecular Biology and Evolution*, *37*(1), 295–299. <https://doi.org/10.1093/molbev/msz197>
- Lauber, C., Seitz, S., Mattei, S., Suh, A., Beck, J., Herstein, J., Börold, J., Salzburger, W., Kaderali, L., Briggs, J. A. G., & Bartenschlager, R. (2017). Deciphering the Origin and Evolution of Hepatitis B Viruses by Means of a Family of Non-enveloped

- Fish Viruses. *Cell Host and Microbe*, 22(3), 387-399.e6.
<https://doi.org/10.1016/j.chom.2017.07.019>
- Lee, S. H., Cha, E. J., Lim, J. E., Kwon, S. H., Kim, D. H., Cho, H., & Han, K. H. (2012). Structural characterization of an intrinsically unfolded mini-HBX protein from hepatitis B virus. *Molecules and Cells*, 34(2), 165–169.
<https://doi.org/10.1007/s10059-012-0060-z>
- Lemasters, J. J. (2014). Variants of mitochondrial autophagy: Types 1 and 2 mitophagy and micromitophagy (Type 3). *Redox Biology*, 2(1), 749–754.
<https://doi.org/10.1016/j.redox.2014.06.004>
- Letunic, I., & Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research*, 47(W1), W256–W259.
<https://doi.org/10.1093/nar/gkz239>
- Li, T., Robert, E. I., van Breugel, P. C., Strubin, M., & Zheng, N. (2010). A promiscuous α -helical motif anchors viral hijackers and substrate receptors to the CUL4–DDB1 ubiquitin ligase machinery. *Nature Structural and Molecular Biology*, 17(1), 105–112. <https://doi.org/10.1038/nsmb.1719>
- Lok, A. S.-F. (2000). Hepatitis B infection: Pathogenesis and management. *Journal of Hepatology*, 32, 89–97. [https://doi.org/10.1016/S0168-8278\(00\)80418-3](https://doi.org/10.1016/S0168-8278(00)80418-3)
- Lucifora, J., Arzberger, S., Durantel, D., Belloni, L., Strubin, M., Levrero, M., Zoulim, F., Hantz, O., & Protzer, U. (2011). Hepatitis B virus X protein is essential to initiate and maintain virus replication after infection. *Journal of Hepatology*, 55(5), 996–1003. <https://doi.org/10.1016/j.jhep.2011.02.015>
- Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A. R. N., Potter, S. C., Finn, R. D., & Lopez, R. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research*, 47(W1), W636–W641. <https://doi.org/10.1093/nar/gkz268>
- Marusawa, H., Matsuzawa, S. ichi, Welsh, K., Zou, H., Armstrong, R., Tamm, I., & Reed, J. C. (2003). HBXIP functions as a cofactor of survivin in apoptosis

- suppression. *EMBO Journal*, 22(11), 2729–2740.
<https://doi.org/10.1093/emboj/cdg263>
- Melegari, M., Scaglioni, P. P., & Wands, J. R. (1998). Cloning and characterization of a novel hepatitis B virus x binding protein that inhibits viral replication. *Journal of Virology*, 72(3), 1737–1743. <https://doi.org/10.3748/wjg.v11.i36.5666>
- Mertins, P., Mani, D. R., Ruggles, K. V., Gillette, M. A., Clauser, K. R., Wang, P., Wang, X., Qiao, J. W., Cao, S., Petralia, F., Mundt, F., Krug, K., Tu, Z., Lei, J. T., Gatza, M. L., Wilkerson, M., Perou, C. M., Yellapantula, V., Huang, K., ... Cptac, N. C. I. (2016). Proteogenomics connects somatic mutations to signaling in breast cancer. *Nature*, 534(7605), 55–62. <https://doi.org/10.1038/nature18003>. Proteogenomics
- Ming-Yuan Su, Morris, K. L., Do Jin Kim, Fu, Y., Lawrence, R., Stjepanovic, G., Zoncu, R., & Hurley, J. H. (2017). Hybrid Structure of the RagA/C-Ragulator mTORC1 Activation Complex. *J Autism Dev Disord*, 47(3), 549–562.
<https://doi.org/10.1097/CCM.0b013e31823da96d>. Hydrogen
- Mizushima, N., & Komatsu, M. (2011). Autophagy: Renovation of cells and tissues. *Cell*, 147(4), 728–741. <https://doi.org/10.1016/j.cell.2011.10.026>
- Murakami, S., Cheong, J. H., & Kaneko, S. (1994). Human hepatitis virus X gene encodes a regulatory domain that represses transactivation of X protein. *Journal of Biological Chemistry*, 269(21), 15118–15123.
- Murrell, B., Weaver, S., Smith, M. D., Wertheim, J. O., Murrell, S., Aylward, A., Eren, K., Pollner, T., Martin, D. P., Smith, D. M., Scheffler, K., & Kosakovsky Pond, S. L. (2015). Gene-wide identification of episodic selection. *Molecular Biology and Evolution*, 32(5), 1365–1371. <https://doi.org/10.1093/molbev/msv035>
- Murrell, B., Wertheim, J. O., Moola, S., Weighill, T., Scheffler, K., & Kosakovsky Pond, S. L. (2012). Detecting individual sites subject to episodic diversifying selection. *PLoS Genetics*, 8(7). <https://doi.org/10.1371/journal.pgen.1002764>
- Netter, R., Hildt, E., Schuster, R., Chang, S., Ju, H., Schaefer, S., Hsu, Y., Rang, A., Will, H., & Irol, J. V. (2001). Duck Hepatitis B Virus Expresses a Regulatory HBx-

- Like Protein from a Hidden Open Reading Frame. *Journal of Virology*, 75(1), 161–170. <https://doi.org/10.1128/JVI.75.1.161>
- Nguyen, T. N., Padman, B. S., & Lazarou, M. (2016). Deciphering the Molecular Signals of PINK1/Parkin Mitophagy. *Trends in Cell Biology*, 26(10), 733–744. <https://doi.org/10.1016/j.tcb.2016.05.008>
- Notredame, C., Higgins, D. G., & Heringa, J. (2000). T-coffee: a novel method for fast and accurate multiple sequence alignment 1 Edited by J. Thornton. *Journal of Molecular Biology*, 302(1), 205–217. <https://doi.org/10.1006/jmbi.2000.4042>
- Nuismer, S. L., & Thompson, J. N. (2006). Coevolutionary Alternation in Antagonistic Interactions. *Evolution*, 60(11), 2207. <https://doi.org/10.1554/06-111.1>
- Orenstein, S. J., & Cuervo, A. M. (2010). Chaperone-mediated autophagy: Molecular mechanisms and physiological relevance. *Seminars in Cell & Developmental Biology*, 21(7), 719–726. <https://doi.org/10.1016/j.semcdb.2010.02.005>
- Park, S., Choi, S. G., Yoo, S. M., Nah, J., Jeong, E., Kim, H., & Jung, Y. K. (2015). Pyruvate stimulates mitophagy via PINK1 stabilization. *Cellular Signalling*, 27(9), 1824–1830. <https://doi.org/10.1016/j.cellsig.2015.05.020>
- Pascucci, B., D'Errico, M., Romagnoli, A., De Nuccio, C., Savino, M., Pietraforte, D., Lanzafame, M., Calcagnile, A. S., Fortini, P., Baccarini, S., Orioli, D., Degan, P., Visentin, S., Stefanini, M., Isidoro, C., Fimia, G. M., & Dogliotti, E. (2017). Overexpression of parkin rescues the defective mitochondrial phenotype and the increased apoptosis of Cockayne Syndrome A cells. *Oncotarget*, 8(61), 102852–102867. <https://doi.org/10.18632/oncotarget.9913>
- Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*, 14(4), 417–419. <https://doi.org/10.1038/nmeth.4197>
- Phanstiel, D. H., Brumbaugh, J., Wenger, C. D., Tian, S., Probasco, M. D., Bailey, D. J., Swaney, D. L., Tervo, M. A., Bolin, J. M., Ruotti, V., Stewart, R., Thomson, J. A., & Coon, J. J. (2011). Proteomic and phosphoproteomic comparison of human ES and

- iPS cells. *Nature Methods*, 8(10), 821–827. <https://doi.org/10.1038/nmeth.1699>
- Piculell, B. J., Hoeksema, J. D., & Thompson, J. N. (2008). Interactions of biotic and abiotic environmental factors in an ectomycorrhizal symbiosis, and the potential for selection mosaics. *BMC Biology*, 6, 1–11. <https://doi.org/10.1186/1741-7007-6-23>
- Rahmani, Z., Huh, K.-W., Lasher, R., & Siddiqui, A. (2000). Hepatitis B Virus X Protein Colocalizes to Mitochondria with a Human Voltage-Dependent Anion Channel, HVDAC3, and Alters Its Transmembrane Potential. *Journal of Virology*, 74(6), 2840–2846. <https://doi.org/10.1128/jvi.74.6.2840-2846.2000>
- Rasheed, N., Lima, T. B., Mercaldi, G. F., Nascimento, A. F. Z., Silva, A. L. S., Righetto, G. L., Bar-Peled, L., Shen, K., Sabatini, D. M., Gozzo, F. C., Aparicio, R., & Smetana, J. H. C. (2019). C7orf59/LAMTOR4 phosphorylation and structural flexibility modulate Ragulator assembly. *FEBS Open Bio*, 9(9), 1589–1602. <https://doi.org/10.1002/2211-5463.12700>
- Ratnayake, L., Adhvaryu, K. K., Kafes, E., Motavaze, K., & Lakin-Thomas, P. (2018). A component of the TOR (Target Of Rapamycin) nutrient-sensing pathway plays a role in circadian rhythmicity in *Neurospora crassa*. *PLOS Genetics*, 14(6), e1007457. <https://doi.org/10.1371/journal.pgen.1007457>
- Robert A, S., & David M, S. (2017). mTOR Signaling in Growth, Metabolism and Disease. *Cell*, 168(6), 960–976. <https://doi.org/10.1016/j.cell.2017.02.004.mTOR>
- Sahu, R., Kaushik, S., Clement, C. C., Cannizzo, E. S., Scharf, B., Follenzi, A., Potolicchio, I., Nieves, E., Cuervo, A. M., & Santambrogio, L. (2011). Microautophagy of Cytosolic Proteins by Late Endosomes. *Developmental Cell*, 20(1), 131–139. <https://doi.org/10.1016/j.devcel.2010.12.003>
- Shaw, R. J., & Cantley, L. C. (2006). Ras, PI(3)K and mTOR signalling controls tumour cell growth. In *Nature* (Vol. 441, Issue 7092, pp. 424–430). <https://doi.org/10.1038/nature04869>
- Shirakata, Y., & Koike, K. (2003). Hepatitis B virus X protein induces cell death by causing loss of mitochondrial membrane potential. *Journal of Biological Chemistry*,

278(24), 22071–22078. <https://doi.org/10.1074/jbc.M301606200>

Sir, D., Tian, Y., Chen, W. L., Ann, D. K., Yen, T. S. B., & Ou, J. H. J. (2010). The early autophagic pathway is activated by hepatitis B virus and required for viral DNA replication. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(9), 4383–4388. <https://doi.org/10.1073/pnas.0911373107>

Slagle, B. L., & Bouchard, M. J. (2016). Hepatitis B virus X and regulation of viral gene expression. *Cold Spring Harbor Perspectives in Medicine*, *6*(3), 1–20. <https://doi.org/10.1101/cshperspect.a021402>

Spielman, S. J., & Wilke, C. O. (2015). The relationship between dN/dS and scaled selection coefficients. *Molecular Biology and Evolution*, *32*(4), 1097–1108. <https://doi.org/10.1093/molbev/msv003>

Strotz, L. C., Simões, M., Girard, M. G., Breitkreuz, L., Kimmig, J., & Lieberman, B. S. (2018). Getting somewhere with the red queen: Chasing a biologically modern definition of the hypothesis. *Biology Letters*, *14*(5). <https://doi.org/10.1098/rsbl.2017.0734>

Suh, A., Brosius, J., Schmitz, J., & Kriegs, J. O. (2013). The genome of a Mesozoic paleovirus reveals the evolution of hepatitis B viruses. *Nature Communications*, *4*, 1–7. <https://doi.org/10.1038/ncomms2798>

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., Jensen, L. J., & Mering, C. von. (2019). STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*, *47*(D1), D607–D613. <https://doi.org/10.1093/nar/gky1131>

Teng, C.-F., Wu, H.-C., Tsai, H.-W., Shiah, H.-S., Huang, W., & Su, I.-J. (2011). Novel feedback inhibition of surface antigen synthesis by mammalian target of rapamycin (mTOR) signal and its implication for hepatitis B virus tumorigenesis and therapy. *Hepatology*, *54*(4), 1199–1207. <https://doi.org/10.1002/hep.24529>

Thompson, J. N. (2010). Four Central Points about Coevolution. *Evolution: Education*

- and Outreach*, 3(1), 7–13. <https://doi.org/10.1007/s12052-009-0200-x>
- Thul, P. J., Åkesson, L., Wiking, M., Mahdessian, D., Geladaki, A., Ait Blal, H., Alm, T., Asplund, A., Björk, L., Breckels, L. M., Bäckström, A., Danielsson, F., Fagerberg, L., Fall, J., Gatto, L., Gnann, C., Hober, S., Hjelmare, M., Johansson, F., ... Lundberg, E. (2017). A subcellular map of the human proteome. *Science*, 356(6340), eaal3321. <https://doi.org/10.1126/science.aal3321>
- Tsurumi, T., Kobayashi, A., Tamai, K., Yamada, H., Daikoku, T., Yamashita, Y., & Nishiyama, Y. (1996). Epstein–Barr Virus Single-Stranded DNA-Binding Protein: Purification, Characterization, and Action on DNA Synthesis by the Viral DNA Polymerase. *Virology*, 222(2), 352–364. <https://doi.org/10.1006/viro.1996.0432>
- Vives-bauza, C., Zhou, C., Huang, Y., Cui, M., Vries, R. L. A. De, Kim, J., & May, J. (2009). *PINK1-dependent recruitment of Parkin to mitochondria in mitophagy*. 1–6. <https://doi.org/10.1073/pnas.0911187107>
- Weaver, S., Shank, S. D., Spielman, S. J., Li, M., Muse, S. V., & Kosakovsky Pond, S. L. (2018). Datamonkey 2.0: A Modern Web Application for Characterizing Selective and Other Evolutionary Processes. *Molecular Biology and Evolution*, 35(3), 773–777. <https://doi.org/10.1093/molbev/msx335>
- Wen, Y., Golubkov, V. S., Strongin, A. Y., Jiang, W., & Reed, J. C. (2008). Interaction of hepatitis B viral oncoprotein with cellular target HBXIP dysregulates centrosome dynamics and mitotic spindle formation. *Journal of Biological Chemistry*, 283(5), 2793–2803. <https://doi.org/10.1074/jbc.M708419200>
- Wolfson, R. L., & Sabatini, D. M. (2017). The dawn of the age of amino acid sensors for the mTORC1 pathway. *Cell Metabolism*, 176(1), 139–148. <https://doi.org/10.1016/j.physbeh.2017.03.040>
- Yang, B., & Bouchard, M. J. (2012). The Hepatitis B Virus X Protein Elevates Cytosolic Calcium Signals by Modulating Mitochondrial Calcium Uptake. *Journal of Virology*, 86(1), 313–327. <https://doi.org/10.1128/jvi.06442-11>
- Yonehara, R., Nada, S., Nakai, T., Nakai, M., Kitamura, A., Ogawa, A., Nakatsumi, H.,

- Nakayama, K. I., Li, S., Standley, D. M., Yamashita, E., Nakagawa, A., & Okada, M. (2017). Structural basis for the assembly of the Ragulator-Rag GTPase complex. *Nature Communications*, *8*(1), 1–10. <https://doi.org/10.1038/s41467-017-01762-3>
- Yoo, S. M., & Jung, Y. K. (2018). A molecular approach to mitophagy and mitochondrial dynamics. *Molecules and Cells*, *41*(1), 18–26. <https://doi.org/10.14348/molcells.2018.2277>
- Yu, L., McPhee, C. K., Zheng, L., Mardones, G. A., Rong, Y., Peng, J., Mi, N., Zhao, Y., Liu, Z., Wan, F., Hailey, D. W., Oorschot, V., Klumperman, J., Baehrecke, E. H., & Lenardo, M. J. (2010). Termination of autophagy and reformation of lysosomes regulated by mTOR. *Nature*, *465*(7300), 942–946. <https://doi.org/10.1038/nature09076>
- Zhang, T. Y., Chen, H. Y., Cao, J. L., Xiong, H. L., Mo, X. B., Li, T. L., Kang, X. Z., Zhao, J. H., Yin, B., Zhao, X., Huang, C. H., Yuan, Q., Xue, D., Xia, N. S., & Yuan, Y. A. (2019). Structural and functional analyses of hepatitis B virus X protein BH3-like domain and Bcl-xL interaction. *Nature Communications*, *10*(1). <https://doi.org/10.1038/s41467-019-11173-1>

Anexo 2. Declaração de bioética e/ou biossegurança



COORDENADORIA DE PÓS-GRADUAÇÃO
INSTITUTO DE BIOLOGIA
Universidade Estadual de Campinas
Caixa Postal 6109. 13083-970, Campinas, SP, Brasil
Fone (19) 3521-6378. email: cpjib@unicamp.br



DECLARAÇÃO

Em observância ao **§5º do Artigo 1º da Informação CCPG-UNICAMP/001/15**, referente a Bioética e Biossegurança, declaro que o conteúdo de minha Dissertação de Mestrado, intitulada "**ESTUDOS DAS RELAÇÕES EVOLUTIVAS E FUNCIONAIS DA PROTEINA VIRAL HBX COM AS SUBUNIDADES DO COMPLEXO RAGULATOR**", desenvolvida no Programa de Pós-Graduação em Genética e Biologia Molecular do Instituto de Biologia da Unicamp, não versa sobre pesquisa envolvendo seres humanos, animais ou temas afetos a Biossegurança.

Assinatura: _____

Nome do(a) aluno(a): Jordy Alexander Larco Lasso

Assinatura: _____

Nome do(a) orientador(a): Juliana Helena Costa Smetana

Data: 29/03/2021

Anexo 3 Documento do comite de bioética e/ou biossegurança

Uso exclusivo da CIBio:

Número de projeto / processo: 2019-11

Formulário de encaminhamento de projetos de pesquisa com OGMs para análise da CIBio - CNPEM

1. Título do projeto: Estudo bioinformático e proteômico das isoformas da proteína HBXIP e sua relação com a proteína viral HBx

2. Pesquisador responsável: Juliana Helena Costa Smetana

3. Experimentador(es):
Jordy Alexander Lasso Larco
Mariana Piccoli Gonçalves

Nível do treinamento do experimentador: -Iniciação científica, JALL]-mestrado, -doutorado, MPG]-doutorado direto, -pós-doutorado, -nível técnico, -outro, especifique:

4. Unidade operativa: LNLS LNNano LNBR LNBio

5. Maior Classe de risco de OGM deste projeto: Risco I Risco II Risco III Risco IV

6. O projeto é confidencial? não sim

7. No caso de projeto confidencial, o título do projeto pode constar em lista aberta no CNPEM? não sim

8. Qual é o objetivo do projeto? Estudo estrutural e funcional da interação entre a proteína HBx do vírus da hepatite B e a proteína humana HBXIP /Lamtor5 (HBx Interacting Protein), comparando suas isoformas Long e Short.

9. Informe um número e nome para cada OGM, organismo receptor, organismo doador, o transgene e classe de risco do OGM.

- (1) HBV 1.3-mer P-null replicon, células humanas em cultura (HepG2), vírus Hepatite B, genoma do vírus da hepatite B deficiente na proteína do capsídeo, risco I
- (2) pCDNA3-HA-HBXIP, células humanas em cultura (HEK293), humano, proteína HBXIP/Lamtor5, risco I
- (3) pRK5-FLAG -HBXIP, células humanas em cultura (HEK293), humano, proteína HBXIP/Lamtor5, risco I
- (4) pRK5-p18-HA, células humanas em cultura (HEK293), humano, proteína p18/Lamtor1, risco I
- (5) pRK5-FLAG-p14, células humanas em cultura (HEK293), humano, proteína p14/Lamtor2, risco I
- (6) pRK5-HA-MP1, células humanas em cultura (HEK293), humano, proteína MP1/Lamtor3, risco I
- (7) pRK5-p18-FLAG, células humanas em cultura (HEK293), humano, proteína p18/Lamtor1, risco I
- (8) pCDNA-FLAG-c7orf59, células humanas em cultura (HEK293), humano, proteína c7orf59/Lamtor4, risco I
- (9) pGFP-HBx, células humanas em cultura (HEK293), vírus HBV, proteína HBx e GFP, risco I
- (10) pGFP-HBx NESM, células humanas em cultura (HEK293), vírus HBV, proteína HBx (mutante do sinal de exportação nuclear) e GFP, risco I
- (11) pcDNA3.1-Flag HBx, células humanas em cultura (HEK293), vírus HBV, proteína HBx, risco I
- (12) pcDNA3.1-Flag HBx (G124L, I127A), células humanas em cultura (HEK293), vírus HBV, proteína HBx (mutante do motivo BH3), risco I
- (13) pET-Duet-HBXIP(long)-c7orf59, bactérias E. coli, humano, proteína HBXIP/Lamtor5 e c7orf59/Lamtor4, risco I
- (14) pET-Duet-HBXIP(short)-c7orf59, bactérias E. coli, humano, proteína HBXIP/Lamtor5 e c7orf59/Lamtor4, risco I
- (15) pACYC-HBXIP(short)-c7orf59, bactérias E. coli, humano, proteína HBXIP/Lamtor5 e c7orf59/Lamtor4, risco I
- (16) pETSUMO-HBx, bactérias E. coli, vírus HBV, proteína HBx, risco I.

10. Descreva brevemente a função dos transgenes de cada OGM:

- (1) o replicon contém o genoma do vírus da hepatite B para estudos de resposta à infecção em cultura de células. O defeito na proteína do capsídeo ("P-null") impede a montagem da partícula viral. Dessa forma não há produção de partículas virais. O vetor serve apenas para a expressão de genes virais em cultura de células, sem a produção de partículas virais. A célula transfectada não é patogênica nem infecciosa.
- (2) a : as proteínas MP1, p14, p18, HBXIP e c7orf59 são subunidades do complexo Ragulator, que ativa a via mTORC1 na presença de aminoácidos.
- () a proteína HBx do vírus da hepatite B tem a função de modular processos celulares como apoptose e sinalização por cálcio, sendo importante na tumorigênese causada por esse vírus.

Uso exclusivo da CIBio:

Número de projeto / processo: 2019-11

Formulário de encaminhamento de projetos de pesquisa com OGMs para análise da CIBio - CNPEM

(9) (10) GFP: proteína fluorescente verde usada como repórter

11. Algum OGM produz proteína tóxica, oncogênica ou pode gerar produtos deletérios para saúde humana, animal ou meio ambiente? As proteínas HBx, HBXIP e possivelmente outras subunidades do Regulator são oncoproteínas.

12. Algum OGM é agente patogênico esporulante? | | Não | | Sim: _____

13. Algum OGM é agente patogênico e pode se propagar pelo ar? | | Não | | Sim : _____

14. Algum transgene confere infectividade ou patogenicidade para os OGMs? Descreva. Todos os OGMs apresentam resistência ao antibiótico ampicilina, cloranfenicol ou kanamicina. Nenhum transgene confere infectividade ou patogenicidade ao organismo receptor.

15. Com relação aos cuidados preventivos associados a manipulação dos organismos, será necessária alguma avaliação médica periódica para experimentadores? | | Não | | Sim. Que tipo de avaliação? (Ex: consulta com médico, exames laboratoriais etc...) Qual periodicidade? Onde será realizada esta avaliação?

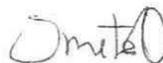
16. Com relação aos cuidados preventivos associados a manipulação dos organismos, será necessária alguma vacinação preventiva para experimentadores? | | Não | | Sim. Qual periodicidade? Onde será realizada esta vacinação?

17. No caso de uma eventual contaminação com organismos patogênicos ou toxinas, descreva medidas emergenciais para tratamento de pessoas envolvidas, descontaminação de equipamentos, instalações e meio ambiente. Não se aplica

18. Projetos que façam uso de organismos ou genes associados ao patrimônio genético brasileiro precisam de cadastro na plataforma SISGEN (www. sisgen.gov.br). É de total responsabilidade do pesquisador responsável esse cadastramento e cumprimento da legislação. O projeto envolve manipulação, transferência, modificação, armazenamento, coleta de Organismos e derivados relativos ao patrimônio genético brasileiro? () SIM, () Não. No caso de responder sim, mencionar a seguir quais os códigos de acesso do cadastro no SISGEN: _____

O pesquisador principal tem conhecimento de que conforme a RDC 50 de 21/02/2002 da Anvisa, é responsável por determinar a classificação de riscos de seu projeto, assim como determinar EPIs e medidas de segurança necessárias para prevenir a contaminação de experimentadores, equipamentos, instalações, terceiros e meio ambiente. O pesquisador responsável também precisará providenciar rotina para realização de exames médicos e laboratoriais para sua equipe, bem como vacinações quando aplicável. Todos os experimentadores envolvidos devem ser supervisionados pelo pesquisador principal, que é o responsável pelo treinamento de biossegurança adequado às suas necessidades para a manipulação, armazenamento, descarte e transporte de OGMs, atendendo a legislação e normativas preconizadas pela CTNBio, Anvisa e outros órgãos e agências regulamentadoras e fiscalizadoras.

Assinatura eletrônica do pesquisador responsável:



Uso exclusivo da CIBio:

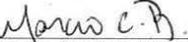
Número de projeto / processo: 2019-11

Formulário de encaminhamento de projetos de pesquisa com OGMs para análise da CIBio - CNPEM

A CIBio analisou este projeto em reunião realizada no dia: 27/6/2019.

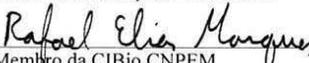
Parecer final: -projeto aprovado, []-projeto recusado, []-projeto com deficiências.

comentários da CIBio:

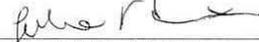

 Presidente da CIBio CNPEM
 Marcio Chaim Bajgelman


 Membro da CIBio CNPEM
 Celso Eduardo Benedetti

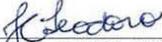
Membro da CIBio CNPEM
 Rafael Elias Marques Pereira Silva


 Membro da CIBio CNPEM
 Douglas Galante




 Membro da CIBio CNPEM
 Juliana Velasco de Castro Oliveira


 Membro da CIBio CNPEM
 Daniel Kolling


 Membro da CIBio CNPEM
 Juliana Conceição Teodoro


 Membro da CIBio CNPEM
 Mateus Borba Cardoso

Anexo 4 Termo de direitos autorais

DIREITOS AUTORAIS

Declaração

As cópias de artigos de minha autoria ou de minha co-autoria, já publicados ou submetidos para publicação em revistas científicas ou anais de congressos sujeitos a arbitragem, que constam da minha Dissertação/Tese de Mestrado/Doutorado, intitulada Studies of the evolutionary and functional relationship of viral protein HBx with Ragulator subunits, não infringem os dispositivos da Lei n.º 9.610/98, nem o direito autoral de qualquer editora.

Campinas,

16/03/2021

Assinatura : _____

Nome do(a) autor(a): Jordy Alexander Larco Lasso

RMN n.º 1 G2783532

Assinatura : _____

Nome do(a) orientador(a): Juliana Helena Costa Smetana

RG n.º 32.221.649-7