XBP1s activity from RNA-seq samples

RESEARCH REPORT

Gabriel Esmiol

Summer 2024

Under the supervision of Dr. Jérémie Breda & Prof. Felix Naef

Institute of Bioengineering: Laboratory of Computational and Systems Biology
EPFL

Abstract

XBP1s is a transcription factor for many proteins. It’s an isoform of the protein XBP1 which is spliced during the Unfolded Protein Response (UPR) after a large production of protein in the endoplasmic reticulum (ER). Research about XBP1 is recent and we don’t know the precise mechanism behind its operation. This is a computational biology study that focuses on the expression of XBP1s around the clock to find repetitive patterns and differences/similarities between male and female.

Introduction

X-box binding protein 1 (XBP1) has a large scale of effects and touches several primordial metabolic mechanisms of cells. For this reason, XBP1 may draw attention to different parts of its action area such as ER stress response, secretory function, lipid metabolism, glucose homeostasis, inflammatory response [1] and immunity differentiation. Because of that, there are also a lot of related diseases like cancer (breast, liver …), diabetes and neurodegenerative diseases like Alzheimer's disease. In this report we will talk first on the mechanism behind the splice of XBP1 then about the method of research and finally we will analyze the results.

I XBP1 splicing events

1 UPR mechanism

XBP1 plays a significant role in the unfolded protein response (UPR). The endoplasmic reticulum (ER) has an important role in the synthesis, assembly and folding of proteins. Perturbation of the homeostasis of the ER leads to misfolding proteins which provoque ER stress (ERS). This stress can be caused by multiple factors such as oxidative stress, hypoxia, protein overexpression, nutriment deprivation, calcium depletion. The ER chaperone-binding immunoglobulin protein (BiP or GRP78), normally binds to the sensor proteins which inhibit their activity, will dissociate from the transducers in response to the accumulation of misfolded proteins. It will provoque the activation of

inositol-requiring enzyme 1 (IRE1) and protein kinase R (PKR)-like ER kinase (PERK). Activating transcription factor 6 (ATF6) liberated from BiP will translocate to the Golgi apparatus. It will then be cleaved to release the N-terminal domain that acts as a transcription factor (TF) to initiate UPR [1].
This response is constituted of three signaling pathways which are recognized by the following ER sensor proteins in mammalian cells: IRE1, PERK, and ATF6 [2]. By those pathways, the UPR will promote the transcription of chaperones and folding enzymes that will increase the protein folding capacity of the ER, inhibit protein translation, and transport misfolded proteins to the cytosol in order to degrade them. It can also cause apoptosis if it can't relieve ERS.

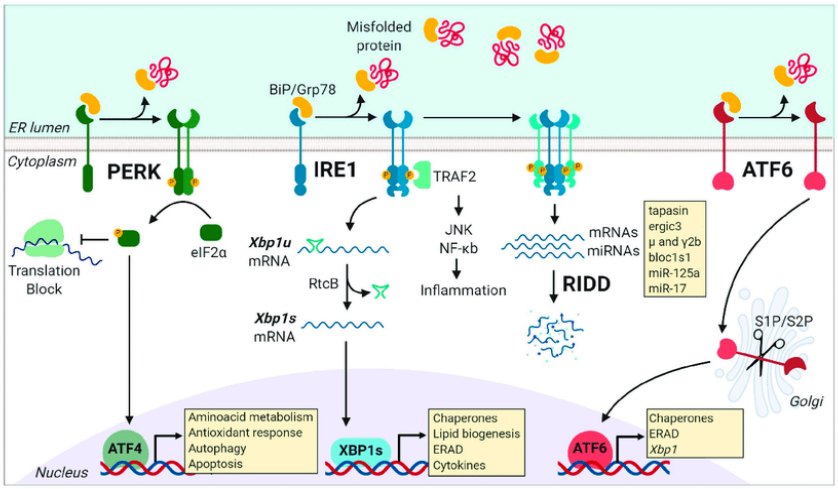


Fig.1: *Activation of the three UPR pathways [3]*

2 XBP1 in UPR

In the UPR, ATF6 will induce XBP1 transcription by binding to the ER stress response element (ERSE) [5] but It's IRE1 that will activate XBP1. Mammalian cells express two isoforms of IRE1: α and β. During ERS periods, activated IREα will recognize the stem-loop structure of XBP1 mRNA containing the consensus sequence (5'-CUGCAG-3') and promote splicing by cleaving 26 intronic nucleotides. [1] [4]. This spliced mRNA will be translated into XBP1 spliced (XBP1s) instead of an unspliced XBP1 (XBP1u). XBP1s is an transcription factor (TF) (active) which regulates expressions of UPR target genes (ER chaperones:Dnajb9, Dnajb11, Pdia3 and Dnajc3, ERAD component: Edem1, Herpud1, and Hrd1, folding enzyme: Pdia6, ER translocon: Sec61a1) [1].

3 Splicing event

This very unconventional splicing of XBP1 occurs in the cytosol allowing a quick production unlike conventional one that occurs in the nucleus dependent on the catalytic function of spliceosomes. It is cleaving 26 intronic nucleotides and translates from 267 amino acids (mice) in XBP1u to 371 in XBP1s by a frameshift (fig.3A) that alters the position of the codon stop resulting in a larger protein (fig.3B). It causes a total modification of the C terminus of XBP1s that enables it to regulate target gene transcription and leads to a different biological function than XBP1u (fig.2) [6]. *"While IRE1α can form large clusters that have been proposed to function as XBP1 processing centers on the ER, the actual oligomeric state of active IRE1α complexes as well as the targeting mechanism that recruits XBP1 to IRE1α oligomers remains unknown."*[7]

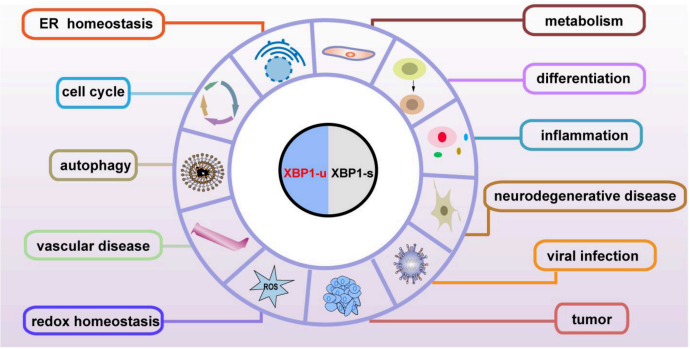


Fig. 2: *Pathophysiological functions of XBP1-u and XBP1-s [6]*

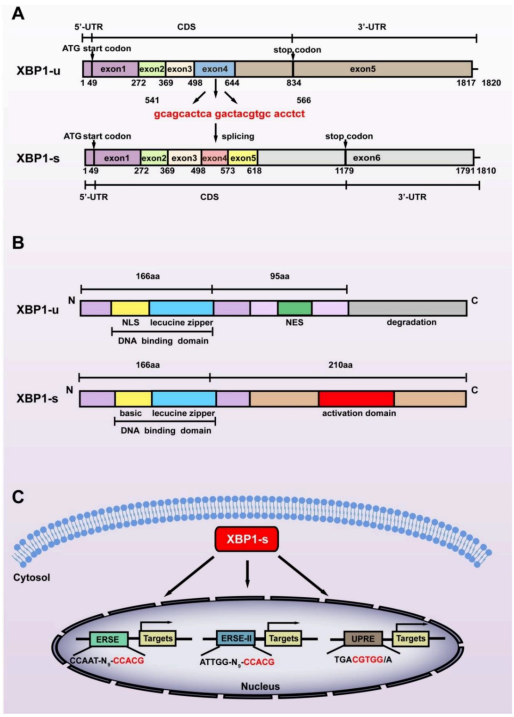


Fig. 3: *Structural differences between the two splicing isoforms of XBP1 [6]*

II Methods of research

We analyzed the expression of XBP1s in RNA-seq (sequencing) samples from different datasets. For that, scripts were written using tools Kallisto and STAR for the quantification of the data and snakemake to optimize the script. The datasets that we used are the Atger dataset [8], Zhang dataset [9], Weger dataset [10] and others from the

NCBI [11] [12]. We used IGV to visualize some sashimi plots and UCSC browser to see different XBP1 isoforms (Xbp1s and Xbp1u). During this research we tried to answer 3 main questions:
1.Does the expected expression appear ?
2.Can we see some distinctable pattern in the expression of Xbp1s around the clock ?
3.Is the expression of XBP1s different between male and female ?

III Results and interpretation

Results that you will see in 1-2-3 paragraphs were made by a script that quantified Xbp1s and XBP1u expressions using Kallisto and Snakemake (for optimization).

1 Atger dataset

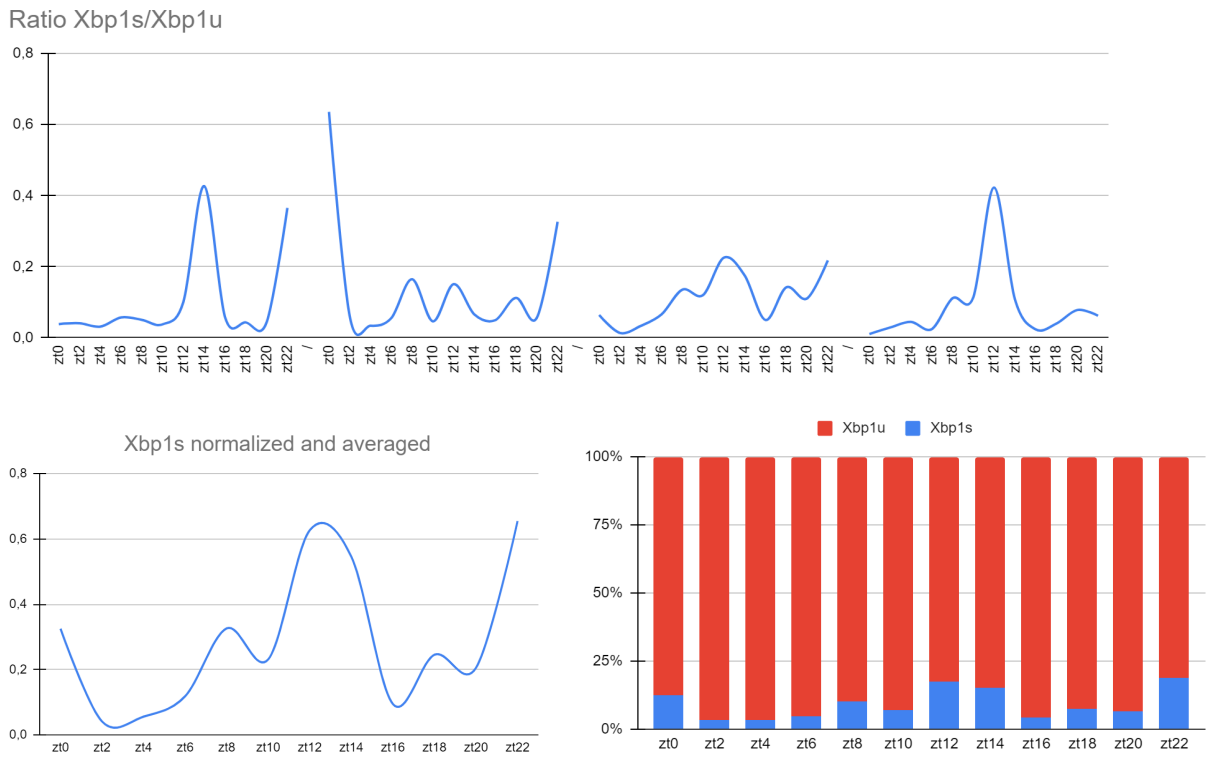


Fig. 4: *Ratio of Xbp1s/Xbp1u of Ad Libitum samples from Atger dataset*
Fig. 5: *Xbp1s expression of Ad Libitum samples from Atger dataset normalized and averaged*
Fig. 6: *Proportion of Xbp1s/Xbp1u of Ad Libitum samples from Atger dataset*

Those samples are very interesting (Fig. 4) because of the small gap between each sample (2 hours) that allowed a good representation of the expression of XBP1s. They are from mouse liver with an Ad Libitum diet. After some treatment: min/max normalization and averaged we obtain the Fig.5. This graph is very interesting because we can distinguish a pattern of around 11 to 12 hours and not 24 hours as many circadian genes are (note that Xbp1 is not proven to be a circadian gene for now). Which leads to the question: Is XBP1 a circadian gene or is it independent of the circadian clock ?
Fig. 6 shows us that a small concentration of Xbp1s compared to Xbp1u is needed to make an adequate response. We can explain this by the fact that Xbp1u facilitates UPR activation and prevents overactivation [13].

2 Zhang dataset

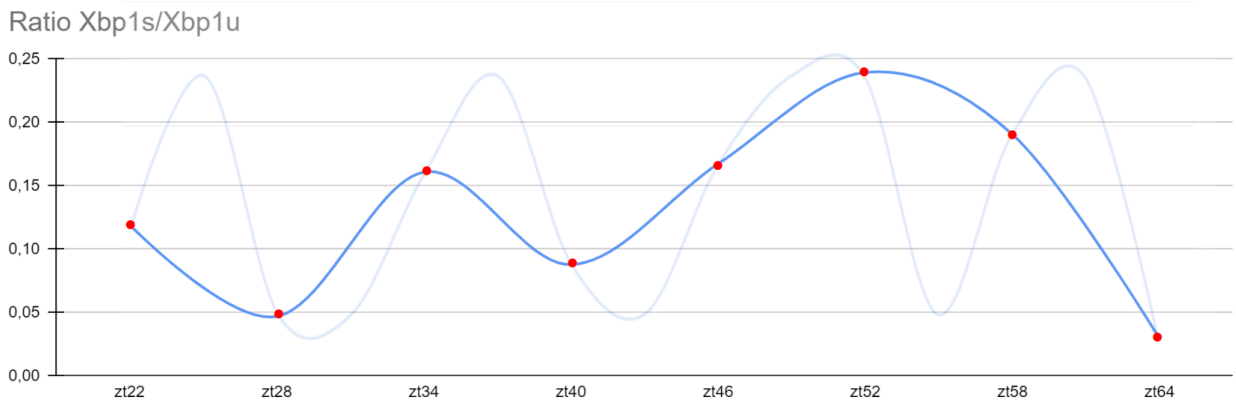
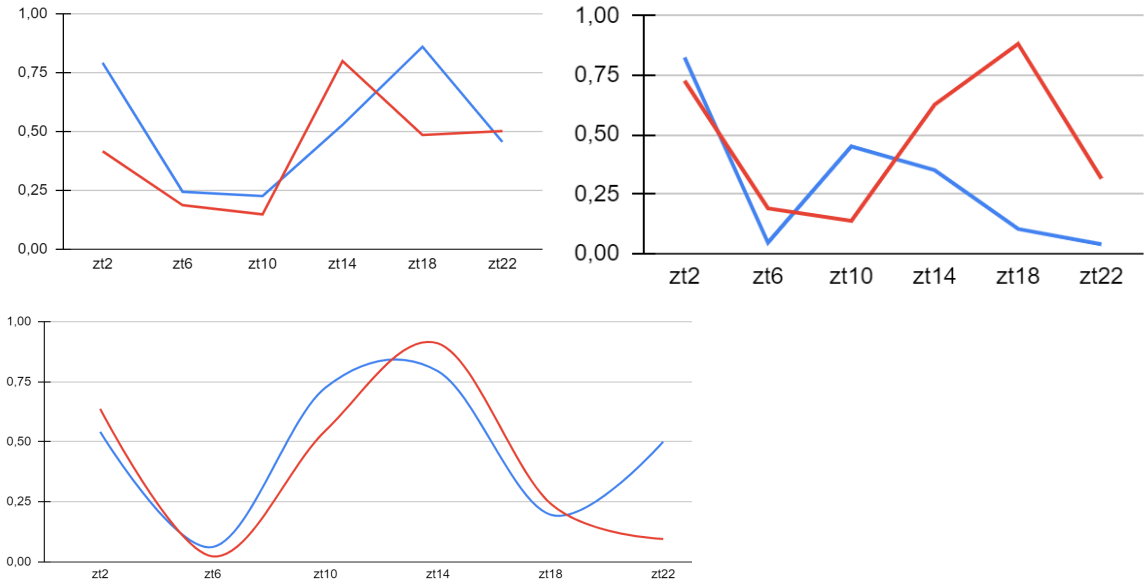


Fig. 7: *Ratio of Xbp1s/Xbp1u from Zhang dataset in blue, light blue curve is an artificial curve that doesn't represent a real curve*

From this database we see a pattern (Fig. 7) of 17 hours which is a bit far from what we had before with the Atger dataset. It can be explained by the fact that the sampled frequency is 6 hours and that we expected a frequency of 12 hours. The fact that the real frequency is about 2 times the sampled frequency means that the signal is probably aliased and doesn't represent reality. We need to have a sampled frequency less than 6 hours to have a correct representation of the expression of Xbp1s.

3 male/Female comparison & Weger dataset



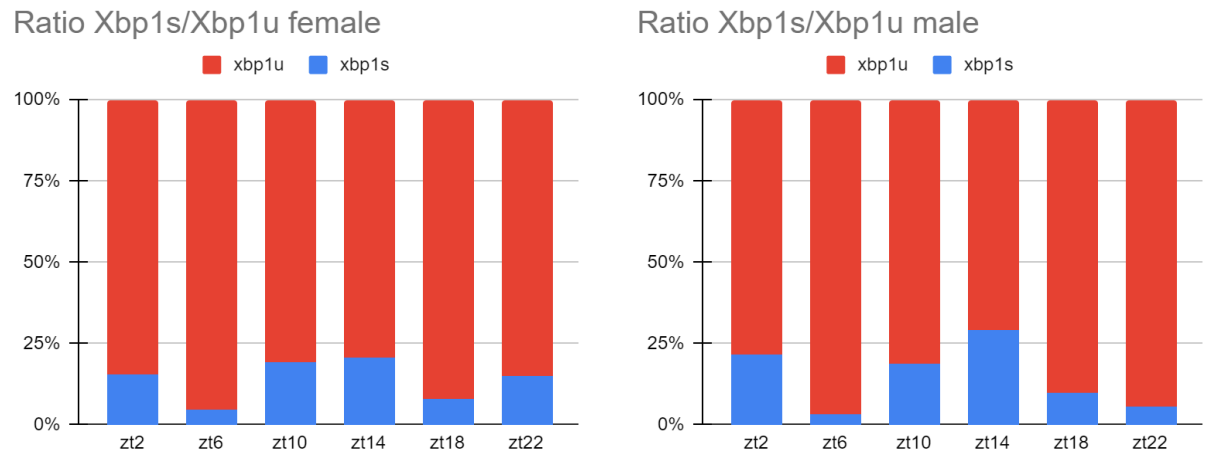


Fig. 8: *Xbp1s expression for male and female of Calorie-restricted diet (mouse) samples from NCBI dataset [11], normalized and averaged*
Fig. 9: *Xbp1s expression for male and female of Ad-libitum diet (mouse) samples from NCBI dataset [12], normalized and averaged*
Fig. 10: *Xbp1s expression for male and female of Restricted feeding diet samples from Weger dataset, normalized and averaged*
Fig. 11: *Proportion of Xbp1s/Xbp1u of Restricted feeding diet female samples from Weger dataset*
Fig. 12: *Proportion of Xbp1s/Xbp1u of Restricted feeding diet male samples from Weger dataset*

With the NCBI Calorie-restricted diet and the Weger dataset we can easily distinguish similarities in the curves of Fig. 8 and in the curves of Fig. 10. Those curves have, pair to pair, the same shape and are synchronized. Hence, It appears that the variation in the

ratio between Xbp1s and Xbp1u is the same for male and females. However, it seems that the gap between min and max values may be larger for male in the Weger dataset (see Fig. 11 and Fig. 12). This tendency can be observed too with the NCBI dataset (Fig. 8 and Fig. 9).

4 Sashimi plot

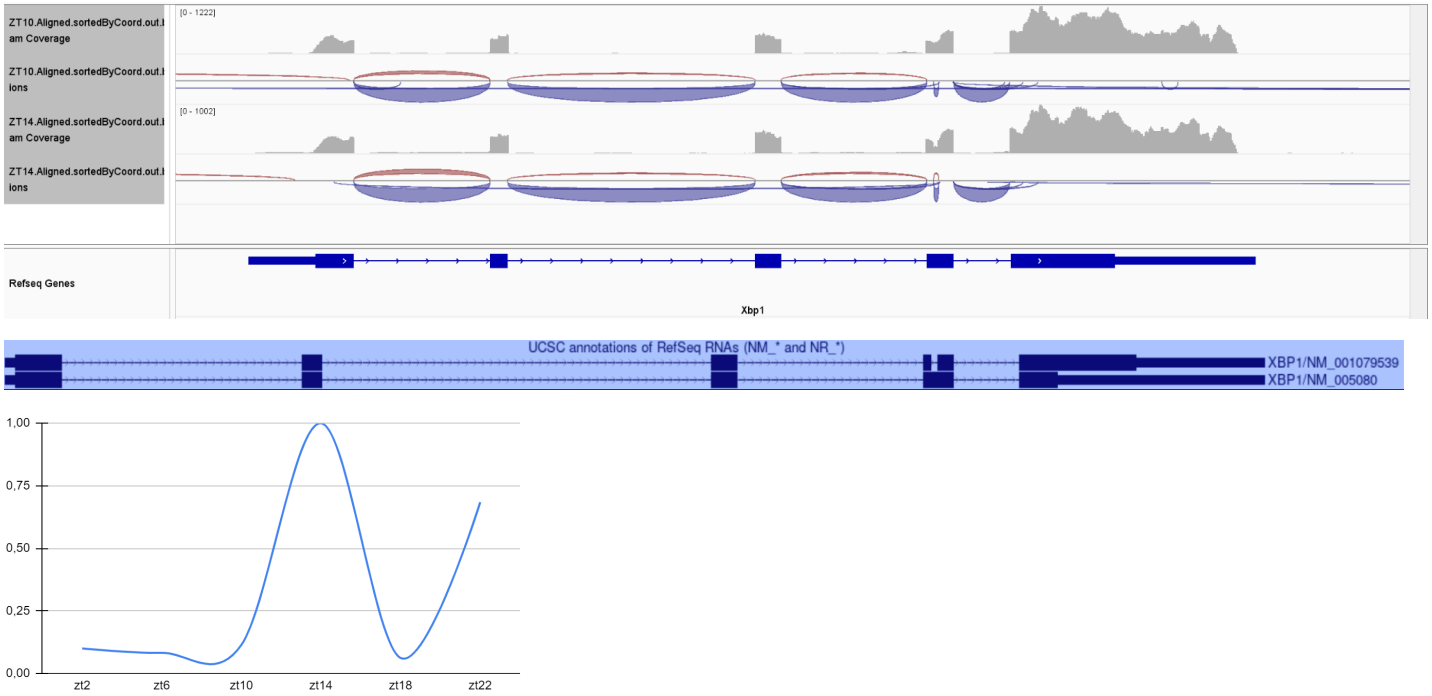


Fig. 13: *Sashimi plot of aligned RNA-seq from Atger dataset Restricted feeding, the first one (up) is at 10h and the second one (down) at 14h*
Fig. 14: *Representation of Xbp1s (up) and Xbp1u (down)*
Fig. 15: *Xbp1s expression of Restricted feeding samples from Atger dataset normalized and averaged*

In This part, we've taken the data from Atger dataset Restricted feeding 10h and 14h (same data as Fig. 15) to see in another perspective if our results are coherent. To start, with Fig. 15 we see that the sample of 10 hours and 14 hours have a ratio of Xbp1s/Xbp1u significatively at the opposite. Hence, we expect to have in the sashimi plot a distinction that shows the presence of Xbp1s in the 14 hours sample that is less in the 10 hours sample. In order to see this sashimi plot (Fig. 13), we used the tool STAR and we put the result on IGV to see it. The Fig.14 shows the splice on the exon 4 that should be seen on the sashimi plot. If we observe this sashimi plot (Fig. 13), we distinguish less expression in the middle of the exon 4. The variation is more clear on the 14 hours sample than for the 10 hours sample. Because Xbp1s hasn't this little part that has Xbp1u when the ratio of Xbp1s/Xbp1u increases this part in the middle of the exon 4th is less expressed than the others parts that are expressed by both isoforms. This gap in the exon 4th shows us

well that the 14 hours sample has more Xbp1s than the 10 hours sample (in comparison to Xbp1u) which was what we expected to see.

Conclusion

Xbp1s is a rhythm gene of around 12 hours and It needs to be taken into consideration when choosing the time interval between samples. It means to have an interval smaller than 6 hours. It appears also that a small concentration of Xbp1s (compared to the concentration Xbp1u) is sufficient for the organism to function correctly. As we explained above, Xbp1u facilitates UPR activation and prevents overactivation.

Rhythm patterns in Xbp1s expression are the same for male and female and are synchronized. However, It appears that the difference between the min/max value is stronger for male than for female. It may come from a stronger UPR reaction or a more effective inhibition of IRE1 that spiced XBP1u. More specific research needs to be done.

RÉFÉRENCES

[1] Park, Sung-Min, Tae-Il Kang, et Jae-Seon So. « Roles of XBP1s in Transcriptional Regulation of Target Genes ». Biomedicines 9, nº 7 (juillet 2021): 791. https://doi.org/10.3390/biomedicines9070791

[2] Walter, Peter, et David Ron. « The Unfolded Protein Response: From Stress Pathway to Homeostatic Regulation ». Science 334, nº 6059 (25 novembre 2011): 1081‑86. https://doi.org/10.1126/science.1209038

[3] Flores-Santibáñez, Felipe, Bernardita Medel, José Ignacio Bernales, et Fabiola Osorio. « Understanding the Role of the Unfolded Protein Response Sensor IRE1 in the Biology of Antigen Presenting Cells ». Cells 8, nº 12 (décembre 2019): 1563. https://doi.org/10.3390/cells8121563

[4] Oikawa, Daisuke, Mio Tokuda, Akira Hosoda, et Takao Iwawaki. « Identification of a Consensus Element Recognized and Cleaved by IRE1α ». Nucleic Acids Research 38, nº 18 (1 octobre 2010): 6265‑73. https://doi.org/10.1093/nar/gkq452

[5] Yoshida, Hiderou, Toshie Matsui, Akira Yamamoto, Tetsuya Okada, et Kazutoshi Mori. « XBP1 mRNA Is Induced by ATF6 and Spliced by IRE1 in Response to ER Stress to Produce a Highly Active Transcription Factor ». Cell 107, nº 7 (décembre 2001): 881‑91. https://doi.org/10.1016/S0092-8674(01)00611-0

[6] Luo, Xinxin, Leader Alfason, Mankun Wei, Shourong Wu, et Vivi Kasim. « Spliced or Unspliced, That Is the Question: The Biological Roles of XBP1 Isoforms in Pathophysiology ». International Journal of Molecular Sciences 23, nº 5 (2 mars 2022): 2746. https://doi.org/10.3390/ijms23052746

[7] « XBP1 mRNA Splicing on the ER ». Consulté le 3 juillet 2024. https://www.mls.uzh.ch/en/research/voigt/research/XBP1-mRNA-splicing-on-the-ER.html

[8] Atger, Florian, Cédric Gobet, Julien Marquis, Eva Martin, Jingkui Wang, Benjamin Weger, Grégory Lefebvre, Patrick Descombes, Felix Naef, et Frédéric Gachon. « Circadian and Feeding Rhythms Differentially Affect Rhythmic mRNA Transcription and Translation in Mouse Liver ». Proceedings of the National Academy of Sciences 112, nº 47 (24 novembre 2015). https://doi.org/10.1073/pnas.1515308112

[9] Zhang, Ray, Nicholas F. Lahens, Heather I. Ballance, Michael E. Hughes, et John B. Hogenesch. « A Circadian Gene Expression Atlas in Mammals: Implications for Biology and Medicine ». Proceedings of the National Academy of Sciences 111, nº 45 (11 novembre 2014): 16219‑24. https://doi.org/10.1073/pnas.1408886111

[10] Weger, Benjamin D., Cédric Gobet, Fabrice P. A. David, Florian Atger, Eva Martin, Nicholas E. Phillips, Aline Charpagne, Meltem Weger, Felix Naef, et Frédéric Gachon. « Systematic Analysis of Differential Rhythmic Liver Gene Expression Mediated by the Circadian Clock and Feeding Rhythms ». Proceedings of the National Academy of Sciences 118, nº 3 (19 janvier 2021): e2015803118. https://doi.org/10.1073/pnas.2015803118

[11] Consulté le 16 septembre 2024. https://www.ncbi.nlm.nih.gov/Traces/study/?acc=PRJNA893576&o=acc_s%3Aa

[12] Consulté le 16 septembre 2024. https://www.ncbi.nlm.nih.gov/Traces/study/?acc=PRJNA873076&o=acc_s%3Aa

[13] Yücel, Sara Suna, Walter Stelzer, Alessandra Lorenzoni, Manfred Wozny, Dieter Langosch, et Marius K. Lemberg. « The Metastable XBP1u Transmembrane Domain Defines Determinants for Intramembrane Proteolysis by Signal Peptide Peptidase ». Cell Reports 26, nº 11 (12 mars 2019): 3087-3099.e11. https://doi.org/10.1016/j.celrep.2019.02.057