Volume 11 | Issue 2 | December 2021 ISSN 2182-0287

Journal of Integrated OMICS

a methodological journal

JIOMICS

Journal of Integrated OMICS

Focus and Scope

Journal of Integrated OMICS, JIOMICS, provides a forum for the publication of original research papers, preliminary communications, technical notes and critical reviews in all branches of pure and applied "-omics", such as genomics, proteomics, lipidomics, metabolomics or metallomics. The manuscripts must address methodological development. Contributions are evaluated based on established guidelines, including the fundamental nature of the study, scientific novelty, and substantial improvement or advantage over existing technology or method. Original research papers on fundamental studies, and novel sensor and instrumentation development, are especially encouraged. It is expected that improvements will also be demonstrated within the context of (or with regard to) a specific biological question; ability to promote the analysis of molecular mechanisms is of particular interest. Novel or improved applications in areas such as clinical, medicinal and biological chemistry, environmental analysis, pharmacology and materials science and engineering are welcome.

Editors-in-Chief

Regional editors

Associated editors

AUSTRALIA AND NEW ZEALAND

EUROPE

SOUTH AMERICA

JOURNAL OF INTEGRATED OMICS

A methodological Journal

SPECIAL ISSUE: SELECTED ABSTRACTS OF THE IV INTERNATIONAL CAPARICA CONFERENCE IN SPLICING (SPLICING 2021) AND THE VII INTERNATIONAL CAPARICA CONFERENCE ON ANALYTICAL PROTEOMICS (ICAP 2021)

REVIEWS

SPECIAL ISSUE

SELECTED ABSTRACTS OF THE IV INTERNATIONAL CAPARICA CONFERENCE IN SPLICING (SPLICING 2021) AND THE VII INTERNATIONAL CAPARICA CONFERENCE ON ANALYTICAL PROTEOMICS (ICAP 2021)

JOURNAL OF INTEGRATED OMICS
A METHODOLOGICAL JOURNAL
HTTP://WWW.JIOMICS.COM

SPECIAL ISSUE: SELECTED ABSTRACTS OF THE IV INTERNATIONAL CAPARICA CONFERENCE IN SPLICING (SPLICING 2021)

The Interplay between ER Stress, RNA Splicing and Intracellular pH Control in Arabidopsis Immunity

Karolina M. Pajerowska-Mukhtar[1]*

[1]University of Alabama at Birmingham, Dept. of Biology, Birmingham, AL 35294, USA

Available Online: 31 December 2021

ABSTRACT

Abiotic and biotic stresses can severely perturb endoplasmic reticulum (ER) function. The unfolded protein response (UPR) is a three-pronged signalling axis dedicated to preserving ER homeostasis. If these mechanisms of adaptation and survival are insufficient to recover the ER homeostasis, cells will initiate apoptosis. While the importance of pH in basic cellular physiology has been recognized for over a century, it remains poorly understood how alterations in the pH level affect host abiotic and biotic stress responses in Arabidopsis plants. We obtained preliminary data that led us to hypothesize that pathogenic microbes and environmental stress such as heat and drought, affect a universal eukaryotic cellular stress sensor IRE1 (Inositol-requiring enzyme 1) to modulate plant stress responses. Arabidopsis IRE1 directly cleaves bZIP60 (basic leucine zipper 60) mRNA in response to various environmental stresses, leading to the production of an active transcription factor that promotes the expression of multiple ER stress-responsive genes. IRE1 is also known to engage in cleavage and bulk degradation cleave of certain mRNAs in a process called Regulated IRE1-Dependent Decay (RIDD). The exact nature and function of the degraded mRNAs are not completely understood, though it is known that RIDD is integral in the determination of IRE1 to switch between pro-survival and pro-death signalling functions.

We performed genome-wide transcriptomic analyses of RIDD pathways upon pathogen infection (Pseudomonas syringae DC3000 -avrRpm1) in wild-type Arabidopsis Col-0 plants following treatment with Actinomycin D (inhibitor of de novo transcription). Pathway analysis on transcripts stabilized in the ire1a/ire1b double mutant plants (=RIDD targets) identified a specific enrichment of proline-related genes that are implicated in the cellular pH homeostasis. RNA secondary structure analysis of these genes predicted the formation of hallmark features of IRE1 cleavage clients: a double hairpin structure and presence of nucleotide consensus sequences corresponding to IRE1 cleavage sites. We tested a number of developmental, stress and immune phenotypes of ire1a/ire1b double mutants following treatments with proline and we observed altered intracellular and apoplastic pH, and severe growth retardation. The ire1a/ire1b plants also susceptibility to bacterial infection and heat sensitivity, which are further magnified by pH-altering treatments. In summary, our data show that Arabidopsis IRE1a and IRE1b participate with various degrees in the control of intracellular pH by modulating the transcript accumulation of Pro-related genes in response to biotic and abiotic stresses.

Keywords: ER stress; IRE1; bZIP60; RIDD; pH control

Acknowledgments:

This work was supported by the National Science Foundation CAREER award (IOS-1350244) to K.P.M.

References:

[1] Verchot J and Pajerowska-Mukhtar KM. Current Opinion in Virology, 47 (2021), 9-17.
[2] Afrin T, Seok M, Terry B, Pajerowska-Mukhtar KM. Scientific Reports, 10 (2020), 19154.
[3] Afrin T, Diwan D, Sahawneh K, Pajerowska-Mukhtar KM. Journal of Experimental Botany, 71 (2020), (5), 1659-1667.

Correspondence: Email - kmukhtar@uab.edu

JOURNAL OF INTEGRATED OMICS
A METHODOLOGICAL JOURNAL
HTTP://WWW.JIOMICS.COM

SPECIAL ISSUE: SELECTED ABSTRACTS OF THE IV INTERNATIONAL CAPARICA CONFERENCE IN SPLICING (SPLICING 2021)

An Activator of G Protein-Coupled Receptor and MEK1/2-ERK1/2 Signaling Suppresses HIV-1 Replication by Altering Viral RNA Processing

Raymond W. Wong[1*], Ahalya Balachandran[2], Peter K. Cheung[3], Ran Cheng[2], Qun Pan[2,4], Peter Stoilov[5], P. Richard Harrigan[3,6], Benjamin J. Blencowe[2,4], Donald R. Branch[1,7,8,9,10], and Alan Cochrane[2,10]

[1]Dept. of Medicine, University of Toronto, Canada; [2]Dept. of Molecular Genetics, University of Toronto, Canada; [3]British Columbia Centre for Excellence in HIV/AIDS, Canada; [4]Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Canada; [5]Dept. of Biochemistry, West Virginia University, United States of America; [6]Dept. of Medicine, The University of British Columbia, Canada; [7]Dept. of Laboratory Medicine and Pathobiology, University of Toronto, Canada; [8]Division of Advanced Diagnostics, Infection and Immunity Group, Toronto General Hospital Research Institute, Canada; [9]Centre for Innovation, Canadian Blood Services, Canada; [10]Institute of Medical Science, University of Toronto, Canada

Available Online: 31 December 2021

ABSTRACT

The ability of HIV-1 to evolve resistance to combined antiretroviral therapies (cARTs) has stimulated research into alternative means of controlling this infection. We assayed >60 modulators of RNA alternative splicing to identify new inhibitors of HIV-1 RNA processing—a segment of the viral lifecycle not targeted by current drugs—and discovered compound N-[4-chloro-3-(trifluoromethyl)phenyl]-7-nitro-2,1,3-benzoxadiazol-4-amine (5342191) as a potent inhibitor of both wild-type (Ba-L, NL4-3, LAI, IIIB, and N54) and drug-resistant strains of HIV-1 (IC50: ~700 nM) with no significant effect on cell viability at doses tested. 5342191 blocks expression of four essential HIV-1 structural and regulatory proteins (Gag, Env, Tat, and Rev) without affecting total protein synthesis of the cell. This response is associated with altered unspliced (US) and singly-spliced (SS) HIV-1 RNA accumulation (~60% reduction) and transport of these RNAs to the cytoplasm (loss of Rev; Fig. 1) whereas parallel analysis of cellular RNAs revealed alterations of less than 0.7% to host alternative splicing events (0.25-0.67% by ≥ 10-20%), gene expression (0.01-0.46% by ≥ 2-5 fold), and protein abundance (0.02-0.34% by ≥ 1.5-2 fold). Decreased expression of Tat, but not Gag or Env, upon 5342191 treatment was reversed by a proteasome inhibitor, suggesting that this compound alters the synthesis/degradation of this key viral factor. Consistent with an affect on HIV-1 RNA processing, 5342191 treatment of cells altered the abundance and phosphorylation of serine/arginine-rich splicing factor (SRSF) 1, 3, and 4. Despite the activation of several intracellular signaling pathways by 5342191 (Ras, MEK1/2-ERK1/2, and JNK1/2/3), inhibition of HIV-1 gene expression by this compound could be reversed by pretreatment with either a G-protein α-subunit inhibitor or two different MEK1/2 inhibitors (confirming the importance of this inhibitory signal)[1,2]. These observations demonstrate enhanced sensitivity of HIV-1 gene expression to small changes in host RNA processing and highlights the potential of modulating host intracellular signaling as an alternative approach for controlling HIV-1 infection[1,3].

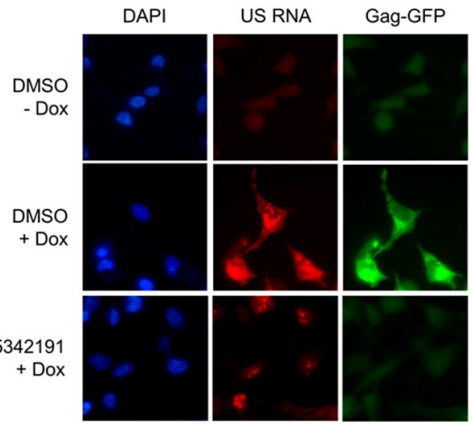

Figure 1 | Trafficking of US RNAs (labeled with Texas Red) detected by fluorescent in situ hybridization (representative of n ≥ 3). Nuclei were detected by DAPI stain, Gag-GFP expression by GFP, and images captured at 630x magnification.

Keywords: HIV-1 RNA processing, alternative splicing; intracellular signaling; small molecule inhibitor; gene expression; antiretroviral therapies

Correspondence: Email - rw.wong@mail.utoronto.ca; alan.cochrane@utoronto.ca

Acknowledgments:

This work was published in Wong et al. (2020)[4] and supported by a Canadian Institutes of Health Research (CIHR) Fellowship - Priority Announcement: HIV/AIDS Research Initiative (158250) to RWW, CIHR Doctoral Award – Frederick Banting and Charles Best Canada Graduate Scholarship (128767) to RWW, CIHR Operating Grant - Priority Announcement: HIV/AIDS Research Initiative (HOP-134065) to AC, National Institutes of Health Grant (R01, EY025536) to PS, and Ontario Graduate Scholarship to AB.

References:

[1] R.W. Wong, C.A. Lingwood, M.A. Ostrowski, T. Cabral, A. Cochrane, Scientific Reports 8 (2018) 850.

[2] R.W. Wong, A. Balachandran, M.A. Ostrowski, A. Cochrane, PLoS Pathogens 9 (2013) e1003241.

[3] S.D.S. McCarthy, D. Leontyev, P. Nicoletti, B. Binnington, H.N. Kozlowski, M. Ostrowski, A. Cochrane, D.R. Branch, R.W. Wong, Journal of Acquired Immune Deficiency Syndromes 82 (2019) 407–415.

[4] R.W. Wong, A. Balachandran, P.K. Cheung, R. Cheng, Q. Pan, P. Stoilov, P.R. Harrigan, B.J. Blencowe, D.R. Branch, A. Cochrane, PLoS Pathogens 1 (2020) e1008307.

JOURNAL OF INTEGRATED OMICS
A METHODOLOGICAL JOURNAL
HTTP://WWW.JIOMICS.COM

SPECIAL ISSUE: SELECTED ABSTRACTS OF THE IV INTERNATIONAL CAPARICA CONFERENCE IN SPLICING (SPLICING 2021)

Structural and functional annotation of alternative proteomes derived from splicing

Zarifa Osmanli[1,2*], Theo Falgarone[1], Jeremy Leclercq[1], Andrey V. Kajava1

[1]University Group of Structural Bioinformatics and Molecular Modelling, CRBM, CNRS, University of Montpellier, Montpellier, France; [2]Integrative Biology laboratory, Biophysics Institute, ANAS, Baku, Azerbaijan.

Available Online: 31 December 2021

ABSTRACT

The dramatic growth of genomic data presents new challenges for scientists. Making sense of millions of protein sequences requires information about their 3D structure as well as their evolutionary and functional relationships. In line with this tendency, we are developing a pipeline of bioinformatics tools for the large-scale, structural and functional annotation of proteomes. Based on the amino acid sequences, the pipeline is able to detect all principal structural states of proteins including globular domains, transmembrane regions, intrinsically disordered regions, low complexity regions, tandem repeats and aggregation-prone motifs [1]. The objective of this work is to apply this pipeline for the analysis of alternative proteomes and to compare structural states of canonical products of translation with ones obtained by alternative splicing, alternative promoter usage or frameshifting [2,3]. We expected that the large-scale, structural-functional annotation of alternative proteomes derived from alternative splicing may provide an insight into the underestimated coding potential of genomes and bring new data about the regions under increased evolutionary pressure. For this purpose, we selected canonical and non-canonical proteomes of 58 eukaryotic species from the UniProt database, which include mammalians, other vertebrates, invertebrates, plants, fungi and other unicellular eukaryotes. The canonical proteome is a complete set of proteins derived from splicing, while the noncanonical/alternative proteome is a set of proteins that are products of alternative splicing and alternative promoter usage or frameshifting. Our comparative analysis revealed some differences in the structural states of proteins from these two sets. For example, the analysis suggests that isoforms, which are derived from alternative splicing, have higher coverage of aggregation-prone regions than the canonical proteins (Figure 1). Moreover, products of splicing and alternative splicing events differ in the coverage of intrinsically disordered regions. These and the other results of our analysis will be present.

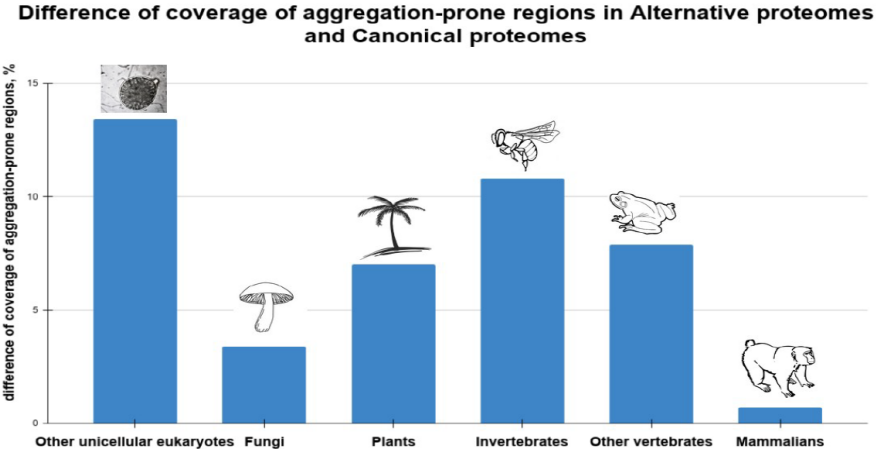

Figure 1 | Difference of coverages of aggregation-prone regions in alternative and canonical proteomes of 58 eukaryotic species, which are grouped in 6 subsets: mammalians, other vertebrates, invertebrates, plants, fungi and other unicellular eukaryotes. Aggregation-prone regions are predicted by using ArchCandy program [1].

Keywords: Dual coding regions; Alternative splicing; Alternative proteome; Aggregation-prone motifs; Structural-functional annotation

Correspondence: Email - zarifa.osmanli@crbm.cnrs.fr

Acknowledgments:

We thank Azerbaijan National Academy of Sciences and CBS2 doctoral school of the University of Montpellier for funding of this project.

References:

[1] A.B. Ahmed, N.Znassi, M.T. Château, A.V. Kajava, Alzheimers & Dementia 11 (2015) 681-690.

[2] E. Klerk, P. Hoen, P.A.C.T. Hoen, Trends in Genetics 31 (2015) 128-139.

[3] P. Wang, B. Yan, J. Guo, C. Hicks, Y. Xu, Proc Natl Acad Sci USA 102 (2005) 18920–18925.

JOURNAL OF INTEGRATED OMICS
A METHODOLOGICAL JOURNAL
HTTP://WWW.JIOMICS.COM

SPECIAL ISSUE: SELECTED ABSTRACTS OF THE IV INTERNATIONAL CAPARICA CONFERENCE IN SPLICING (SPLICING 2021)

An alternative splicing mechanism generates novel translation initiation sites, producing a novel extracellular form of FUS in amyotrophic lateral sclerosis (ALS)/ motor neuron disease (MND).

Julie Atkin[1]*, Marta Vidal[1], Sonam Parakh[1], Anna Konopka[1] , Mariana Brocardo[1]

[1] Macquarie Centre for MND Research, Macquarie University, Sydney, Australia

Available Online: 31 December 2021

ABSTRACT

An alternative splicing mechanism has been recently described that produces a gene transcript which differs from the canonical transcript by deletion of a length of sequence not divisible by three. Translation can then be rescued by an alternative start codon, resulting in a predicted protein in which the amino terminus differs markedly in sequence from the canonical protein because it is translated from an alternative reading frame. Whilst thousands of protein splice variants are known, these isoforms have been overlooked and are under-represented in current databases. Many of these isoforms also result in alterations in localisation compared to the canonical version of each protein.

Amyotrophic Lateral Sclerosis (ALS), also known as motor neuron disease (MND), is a fatal and rapidly progressing neurodegenerative disorder. Fused in sarcoma (FUS) is an important pathological protein in in both familial and sporadic (90% of all cases) forms of ALS, and mutations in FUS account for a particularly aggressive, juvenile form of the disease. Hence, elucidating the pathological role of FUS in ALS is crucial for understanding this disorder. FUS is an DNA/RNA binding protein with important cellular functions in RNA metabolism and the DNA damage response. However, in ALS, FUS mis-localises from its normal location in the nucleus to the cytoplasm, where it forms misfolded inclusions, and both loss and gain of function mechanisms are implicated in pathophysiology. While the mechanisms underlying neurodegeneration in ALS are not fully understood, defects in alternative splicing have been previously associated with disease.

Here we have identified a novel isoform of FUS, 'EC-FUS', produced by this novel mechanism of alternative splicing. This new isoform bears a unique N-terminus containing a secretory leader peptide, hence EC-FUS is present in the extracellular space, unlike canonical FUS which is normally intracellular. Importantly, EC-FUS is the only major pathological protein linked to ALS/FTD that is normally localised extracellularly. EC-FUS shares its C-terminus with canonical FUS, where most of the ALS mutations are located. We have validated the expression of EC-FUS at both the mRNA and protein level and have detected it in multiple tissues. Furthermore, expression of the mutated form of EC-FUS produces features characteristic of ALS, including misfolded protein inclusions. In addition, we have found that the levels of EC-FUS are dysregulated in sporadic ALS patient lumbar spinal cords, cerebrospinal fluid, and serum, linking this isoform to ALS pathology. The discovery of a new alternatively spliced isoform of FUS provides novel insights into the pathogenic mechanisms implicated in ALS/MND.

Keywords: Alternative splicing; ALS; motor neuron disease; FUS

Correspondence: Email - julie.atkin@mq.edu.au

JOURNAL OF INTEGRATED OMICS
A METHODOLOGICAL JOURNAL
HTTP://WWW.JIOMICS.COM

SPECIAL ISSUE: SELECTED ABSTRACTS OF THE IV INTERNATIONAL CAPARICA CONFERENCE IN SPLICING (SPLICING 2021)

Exploring the effect of the pro-inflammatory microenvironment on the expression of the tumor-related RAC1B splice variant in colorectal cancer cells

Cláudia Bessa[1,2,#*], Joana F S Pereira[1,2,#], Vânia Gonçalves[1,2], Paulo Matos[1,2], Peter Jordan[1,2]

[1]Department of Human Genetics, National Institute of Health 'Dr. Ricardo Jorge', Lisbon, Portugal; [2]BioISI–Biosystems & Integrative Sciences Institute, Faculty of Sciences, University of Lisbon, Portugal. [#]equal contribution

Available Online: 31 December 2021

ABSTRACT

An inflammatory tumor cell microenvironment has been identified as a critical tumor-promoting condition providing survival signals to which cancer cells respond with changes in their gene expression. Interestingly, alternative splicing (AS) is one key gene regulatory mechanism that responds to extracellular signals directly affecting cancer progression [1, 2]. For example RACB1, a RAC1 AS variant, previously identified by our group in a subset of BRAF-mutated colorectal tumors [3], was found increased in colon mucosa under inflammatory conditions, both in samples from inflammatory bowel disease patients or following experimentally-induced acute colitis in a mouse model [4].

Based on these findings, the main goal of this work was to determine which pro-inflammatory signals from stromal cells lead to an increase in RAC1B expression levels in colorectal cancer (CRC) cells. For that, we use a more physiologically relevant culture model consisting of a 2D polarized monolayer of Caco-2 CRC cells grown on porous membranes, and then co-cultured with stromal cells, including fibroblasts, monocytes and macrophages. RAC1B expression was analyzed in Caco-2 cells by RT-qPCR, Western blot and confocal fluorescence microscopy. In fact, co-culture experiments revealed that the combined presence of fibroblasts and/or M1 macrophages induced a transient increase in RAC1B mRNA and protein levels in Caco-2 cells, accompanied by a loss of epithelial organization. Besides RAC1B, a panel of 6 additional pairs of tumor-related AS variants was analyzed, being observed an alteration of the expression ratios in 3 out of the 6 variants tested. Moreover, using a human inflammation array, we were able to identify from the conditioned co-culture media three cytokines that associated with increased RAC1B: IL-1b, IL-6 and IL-8. Remarkably, the incubation of polarized Caco-2 cells with described purified cytokines was sufficient per se to trigger an increase in RAC1B expression in a dose-dependent manner. As a whole, our data indicate that pro-inflammatory signals can modulate AS in colon epithelial cells, particularly leading to an increase in RAC1B levels. Future identification of the interleukin-regulated signaling pathways involved in altered AS should provide details on how these talk to the splicing machinery, and may represent a promising approach to discovery novel therapeutic targets for suppressing oncogenic splice variants.

Keywords: Pro-inflammatory microenvironment; Colorectal cancer; RAC1B; Cytokines

Acknowledgments:

Work in the authors' laboratory was supported by Fundação para a Ciência e Tecnologia (FCT) (through grant UID/MULTI/04046/2019 to Research Unit BioISI and fellowship BD/109162/2015 to JFP) and by the Portuguese association for intestinal inflammatory disease (grant GEDII 2013 to PJ).

References:

[1] S. Bonomi, S. Gallo, M. Catillo, D. Pignataro, G. Biamonti, C. Ghigna, Int J Cell Biol (2013), 962038.,
[2] C.J. David, J.L. Manley, Genes Dev 24 (2010), 2343-64.
[3]P. Matos, C. Oliveira, S. Velho, V. Goncalves, T.L. da Costa, M.P. Moyer, R. Seruca, P. Jordan, Gastroenterology 135 (2008), 899-906
[4] P. Matos, L. Kotelevets, V. Goncalves, A.F. Henriques, A. Henriques, P. Zerbib, M.P. Moyer, E. Chastre, P. Jordan, Neoplasia 15 (2013), 102-111.

Correspondence: Email - claudia.bessa@insa.min-saude.pt

JOURNAL OF INTEGRATED OMICS
A METHODOLOGICAL JOURNAL
HTTP://WWW.JIOMICS.COM

SPECIAL ISSUE: SELECTED ABSTRACTS OF THE IV INTERNATIONAL CAPARICA CONFERENCE IN SPLICING (SPLICING 2021)

CTCF & Alternative Splicing: Look who joined the party!

John E. J. Rasko[1]*

[1]Sydney Medical School, University of Sydney; [2]Gene and Stem Cell Therapy Program, Centenary Institute; [3]Cell & Molecular Therapies, Royal Prince Alfred Hospital, Sydney, Australia

Available Online: 31 December 2021

ABSTRACT

At the 3rd International Caparica Conference on Splicing we discussed the importance of intron retention (IR) as a once-neglected form of alternative splicing. Since our original demonstration that IR is a conserved mechanism of gene expression control in mammalian biology (Wong et al Cell, 2013; Schmitz et al Genome Biology 2017) we have demonstrated epigenetic control mechanisms and broad relevance to normal and cancer biology. In this presentation we highlight the role of the "Master weaver of the Genome", CTCF, and provide evidence of its increasingly compelling role in co-ordinating alternative splicing (Alharbi et al NAR 2021).

CTCF is a ubiquitously expressed, conserved zinc finger transcription factor with ~100,000 target sites identified genome-wide. It is widely acknowledged as being central to the 3D organisation of chromatin architecture in the nucleus. We have shown that CTCF acts as a tumour suppressor gene (Rasko et al, Cancer Res 2001, Tiffen, Bailey et al, Int J Can 2013). We now perceive that CTCF has a role in orchestrating alternative splicing via diverse mechanisms, including: transcriptional elongation, DNA methylation, chromatin architecture, histone modification, and regulation of splicing factor expression and assembly. We examined the effect of Ctcf haploinsufficiency on gene expression and AS in five tissues from Ctcf hemizygous (Ctcf+/-) mice. An increase in IR was observed in Ctcf+/- liver and kidney, but not in brain, kidney or spleen. In liver, this specifically impacted genes associated with cytoskeletal organisation, splicing and metabolism. Ctcf binding sites were greatly enriched upstream 200–50,000 nt, but not downstream or within introns, of differentially expressed transcripts with retained, compared to non-retained, introns. CTCF not only co-ordinates gene expression regulation but modulates transcriptomic complexity.

References:

Adel B. Alharbi, Ulf Schmitz, Charles G. Bailey and John E. J. Rasko. Nucleic Acids Research, (2021) 1–15

Adel B. Alharbi , Ulf Schmitz , Amy D. Marshall , Darya Vanichkina , Rajini Nagarajah , Melissa Vellozzi , Justin J. L. Wong , Charles G. Bailey & John E. J. Rasko. RNA Biology (2021)

Correspondence: Email - j.rasko@centenary.org.au

JOURNAL OF INTEGRATED OMICS
A METHODOLOGICAL JOURNAL
HTTP://WWW.JIOMICS.COM

SPECIAL ISSUE: SELECTED ABSTRACTS OF THE IV INTERNATIONAL CAPARICA CONFERENCE IN SPLICING (SPLICING 2021)

Acetylation of Histones H4 and H2A.Z by the NuA4 acetyltransferase is Required for Appropriate RNA Splicing in *Saccharomyces cerevisiae*

Tracy L. Kress[1], Mengjia Lin[1], Julie Burkett[1], Olivia Marino[1], Jessica Kopew[1], Adriana Mendizabal[1], Lawrence Palfini[1], Thomas Gunning[1], Rose Maisner[1], Michael A. Gildea[2], Jeffrey A. Pleiss[2]

[1]The College of New Jersey, Department of Biology, Ewing, NJ; [2]Cornell University, Department of Molecular Biology and Genetics, Ithaca, NY.

Available Online: 31 December 2021

ABSTRACT

Efficient and precise gene expression depends on effective coordination between RNA splicing and transcription1, yet the mechanistic details that underlie this coordination have remained relatively unexplored. Recent studies in yeast and metazoan [1, 2] have identified connections between chromatin modification, which alters chromatin dynamics to regulate transcription, and splicing. The NuA4 histone acetyltransferase works the with Swr1 chromatin remodeling enzyme to promote transcription [3-5]. First, NuA4 acetylates histone H4, which is required for the recruitment of Swr1. Next, Swr1 deposits the histone variant H2A.Z, which is subsequently acetylated by NuA4. Interestingly, in Saccharomyces cerevisiae, both NuA4 and Swr1 play a key role in the transcription of the ribosomal protein genes [7], which make up a large fraction of the intron-containing genes. Thus, NuA4, Swr1, and H2A.Z are excellent candidates for coordinating transcription with splicing. We recently identified a role for H2A.Z in splicing [6] and here we present data that support a role for NuA4 in splicing. Interestingly, data from our lab suggest that NuA4 may have functions in RNA splicing that are independent of H2A.Z deposition. Using directed genetic screens we identified interactions between splicing factor gene mutations and mutations that alter H4 and H2A.Z acetylation, including point mutations in H4 and H2A.Z and point mutations that inhibit the catalytic activity of NuA4. Additionally, using whole genome RNA sequencing (MPE-seq [8]) and RT-qPCR we found that these mutations alter the splicing of a subset of RNAs. Our data suggest that NuA4 and the acetylation of H4 and H2A.Z play a role in regulating RNA splicing and we are currently performing chromatin immunoprecipitation experiments to determine whether NuA4 activity or acetylation of H4 and H2A.Z impacts the association of splicing proteins with nascent RNAs during transcription. Together, our data support a model in which NuA4 interacts with the splicing machinery to coordinate transcription with splicing.

Acknowledgments:

We thank the Rine and Cote laboratories for strains and the Pleiss laboratory for assistance with and analysis of the MPE-seq experiments. We thank members of the Kress and Pleiss laboratories for helpful discussion. This work was supported by grants to T.L.K. from the Research Corporation for the Advancement of Science (Cottrell College Science award no. 20186) and the National Institutes of Health (R15GM122026).

References:

[1] Bentley DL. Nat Rev Genet 15 (2014), 163-175.
[2] Luco RF, et al. Cell 144 (2011), 16-26.
[3] Allard S, et al. J EMBO 18 (1999), 5108-5119.
[4] Altaf M, et al. J Biol Chem 285 (2010) 15966-15977.
[5] Krogan NJ, et al. Mol Cell 12 (2003), 1565-1576.
[6] Neves, LT, et al. Genes Dev 31 (2016), (7), 702-717.
[7] Reid JL, et al. Mol Cell 6 (2000), 1297-1307.
[8] Xu, Hansen, et al. Nat Methods 16 (2019), (1), 55-58.

Correspondence: Email - kress@tcnj.edu

JOURNAL OF INTEGRATED OMICS
A METHODOLOGICAL JOURNAL
HTTP://WWW.JIOMICS.COM

SPECIAL ISSUE: SELECTED ABSTRACTS OF THE IV INTERNATIONAL CAPARICA CONFERENCE IN SPLICING (SPLICING 2021)

Purification of fibrinogen gamma A/gamma prima (γA/γ') by fast protein liquid chromatography and its effect on fibrin polymerization

Miriam Elena Cantero Guevara[1]*, Bárbara Cardinali[2], Rita Marchi[3]

[1]Universidad de Córdoba, Colombia; [2]Hospital Policlinico San Martino, Italia; [3]Afiliación, País. 4Instituto Venezolano de Investigaciones Científica, Venezuela

Available Online: 31 December 2021

ABSTRACT

Fibrinogen is a homodymeric plasma protein of 340 kDa consisting of 6 polypeptide chains 2 (Aα, Bβ and γ) linked together by disulfide bonds [1]. A common variant of the γ chain called fibrinogen γ' (γA/γ'), is produced by an alternative splice in intron 9, resulting in the loss of exon 10 and retention of part of intron 9 in mRNA [2]. This alternative chain has the last 4 C-terminal residues replaced by 20 different amino acids, with a high proportion of negatively charged residues. Fibrinogen γA/γ' has an average plasma concentration of 8% to 15% of plasma fibrinogen [2]. Clinically it is associated with opposite effects on thrombosis in arterial and venous circulation [3]. Fibrinogen γA/γ' has been shown to have unique properties that modify the structure of the fibrin clot [4]. The objective of this work was to design a more efficient method of purification of fibrinogen γA/γ' in relation to those described in current literature, from human plasma and to evaluate the kinetics of fibrin polymerization on the surface of HMEC-1 cells with different concentrations of fibrinogen γA/γ'. The purification of fibrinogen γA/γ' was performed from the total fibrinogen obtained from the plasma of apparently healthy individuals by precipitation with β-alanine, with some modifications [5], by means of separation by fast liquid chromatography of proteins, using an anion exchange column. The protein purified by ELISA was quantified and confirmed by Western blot. The kinetics of fibrin formation on HMEC-1 endothelial cells was studied by turbidimetry. Fibrin clots were formed at different concentrations of fibrinogen γA/γ' (3%, 10% and 30%) and the optical density (DO) was read every 2 min at 350 nm in an Infinite 200 M (Tecan, Vienna,

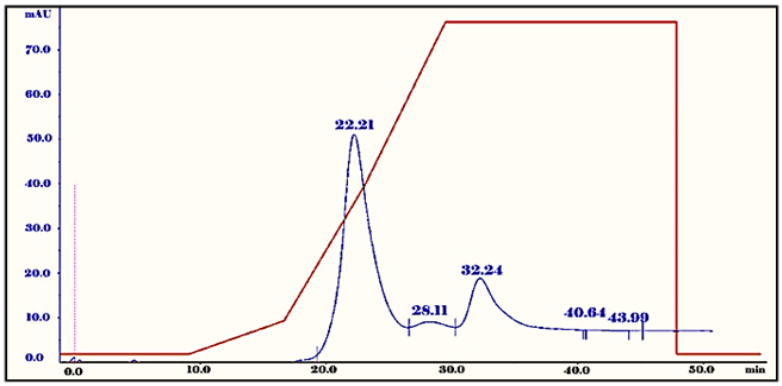

Figure 1 | Elution profile of Fg γA/γ' by fast protein separation liquid chromatography (FPLC). The column used was 20 mL (HiPrepTM DEAE FF 16/10 (GE Healthcare), flow rate 4 mL/min. The red trace corresponds to the gradient used to elute Fg γA/γ'.

Correspondence: Email - miriamcantero@correo.unicordoba.edu.co

Table 1 | Polymerization of fibrin at different concentrations of Fibrinogen γA/γ' on HMEC -1 cells. The results are expressed as the average ± standard error (SE).
*p <0.05: Fibrin formed in the absence of or over a culture of HMEC-1 at 30% γ/γ', compared to that of 3 and 10% γ/γ'

Fibrinogen γA/γ' (%)	Maximum Absorbance: MaxAbs (mOD)		Slope (mOD / min)		Latency time (min)	
	With cells	With no cells	With cells	With no cells	With cells	With no cells
3	440 ± 20	390 ± 10	60 ± 2	45 ± 10	0	0
10	460 ± 50	340 ± 30	60 ± 5	49 ± 6	3	3
30	320 ± 30*	230 ± 30 *	36 ± 2 *	20 ± 6 *	3	3

Austria). The experiments were conducted in triplicate in three independent trials. The kinetics of fibrin formation was characterized by measuring delay time (min), slope (mOD/min) and maximum absorbance (AbsMax, mOD). The γA/γ' fibrinogen purification method showed advantages compared to classical separation methods. For example, the use of smaller sample quantities that could be separated into their components with a high degree of purity in less time (30 min). Two main peaks were obtained, 1 at pH 7.0 : γA/γ' (2.20 ± 0.07 g/L) and 2 at pH 4.5: γA/γ' (0.25 ± 0.05 g/L) (Figure 1). The kinetics of fibrin formation in HMEC-1 with 3% and 10% fibrinogen γA/γ' were similar with respect to the no cells condition. However, with 30% fibrinogen γA/γ' the slope and maximum absorbance (AbsMax) were approximately 50% and 30% lower than 3 and 10% (Table 1). The latency time was equal for 10% and 30% fibrinogen γA/γ', regardless of the presence of the cells. The results obtained allow us to conclude that the technique used to purify the variants of the γ chain of fibrinogen (γA/γA and γA/γ'), is an efficient separation method that allows purifying fibrinogen γA/γ' free of main contaminants, as confirmed by immunoelectrophoresis. Our results demonstrate the pathophysiological impact of increased fibrinogen γA/γ' in fibrin formation on HMEC-1 cells. These results partially explain how high plasma concentrations of fibrinogen γA/γ' increase the risk of thrombosis.

Keywords: Fibrinogen gamma prime (γ'); coagulability; electrophoresis; Fibrin; plasma.

References:

[1] R. F. Doolittle, Z. Yang, I. Mochalkin. Sciences. 936 (2001), (1), 31-43.
[2] D. W. Chung, E. W. Davie. Biochemistry. 23 (1984), (18), 4232-4236.
[3] F. L. Macrae, M. M. Domingues, A. Casini, R. A. S. Ariens. Seminars in Thrombosis and Hemostasis. 42 (2016), (4), 344-55.
[4] S. Kattula, J. R. Byrnes, A. S. J. A. Wolberg. Fibrinogen and fibrin in hemostasis and thrombosis. 37 (2017), (3), e13-e21.
[5] E. Jakobsen, P. Kierulf. Thrombosis research. 3 (1973), (2), 145-159.

JOURNAL OF INTEGRATED OMICS
A METHODOLOGICAL JOURNAL
HTTP://WWW.JIOMICS.COM

SPECIAL ISSUE: SELECTED ABSTRACTS OF THE IV INTERNATIONAL CAPARICA CONFERENCE IN SPLICING (SPLICING 2021)

hnRNP proteins A1 and H1 regulate the alternative splicing of HTLV-1 antisense gene HBZ

Julie Tram[1], Benoît Barbeau[2], Jean-Michel Mesnard[1], Jean-Marie Péloponèse[1]*

[1]IRIM, CNRS, Université de Montpellier, Montpellier, France; [2]Department of Biological Sciences, Université du Québec à Montréal, Montréal, Canada

Available Online: 31 December 2021

ABSTRACT

The human retrovirus HTLV-1 (Human T-cell Leukemia type 1) is the etiological agent of an extremely aggressive cancer named ATL (Adult T-cell Leukemia). Series of evidence have linked the transformation of HTLV-1-infected T cells to the expression of the viral gene hbz (HTLV-1 Basic leucine Zipper transcription factor)1. HBZ expression is initiated by a promoter present in the 3'-LTR2,3 (long terminal repeats) of the provirus and it has been established that it is the only viral protein expressed in every ATL cells1. HBZ pre-mRNA undergoes alternative splicing producing three different mRNAs and thus encoding for three different HBZ isoforms: usHBZ (unspliced), HBZ_SP1 and HBZ_SP2 (spliced)3. Analysis showed that the spliced form HBZ_SP1 is the most expressed isofom in ATL cells. However, the role of each isoform in HTLV-1-linked pathogenesis and the precise mechanisms regulating the alternative splicing phenomenon are yet to be fully understood. To understand this mechanism, we first assessed the expression profile of over twenty different splicing factors (SF), with a predicted binding site on the pre-mRNA of HBZ, both in ATL cell lines and in HTLV-1 infected patients' samples. We observed that a majority of those SF were mis-regulated in the ATL context when compared to the controls. After analysis, we chose to mainly focus on two hnRNPs proteins (heterogeneous nuclear RiboNucleoParticles): hnRNP A1 and H1. Using an RNA Immunoprecipitation (RIP) approach, we first confirmed that hnRNP A1 and H1 are indeed binding onto HBZ pre-mRNA. Next, we investigated the effect of those two splicing factors on HBZ alternative splicing using the K30-as-Luc minigene-assay. In this construction, the second exon of HBZ was replaced by a firefly-luciferase encoding gene. Our results revealed that overexpression of hnRNP A1 efficiently inhibits HBZ splicing. Meanwhile, shRNA-mediated depletion of hnRNP A1, as well as overexpression of hnRNP H1 have the opposite effect, meaning both promote the production of HBZ_SP1 isoform. Mutations of splicing donor and acceptor sites on HBZ pre-mRNA are ongoing to assess in further details the mechanism of HBZ alternative splicing regulation by hnRNP A1 and H1.

ATL prognostic is very pessimistic with an average survival rate extremely low. Until now, therapeutic strategies used to treat ATL were proven inefficient and no strategy targeting directly HBZ is available. Deciphering HBZ alternative splicing regulation mechanism could allow the development of new therapeutic approaches to restore chemosensitivity and fight ATL.

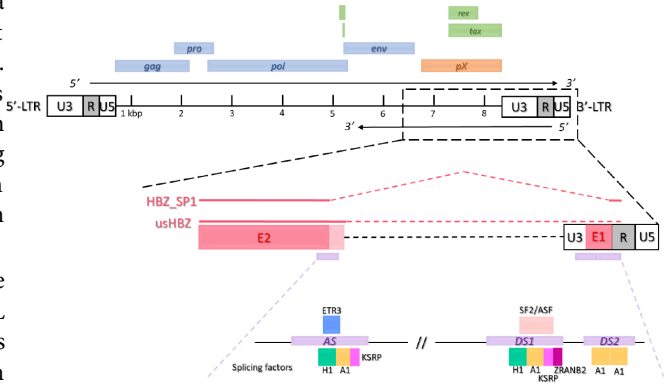

Figure 1 | Recapitulative diagram of HTLV-1 genome. HTLV-1 antisense gene HBZ pre-mRNA is alternatively spliced into three mRNA encoding for three protein isoforms.

References:

Satou, Y., Yasunaga, J. -i., Yoshida, M. & Matsuoka, M. Proc. Natl. Acad. Sci. 103, 720–725 (2006).

Gaudray, G. et al. J. Virol. 76, 12813–12822 (2002).

Cavanagh, M.-H. et al. Retrovirology 3, 15 (2006).

Correspondence: Email - jean-marie.peloponese@irim.cnrs.fr

JOURNAL OF INTEGRATED OMICS
A METHODOLOGICAL JOURNAL
HTTP://WWW.JIOMICS.COM

SPECIAL ISSUE: SELECTED ABSTRACTS OF THE IV INTERNATIONAL CAPARICA CONFERENCE IN SPLICING (SPLICING 2021)

The complex splicing of the alpha-synuclein gene and its clinical meaning

Katrin Beyer[1]*, Aintzane Urbizu[2], Dolores Vilas[3], Lourdes Ispierto[3], Ramiro Alvarez[3]

[1]Dept Pathology, Germans Trias i Pujol Research Institute, Badalona, Barcelona, Spain; [2]Dept. Pathology, University Hospital Germans Trias i Pujol; [3]Dept. Neurology, University Hospital Germans Trias i Pujol, Badalona, Barcelona, Spain.

Available Online: 31 December 2021

ABSTRACT

Alpha-synuclein plays a major role in the development of synucleinopathies. These include Parkinson's disease (PD) which is the most frequent movement disorder, dementia with Lewy bodies (DLB), the second most frequent cause of degenerative dementia after Alzheimer's disease, and multiple system atrophy. Abnormal alpha-synuclein oligomerization and aggregation are central events in these diseases and start up to 20 years before clinical symptoms are evident [1]. Alpha-synuclein is encoded by the 5 coding exon containing SNCA gene that undergoes complex splicing events. Exon-skipping involves two coding exons which can be spliced out individually, or simultaneously. The resulting transcripts have been named by the size of the corresponding protein isoform: SNCA140 gives rise to the main protein, SNCA126 lacks exon 4, SNCA112 exon 6 and SNCA98 both exons 4 and 6. Since exon 4 encodes the most aggregation-prone part of the protein, SNCA126 is expected to have diminished aggregation properties representing a protective isoform. In contrast, exon 6 is located near the C-terminal corresponding to the most soluble part of the protein. Consequently, SNCA112 shows enhanced aggregation properties and it has been proposed that this isoform could participate in the very first steps during the alpha-synuclein aggregation process [2,3]. Furthermore, the SNCA genes contains at least 4 alternative initial exons, 1 and 2a-c, of which exon 2c (SNCAtv4) is included only in brain. Of the transcript variants containing either exon 1 (SNCAtv3), 2a (SNCAtv1) or 2b (SNCAtv2), SNCAtv1 is the main transcript, followed by SNCAtv2. Expression of SNCAtv1 and SNCAtv2 is regulated by hyper-methylation of two CpG-islands identified in their promoter region. SNCAtv3 in turn, seems to be regulated by GATA-transcription factors [reviewed in 2]. The analysis of the expression levels of SNCA transcripts in brain of DLB and PD compared to controls revealed, that SNCAtv1 and SNCA112 are overexpressed in DLB, and SNCAtv2 in PD temporal cortices. These results underline that the increase of alpha-synuclein levels in synucleinopathies are accompanied by the increase of the aggregation-prone isoform SNCA112 in vulnerable brain areas [3]. The analysis of SNCA transcript expression levels in blood showed the diminution of all SNCA transcripts in DLB, and of SNCAtv1 and SNCAtv2 in PD with disease onset before 70 years. SNCAtv3 is driven by its own promoter, and opposite expression levels were found in early DLB and PD, so that SNCAtv3 could represent a suitable biomarker for the early diagnosis of DLB [3]. Correlation between blood transcript levels and disease duration was opposite in DLB and PD. It could reflect differences in brain alpha-synuclein aggregation rates which are expected to be lower in early-onset PD than in DLB, because the duration of early-onset PD is usually longer than 15 years, and DLB is an aggressive disease with an overall survival of between 6 and 10 years. In conclusion, alternative SNCA splicing has a direct impact on the development of synucleinopathies on one hand, and may reflect the disease course on the other. The usefulness of SNCAtv3 levels as blood biomarker has to be further investigated and validated.

Keywords: Alpha-synuclein, transcript variants, exon skipping, 5'-UTR splicing, 3'UTR splicing

Correspondence: Email - kbeyer@igtp.cat

Acknowledgments:

We thank the Neurological Tissue Bank of the Biobank-Hospital Clinic-Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, and the Institute of Neuropathology Brain Bank, L'Hospitalet Spain for sample and data procurement. This research was funded by Spain's Ministry of Science and Innovation, grant number PI15/00216 and PI18/00276, integrated in the National R + D + I and funded by the ISCIII and the European Regional Development Fund.

References:

[1] K.A. Jellinger, Acta Neuropathol 116 (2008) 1-16.
[2] A. Gámez-Valero, K. Beyer, Genes 9 (2017) 63.
[3] L. Marsal-García, A. Urbizu, L. Arnaldo, J. Campdelacreu, D. Vilas, et al, Intl J Mol Sci 22 (2021) 725.

JOURNAL OF INTEGRATED OMICS
A METHODOLOGICAL JOURNAL
HTTP://WWW.JIOMICS.COM

SPECIAL ISSUE: SELECTED ABSTRACTS OF THE IV INTERNATIONAL CAPARICA CONFERENCE IN SPLICING (SPLICING 2021)

Identification of new splicing variants in genes for cattle milk oligosaccharides biosynthesis

Alessandra Crisá[1]*, Pasquale Marmo[1], Cinzia Marchitelli[1]

[1]Dept 1Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria (CREA), Research Centre for Animal production and Aquaculture (CREA), Monterotondo (Roma), Italy

Available Online: 31 December 2021

ABSTRACT

Milk sialylated oligosaccharides (SOS) play crucial roles in many biological processes and have functional health benefits for consumers [1]. The most abundant free SOS in milk are 3'-sialyllactose (3'-SL) and 6'-sialyllactose (6'-SL) and sialyltransferases are a subset of glycosyltransferases that catalyse the transfer of sialic acid residues to lactose or terminal positions in the oligosaccharide chains of glycoproteins and glycolipids [2]. The work here reported is part of the MIQUALAT project activities. One of the aims of the project consists in the analysis of some free oligosaccharides in milk of four cattle breeds (Holstein, Simmenthal x Holstein, Simmenthal, Podolica) at two lactation stages (60 and 120 days after calving) and in the genetic characterization of the breeds using genotyping by bead chip, RNA-seq and the study of some specific genes associated with the molecules biosynthesis. The analysis of candidate genes was focused on beta-1,4-galactosyltransferase 1 (B4GALT1) and lactalbumin alpha (LALBA) for lactose (the precursor molecule for 3'-SL and 6'-SL biosynthesis), beta-galactoside alpha-2,3-sialyltransferase 4 (ST3GAL4), beta-galactoside alpha-2,3-sialyltransferase 5 (ST3GAL5), beta-galactoside alpha-2,3-sialyltransferase 6 (ST3GAL6) for 3'-SL, beta-galactoside alpha-2,6-sialyltransferase 1 (ST6GAL1) for 6'-SL. Using RNA extracted from milk somatic cells, cDNAs of the different genes have been amplified. Preliminary sequencing results of the cDNAs have identified new splicing variants in ST3GAL4, ST6GAL1 and LALBA genes. In ST3GAL4 a shorter transcript with a skipped exon7, in ST6GAL1 a shorter transcript with a skipped exon 3, in LALBA a longer transcript with the intron 1 retention were found.

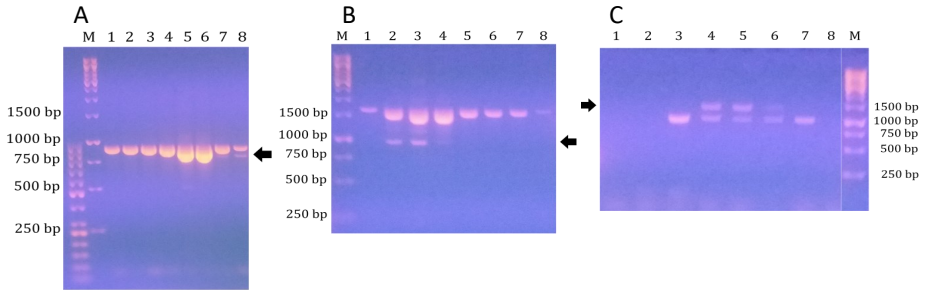

Figure 1 | Agarose gel electrophoresis (1,5%) of ST3GAL4 cDNA (896 bp amplicon), panel A, ST6GAL1 cDNA (1603 bp amplicon), panel B, LALBA cDNA (956 bp amplicon) from the milk somatic cells of eight cows. M = GeneRuler 1 Kb DNA ladder (Thermofisher Scientific). The arrows indicate the spliced or retained fragments.

Keywords: Lactation; Milk; Oligosaccharides; Cattle

Correspondence: Email - alessandra.crisa@crea.gov.it

Acknowledgments:

Acknowledgement: this research was funded by MIPAAF in the national research project MIQUALAT (D.M. 16844/7100/2019)

References:

[1] A. Crisá, Chichester: Wiley (2013) 129–47J .

[2] A. Harduin-lepers, Glycobiology (2010) 29–61.

JOURNAL OF INTEGRATED OMICS
A METHODOLOGICAL JOURNAL
HTTP://WWW.JIOMICS.COM

SPECIAL ISSUE: SELECTED ABSTRACTS OF THE VII INTERNATIONAL CAPARICA CONFERENCE ON ANALYTICAL PROTEOMICS (ICAP 2021)

Plasma β-III tubulin, neurofilament light chain and glial fibrillary acidic protein are associated with neurodegeneration and progression in schizophrenia

Carlos Spuch[1,2*], Daniela Rodrigues-Amorim[1,3], Tania Rivera-Baltanás[1,2], Marta Iglesias-Martínez-Almeida[1,3], Luis Freiría-Martínez[1,3], María Del Carmen Vallejo-Curto[1,3], Cynthia Rodriguez-Jamardo[1,3], Elena de Las Heras[1,2], Carolina Barreiro-Villar[1], Patricia Fernández-Palleiro[1,3], María Álvarez-Ariza[1,2], Marta López-García[1,2,3], Alejandro García-Caballero[1], José Manuel Olivares[1,2]

[1]Galicia Sur Health Research Institute, Spain; [2]CIBERSAM, Spain; [3]University of Vigo, Spain.

Available Online: 31 December 2021

ABSTRACT

Schizophrenia is a progressive disorder characterized by multiple psychotic relapses. After every relapse, patients may not fully recover, and this may lead to a progressive loss of functionality. Pharmacological treatment represents a key factor to minimize the biological, psychological and psychosocial impact of the disorder. The number of relapses and the duration of psychotic episodes induce a potential neuronal damage and subsequently, neurodegenerative processes. Thus, a comparative study was performed, including forty healthy controls and forty-two SZ patients divided into first-episode psychosis and chronic SZ (CSZ) subgroups, where the CSZ sub group was subdivided by antipsychotic treatment. In order to measure the potential neuronal damage, plasma levels of β-III tubulin, neurofilament light chain (Nf-L), and glial fibrillary acidic protein (GFAP) were performed. The results revealed that the levels of these proteins were increased in the SZ group compared to the control group. Moreover, multiple comparison analysis showed highly significant levels of β-III tubulin, Nf-L and GFAP in the subgroup of CSZ clozapine-treated. In conclusion, β-III tubulin, Nf-L and GFAP proteins may be potential biomarkers of neurodegeneration and progression in SZ [1-3].

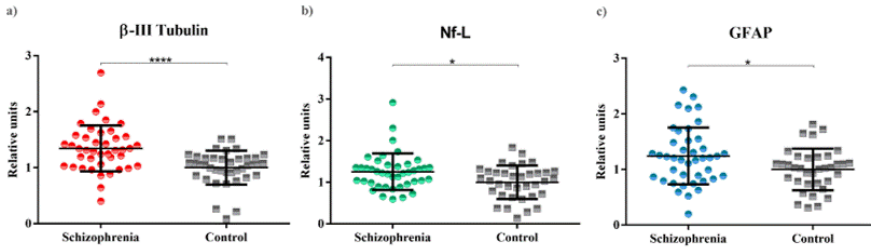

Figure 1 | Plasma levels of β-III tubulin, Nf-L and GFAP proteins in patients with SZ (N=42) and controls (N=40). Scatterplots of plasma levels of structural proteins.

Keywords: Schizophrenia; proteomic; neurodegeneration; biomarkers

Acknowledgments:

Foundation for Science and Technology (FCT) Grant SFRH/BD/135623/2018, the ISCIII Grant P16/00405 and the Ministry of Health, Equality, and Social Policy—Government Delegation for the National Plan on Drugs Grant 2017I054, (ACIS) Grant PRIS2-17 and GAIN Grant IN607C-2017/02, IN607B 2018/17.

References:

[1] Sommer, I. E. et al. npj Schizophr 2 (2016), 16003
[2] Rund, B. R. Nord. J. Psychiatry 63 (2009), 196– 201
[3] Lieberman, J. A. et al. Neuropsychopharmacology 14 (1996), 13S-21S.

Correspondence: Email - cspuch@uvigo.es

REVIEWS

JOURNAL OF INTEGRATED OMICS
A METHODOLOGICAL JOURNAL
HTTP://WWW.JIOMICS.COM

ARTICLE | DOI: 10.5584/jiomics.v11i2.207

Human disease biomarkers: challenges, advances, and trends in their validation

Jemmyson Romário de Jesus[1]*, Marco Aurélio Zezzi Arruda[2,3]*

[1] Research Laboratory in bionanomaterials, LPbio, Chemistry Department, Federal University of Viçosa, UFV, Viçosa, Minas Gerais, Brazil; [2] Spectrometry, Sample Preparation and Mechanization Group, GEPAM, Institute of Chemistry, University of Campinas, UNICAMP, Campinas, Brazil; [3] National Institute of Science and Technology for Bioanalytics.

Received: 20 August 2021 **Accepted:** 05 October 2021 **Available Online**: 31 December 2021

ABSTRACT

Biomarkers are important tools in the medical field, once they allow better prediction, characterization, and treatment of diseases. In this scenario, it is essential that biomarkers are highly accurate. Thus, biomarker validation is an essential part of ensuring the effectiveness of a biomarker. Validation of biomarkers is the process by which biomarkers are evaluated for accuracy and consistency, as well as their ability to inform the condition of health or disease. Although, there is no unique measure that can be used to determine the validity for all biomarkers, there are general criteria that all biomarkers must meet to be useful. In this work, we review the definition of biomarkers and discuss the validity components. We then critically discuss the main methods used to validate biomarkers and consider some examples of biomarkers of the diseases which most killer in the world (cardiovascular diseases, cancer, and viral infections), highlighting the potential biochemical pathways of these biomarkers in the biological system. In addition, we also comment on the omic strategies used in the biomarker discovery process and conclude with information about perspectives in biomarker validation through imaging techniques.

Keywords: *Biomarker, cancer, cardiovascular diseases, viral infection, validation, analytical instrumentation.*

1. Introduction

The development of disease is related to a set of extrinsic factors that involve environmental aspects and lifestyle habits, as well as it may be related to intrinsic factors regulated mainly by genes [1]. It is known that depending on these factors, a great diversity in the clinical expression of the same disease may exist [2]. Obviously, other factors, mainly related to new forms of treatment, can influence the course of the disease, as well as its clinical expression in the biological system. In this sense, biomarkers emerge as dynamic and important tools to understand the cause, type, severity, and treatment monitoring of human diseases [3].

Biomarkers are entities that can be measured experimentally, indicating the normal or pathological condition of an organism, as well as indicating the organism's response to a pharmacological treatment [4]. The identification and validation of different biomarkers of human disease, whether genetic or biochemical, have contributed not only to the diagnosis of diseases related to humans, but also to the understanding of their causes and the response mechanisms to the therapeutic procedure [5]. Biochemical markers (proteins, lipids, metabolites and others) have shown a strong association with different types of diseases, such as: cancer [6–8] , cardiovascular diseases [9,10] and even psychiatric diseases [11–14] and have been widely used as an auxiliary method in the diagnosis of chronic autoimmune and inflammatory diseases, as for example, diabetes [15,16] and degenerative arthritis [17]. In view of technological advances and the application of multidisciplinary research strategies that make it possible to study diseases and their relationship to human health,

***Corresponding author:** J.R. de Jesus, Chemistry Department, Federal University of Viçosa, UFV Viçosa, MG, 36570-900, Brazil. email: jemmyson.jesus@ufv.br M.A.Z. Arruda Chemistry Institute, University of Campinas, UNICAMP Campinas, SP, 13083-970, Brazil. email: zezzi@unicamp.br

genetic markers have been associated with numerous diseases, which confirms the influence of genes on the development of inflammatory and infectious diseases, neoplastic (tumoral) and cardiac [4–6,10].

The development of a biomarker for clinical purposes requires different step of discovery and validation [18,19]. In general, the discovery stage is related to the application of omic strategies, such as: genomics [20], transcriptomics [21], proteomics [22], metabolomics [23] and others that aim to extract and identify possible biomarkers in biological samples (tissue, cell, fluid). The validation step is essential for recognizing a biomolecule as a biological marker. Validation is essentially based on quantitative and qualitative measures to confirm whether the concentration or status of a probable biomarker is significantly different between disease and control, or whether there is a significant response from a therapeutic intervention [5,18]. For a biomolecule to be confirmed as a biological marker it is necessary that it also has important characteristics, such as, reflecting the interaction of the biological system in a qualitative and/or quantitative way with the disease or with a pharmacological agent; be sensitive and specific to the disease or treatment; be qualitatively and quantitatively reproducible, as well as in its analytical measure, present adequate accuracy and precision [19]. Although, in recent times, the biomarker discovery process has gained great emphasis in the scientific field through omic strategies, the validation of such biomolecules has not been seen with the same enthusiasm. However, the validation process is essential for new biomarkers to be introduced in routine clinical practice [5].

In this sense, this review critically discusses the importance of biomarkers in the diagnosis and treatment of human diseases, focusing mainly on cardiovascular diseases, cancer and viral infections, as well as describing the steps to validate such biomolecules, using some of the main validation techniques. This work also comments on the genomic, transcriptomic, proteomic and metabolomic strategies used in the biomarker discovery process.

2. Biomarkers: an overview

The term biomarker (contracted form of biological marker) emerged in the scientific sphere in 1989 as a medical term to indicate biological parameters (for example, the increase or decrease in the concentration of specific enzymes and hormones or the presence of a substance in the biological system) that allow them to be measured quantitatively, serving as an indicator of health and assessments related to physiology such as pregnancy, genetic dysfunction or effects of chemical exposure [4,18]. For the World Health Organization (WHO), biomarkers are defined as any specie, structure or process that can be measured in the body, or in its products and that allows to predict the influence or incidence of disease or results [18]. For the National Institute of Health (NIH) Working Group, biomarkers are a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic process, or pharmacological response to a therapeutic intervention [24]. Although there are different definitions to conceptualize what is a biomarker, in practice, everyone agrees that biomarkers are important tools that can help to understand the cause, diagnosis, progression or regression of symptoms and the outcome of treatment of certain diseases. The main sources of biomarkers are tissues, cells, and biological fluids[4]. However, fluid samples such as, blood (plasma or serum), cerebrospinal fluid (CSF), urine and saliva are the most used in research to discover biomarkers [19]. This fact is justified by the advantages that this type of sample presents, such as, easy accessibility, avoiding the risks of invasive tissue sampling through biopsy, relatively low cost of obtaining it, and its potential for the development of diagnostic/prognostic tests on a large -scale, since biological fluids are used in routine tests [25,26]. The desirable characteristics for a biomarker depend on its application. For the detection of human disease, for example, a biomarker must have (i) high sensitivity and specificity, (ii) adequate precision and accuracy, (iii) be robust and present low cost in its clinical tests [4,18,19].

2.1 Types of biomarkers

According to Wishart et al. [4], biomarkers can be classified as: exposure biomarkers, widely used in predicting risks of toxicological contamination; and effect (or disease) biomarkers used in the prognosis, diagnosis and monitoring (progression or regression) of a given disease [4]. Exposure biomarkers are used to assess and confirm individual or group exposure to a specific substance, establishing a link between external exposure and the quantification of internal exposure in the biological system [5]. Exposure biomarkers can reflect the amount absorbed immediately before sampling, such as the concentration of solvent in the air expelled by the lung or blood during the workday, as well as it can reflect the amount of solvent absorbed the previous day, collecting pulmonary air or blood 16 hours after the end of exposure, in addition, exposure biomarkers may reflect the amount of substance absorbed during months of exposure [24]. In this case, the substance must have a long half-life, as is the case with exposure to heavy metals. Disease biomarkers are biological parameters, which reflect the relationship of disease/biological system or disease/ pharmacological agent [24]. Most of the time, biochemical changes are strong candidates for biological markers [4]. Effect biomarkers have important applications in monitoring the health status of an organism. Among its applications include (i) use as a diagnostic tool to identify abnormal conditions and /or disease, such as, high concentration of glucose in the blood, aiding in the diagnosis of diabetes [27]; (ii) use as a tool for the identification of cancer, measuring carcinoembryonic antigen 125 (CA-125), marker of several cancers [28]; (iii)

help to classify the severity of a disease, as is the case of the assessment of the concentration of the prostate specific antigen (PSA) in the blood, this assessment reflects the degree of tumor growth and metastasis in the body [28]; (iv) use as an indicator of disease prognosis and monitoring of therapeutic response, performed by a clinical intervention, such as, the measurement of cholesterol concentration in the blood, determining the risk of developing cardiovascular diseases [29].

For this work, disease biomarkers will be the focus of discussion. Here, the importance of this type of biomolecule for the diagnosis of human disease will be discussed, as well as the main phases of discovery and validation of such biomolecules will be briefly reported.

2.2 Discovery of biomarkers

The first step to discovery biomarkers involves the application of strategies aimed detecting candidate biomolecules for biological markers [19]. In this sense, genomics, transcriptomics, proteomics and metabolomics are highlights, mainly, investigating possible changes in the gene, messenger RNA, protein, and metabolite that is characterized as a biological marker [4,5]. Doing a search of the last ten years in one of the main search website (Scopus) with the words: omic (genomic, transcriptomic, proteomic, metabolomic) clinical, there is a significant growth in published works, reporting studies to discovery candidate biomolecules to biomarker of diferente diseases in different types of samples. Table 1 presents some of these numbers. This fact shows the high investment by development agencies, as well as demonstrates the interest of the scientific community in conducting research with the objective of discovering biomarkers that help in the diagnosis and treatment of human diseases.

Table 1 | Number of scientific articles published per year, reporting the search for human diseases biomarker.

Omic	Publication/year					
	2009	2011	2013	2015	2017	2019
Genomic	310	441	649	1038	1381	1822
Transcriptomic	65	126	196	333	501	844
Proteomic	562	698	688	768	840	970
Metabolomic	177	317	515	835	1125	1544

2.2.1 Genomics

Genomics is the area of science that studies the genetic patterns that may exist in the genome of a given organism. Genomics allows to assessment of possible alterations in the DNA code that alone or in combination are associated with susceptibility, expression and evolution of the disease, also considering the therapeutic response [20]. Genomics has important tools for gene analysis, such as: the polymorphism of a single nucleotide (SNP) to analyze possible changes in the bases of genetic sequencing, characterizing candidate genes for biological markers of human diseases [20]. For example, in a study developed by Zou et al. SNP technique was used to identify possible changes in the gene extracted from the blood of patients with type 2 diabetes (n = 152). To do this, the gene extracted from the blood of healthy individuals (n = 120) was used as a control. In the study, the researchers identified high levels of methylation in the PRKCZ gene, indicating that PRKCZ may be involved in the pathogenesis of type 2 diabetes [30].

2.2.2 Transcriptomics

Transcriptomics is another omic approach used to discovery biomarker of human diseses. In this strategy, the set of messenger RNAs (mRNAs) of a given biological system is studied under different conditions, such as diseases [31]. Among the main ways to study gene expression are complementary DNA microarray (cDNA) and the reverse transcription polymerase chain reaction (RT-qPCR)[32]. The cDNA arrangement technique is based on the ability of mRNA to pair (hybridize) with the DNA molecule that gave rise to it, allowing to label in an arrangement form and subsequent identification and quantification. RT-qPCR technique allows quantifying the number of copies (cDNA) of the target sequence, comparatively analyzing the number of copies of cDNA and the amount of mRNA that generated it, a fact that allows a global analysis of the gene expression of the disease [32]. Zhang et al. [33] performed a transcriptomic study in the saliva of individuals with breast cancer to determine possible biomarkers of this disease. In the study, the complementary DNA microarray technique was used to assess the profile of women diagnosed with breast cancer (n=10) and healthy women (n=10). Eight mRNAs (S100A8, CSTA, GRM1, TPT1, GRIK1, H6PD, IGF2BP1, MDM4) were identified as differentially expressed and were suggested as candidates for the biomarker of breast cancer [33].

2.2.3 Proteomics

Proteomics studies the set of proteins in an organism under specific conditions [25,26]. One of the applications is the comparative study [12]. In this study, the protein profiles of at least two samples are compared to identify qualitative and quantitative differences between the molecules, which can result in a biomarker [26]. The combination of separation techniques, as for example, electrophoresis (gel and capillary) and liquid chromatography with mass spectrometry is widely used in proteomic studies, as it allows to separate, quantify and identify the possible differential proteins of the disease, for example [12,26]. In this sense, several proteomic studies have been development to discovery biomarkers. For example, a comparative study in human saliva between patients with oral cancer (n = 16) and healthy patients (n = 16) it was performed in order to verify

differential proteins that are candidates for biomarkers [34]. In the study, the liquid chromatography mass spectrometry tandem (LC-MS/MS) technique was used for fractionation and nano-liquid chromatography quadrupole time of flight (nLC-QTOF MS) for identification, both in reverse phase. Fractionation resulted in 35 fractions. The proteins of each fraction were reduced with dithiothreitol (DTT), oxidized with iodocetamide (IAA) and digested with trypsin, then they were analyzed by nLC-QTOF MS. Comparing the spectra obtained from the analysis by nLC-QTOF MS with the Mascot database (score> 25 and p <0.05), it was possible to identify four main proteins (CD59, involucrin, Ras-related protein Rab-7 and moesin) candidates to biological marker of oral cancer [34].

2.2.4 Metabolomics

aIn addition to genomics, transcriptomics and proteomics, changes in the level of concentration of metabolites in an organism are also important data in the search for disease biomarkers [35]. The main tools used in this type of study involve, in addition to nuclear magnetic resonance (NMR) spectroscopy, the combination of separation techniques such as chromatography (gas or liquid) and capillary electrophoresis with mass spectrometry [35]. Daimon et al. [36]used the capillary electrophoresis time of flight (CE-TOF -MS) technique to identify differentially expressed metabolites with a potential biomarker for type 2 diabetes. For the study, the researchers collected blood serum from non-diabetic men (n = 19) and diabetics (n = 17). Samples of non- diabetic individuals were subdivided into individuals with low glycemia (n = 5) and individuals with tolerable normal glycemia (IGNT) (n = 14). Of the 560 metabolites identified, approximately 74 metabolites were quantified in all serum samples. Significant differences between diabetics and IGNT were observed in 24 metabolites, with glycerol-3-phosphate, being the metabolite with the highest concentration in the diabetic group when compared to the IGNT group [36].

3. Analytical validation of a biomarker

Although, significant investments have been made to discovery possible biomarkers of human diseases, the number of new biomarkers approved by the Food and Drug Administration (FDA) has remained extremely low [37]. A comparative survey of publications and biomarker patents between 2009 and 2019 carried out using Scopus and FDA website shows some of this reality (Figure 1).

Figure 1 shows a low number of biomarker patents when compared to publications. Numerous factors have contributed to this low number of biomarkers that reach the clinical assessment stage [24]. Validation is one of the decisive steps to ensure the clinical testing stage. It is responsible for assessing whether there is sufficient evidence for the clinical utility of the candidate biomolecule as a

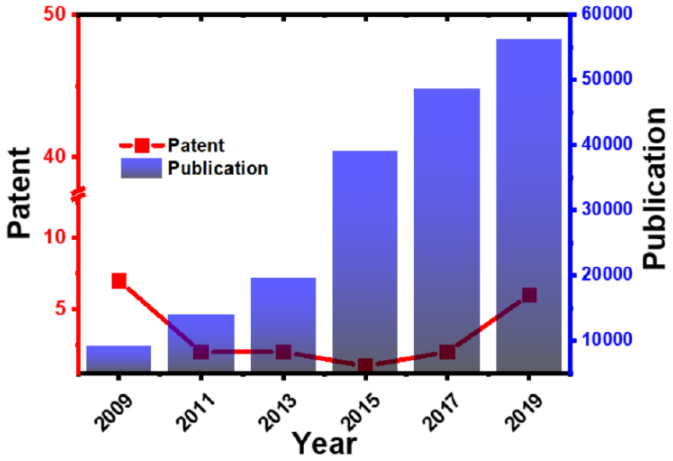

Figure 1 | Overview of the relationship between publications and patents involving biomarker.

biomarker [18,19]. Validation criteria are defined according to the purpose for which it is intended to validate. Generally, for the validation of a biomarker, the following points are used as essential criteria [10, 21]: (i) accuracy; (ii) precision; (iii) sensitivity; and (iv) specificity [18,19,21,22,38].

Accuracy is the proximity between the result of a measurement and the value considered true [38]. Limits for the minimum acceptable accuracy must be established before or during method development [39]. Standard reference materials (SRM) are used in assay validation to estimate intra- and inter-run accuracy [39]. Quality control reference materials are used to accept or reject assay runs [38,39]. Precision is defined as the proximity between measured quantity values obtained by replica measurements under specified conditions [38]. Analytical repeatability and reproducibility are requirements for the implementation of diagnostic tests and treatment. Measurement precision is usually expressed numerically by standard deviation, variance, or coefficient of variation under the specified measurement conditions [38]. Due the lack of well-characterized and well regulated "standard reference materials" or quantitative measures of immune analytes, reference materials often in the forms of biological samples are used to assess relative accuracy of an assay performance [38,39]. Sensitivity refers to the ability of the biomarker to reflect qualitatively and/or quantitatively a change in the biological system through a certain stimulus, such as disease, that is, it measures the probability of obtaining a positive result when the disease is present. Sensitivity is usually assessed experimentally in research groups (patients). Specificity indicates the probability of obtaining a negative result when the individual has no disease, that is, measures the ability to rule out a disease when it is not present. Specificity is assessed, experimentally, in control (healthy) groups [38]. The selectivity/specificity relationship is usually assessed by the receiver operation characterization curve (ROC). The ROC curve allows the quantitative description of the performance of a diagnostic test, resulting in the confirmation or discard of a procedure, in this case: the

validation of a biomolecule as a biomarker [4,38]. In addition to the validation step, it is also necessary to evaluate important aspects of the feasibility of applying the biomarker in clinical practice, such as: validity and availability of the biomarker, agility, simplicity and cost of the experiments [18,19,37].

3.1 Main techniques applied for biomarker validation and their challenges

A precise method to quantify the gene biomarker is the RT -qPCR[32]. With this method it is possible to quantify unique genes or multiple gene sets in a run [40]. PCR is an assay that amplify a DNA (or RNA) target rather than a signal. Using qPCR arrays, up to 384 different transcripts can be analyzed in parallel [40]. There are also other methods available for a holistic screen of gene-expression changes. Until recently, microarray analysis has been the screening method of choice for most gene-expression experiments at the mRNA and miRNA and miRNA level. However, RNA sequencing (RNA-Seq) has been gaining space. RNA-Seq is a new method that allows the sequencing and the quantification of the whole transcriptome of a biological sample. It is a sensitive approach. In addition, a single transcript of a given gene is detectable, and since it is free from guesswork, it is also possible to discover new transcripts or unknown splice variants [41].

Although, RT-qPCR is the current gold standard for sensitive and reproducible miRNA gene-expression analysis, the nature of miRNA molecules poses a challenge for reliable analyses, as for example, (i) members of a miRNA family (e.g., let 7 family) usually differ by only one nucleotide, mainly at the 3'end of the sequence; (ii) the combination of short length of mature miRNAS (~22nt) and a heterogeneous content poses a challenge for cDNA synthesis and primer and probe design since these results in significant difference in the melting temperatures of different miRNAs; (iii) There are no specific guidelines for analyzing and normalizing miRNA expression data [39]. However, strategies to deal with these challenges have been published and are being intensely discussed [38,39,41]. Not only are molecules properties challenging for established technical procedures, but sample matrices also present additional problems [41].

Immunohistochemistry (IHC) is a multi-step process that requires standardized conditions for samples collection, fixation and processing, preparation of the IHC slide, and interpretation of staining results [38]. Tissues are typically the most used sample in the IHC assay. Tissue-based biomarkers can be measured in freshly frozen (FF) samples or formalin fixed paraffin embedded (FFPE) tissue [39]. FFPE tissue blocks often available as archival materials as part of bio-bank samples for conventional IHC, which is the most widely used platform for evaluating biomarker in diagnostic surgical pathology and for retrospective research. However, protein and nucleic acid damage usually occurs though fixation, incorporation, and prolonged storage of FFPE samples [39]. Therefore, to control the pre-analytical requirements of assay performance, it is recommended to run the test on a series of in-house tissues with known IHC performance characteristics, representing known positive and negative tissues (references samples) [38,39].

Another technique used to validate biomarker of human disease is those based on antibodies, such as, enzyme-linked immunosorbent assays (ELISA) and Western blotting [42]. Although different immunoassay formats are available, the sandwich ELISA is the most common assay used in biomarker analysis due to its high specificity and sensitivity [42,43]. In this format, the target protein will be detected using two different antibodies (capture and detection antibodies). However, for many of the candidate biomarkers, a commercially available assay will not exist and specific antibodies against the target of interest and/or the corresponding ELISA need to be developed [43]. The development and optimization of an ELISA assay requires a careful design, as a wide range of variables, ranging from antibody specificity to the concentration and composition of different reagents, can affect the result and therefore the validity of the candidate [42,43]. Jou et al. [44]. Suggested protein transferrin as a biomarker of early-stage oral cancer. In this comparative proteomics study, the researchers used two-dimensional electrophoresis and matrix-assited laser desorption/ionization mass spectrometry (MALDI-TOF MS) to discover differential proteins between saliva of patients diagnosed with oral cancer (n = 11) and healthy patients (n=30). Transferrin was confirmed using ELISA and Western blotting as validation techniques. The specificity and selectivity of the biomarker, according to the authors, were 100% for both criteria [44].

More recently, strategies based on targeted mass spectrometry have been developed to validate biomarker involved in different diseases, offering an alternative [45]. The great advantage of this methodologies on targeted mass spectrometry-based is that they allow the simultaneous accurate and specific quantification of several biomarkers (multiplexing) [45]. Peptides are used as protein surrogates, measured using triple quadrupole instruments in selected reaction monitoring/multiple reaction monitoring (SRM/ MRM) analysis [46]. SRM is an MS-based method performed in two stage of mass analysis [47]. An ion of a specific mass is selected to be detected in the first stage of mass spectrometer an a product ion after fragmentation of the precursor ion is selected in a second detection step [45–47]. MRM is the application of selected reaction monitoring for multiple product ions of one or more precursor ions [45]. The pair of m/z values corresponding to the precursor and fragment ions is called a transition. The intensity of this pair (transition) is recorded and used for absolute quantification [47]. However, the applications of SRM/MRM assays are still in the beginning, therefore, there are, still, limitation, but it is believed that improved automated sample preparation and mass spectrometry technology

(faster instruments with higher selectivity) could make SRM/MRM methods the diagnostic tool of the future [45]. In the work carried out by by Ahn et al. [48], glycoproteins were evaluated to identify possible biomarkers of lung cancer. In the study, using plasma from healthy patients (n = 30) and patients with lung cancer (n = 30), the researchers validated, by SRM-MS, alpha-1-acid glycoprotein and ceruloplasmin as biomarkers of such neoplasia. The specificities and sensitivity of the identified biomarkers were 0.75 and 0.80, respectively [48].

4 General classes of biomarker assays

The American Association of Pharmaceutical Scientists (AAPS) and the US Clinical Ligand Society have listed four general classes of biomarker assays [38]: (i) qualitative assay; (ii) semi-quantitative assay; (iii) relative quantitative assay; (iv) a definitive quantitative assay. Table 2 show a summary of the general classes of biomarker assays, highlighting the performance characteristics required to be evaluated for each type of assay.

A qualitative assay generates categorical data lacks proportionality to the amount of analyte in a sample [49]. Data can be ordinal in the sense that the assay is based on discrete scoring scales such as those often used for IHC or nominal such as the presence or absence of a gene producto [38]. Qualitative assays are only needed to show that they are sufficiently sensitive and specific to detect the target analyte [50]. A semi-quantitative assay does not use a calibration standard but has a continuous response that is expressed in terms of a characteristic of the test sample [49]. Precision can be validated, but not accuracy [38]. A relative quantitative assay uses a concentration-response calibration with reference standards that are not fully representative of the biomarker [49]. As the calibration curve can use a non-certified standard or surrogate matrix or both, studies on parallelism and dilution linearity are necessary [50]. Precision can be validated but accuracy can only be estimated [50]. A definitive quantitative assay makes uses of calibrators and a regression model to calculate absolute quantitative values for unknown samples [49]. The reference standard must be well defined and fully representative of the biomarker [38]. This type of assay can be validated to be accurate and precise [38,49,50].

5. Human diseases and the importance of biomarkers in their diagnosis

According to WHO data, cardiovascular disease, cancer, and viral infection are, at the moment, the diseases that most affect and kill in the world. Thus, in the following topics, a brief description of these diseases is presented, highlighting some of the validated biological markers and their potential biochemical pathways in the biological systems that can be used in the diagnosis and/or treatment of these disease.

5.1 Main techniques applied for biomarker validation and their challenges

Cardiovascular diseases (CVD) are diseases characterized by affecting the circulatory system, that is, the blood vessels and the heart [51]. There are several types of CVD, including myocardial infarction, ischemia, atherosclerosis, and strokes [10,51]. According to the WHO, CVDs are the main cause of mortality in the world, making a total of 7.2 million deaths each year [10]. The incidence of CVDs is related to lifestyle, as well as genetic susceptibility [29]. Thus, studies have provided important information about biological markers that play a role in the diagnosis of CVDs, for example, the LTA and LGALS2 genes and the 5-lipoxygenase and 4D-phosphodiesterase proteins, among others. These biomolecules have been found to be overexpressed in comparative studies with cardiac patients [51,52]. Several molecular mechanism mediate hypertension-induced vascular remodeling [9,10,20,52]. One of the main mechanisms is shown in Figure 2 [51].

The mechanical stretching force exerted by hypertension on the vascular wall promotes the production of reactive oxygen species (ROS) [53], which in turn induce Vascular Smooth Muscle Cell (VSMC) remodeling [54,55]. Hypertension-mediated excessive stretch force also causes changes in the extracellular matrix, activating the RhoA pathway, which in turn promotes actin cytoskeleton remodeling in VSMC [51]; hypertension-induced activation of extracellular signal- regulated kinases 1 and 2 (ERK1/2) and protein kinase B (AKT) also results in vascular remodeling [56,57]. Furthermore, caveolae, which are lipid raft investigations in the plasma membrane, mediate the modeling of hypertension induced VSMC via endothelial nitric oxide synthase (eNOS) and endothelin receptor type A (ETA) [58,59]. Studies have also shown that angiotensin II type 1 receptor (AT1), platelet-derived growth factor receptor (PDGF-R), and specific ion channels, such as voltage-gated calcium channels, are implicated in hypertension-induced VSMC remodeling [60,61]. The forces

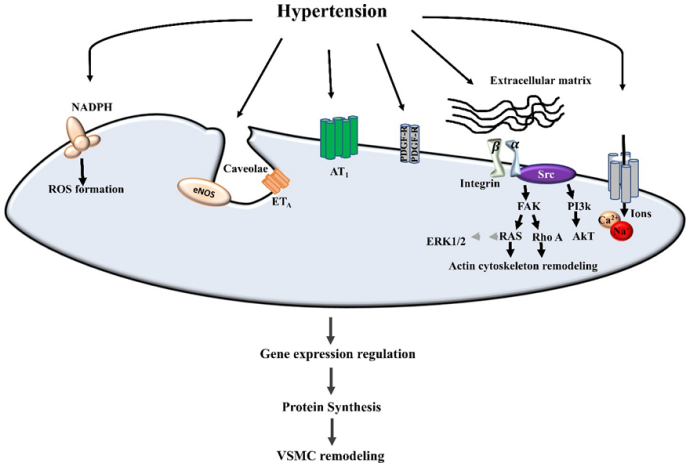

Figure 2 | Scheme of vascular smooth muscle cell (VSMC) remodeling in response to hypertension .

Table 2 | Overview of the general classes of biomarker assays, highlighting the performance characteristics that should be evaluated for each type of assay

Assay type	Definition	Examples	Assay performance characteristics			
			Accuracy	Precision	Sensitivity	Specificity
Qualitative	No calibration standards, and the bioanalytical results are expressed in terms of test discrete sample characteristic (discontinuous) and reported in any ordinal or nominal formats.	Immunohistochemistry (IHC), imaging	-	+	+	+
Semi-quantitative	Does not use a reference standard (or calibration curve). However, the analytical response is continuous (numeric), and the bioanalytical result are expressed in terms of a characteristic of the test sample.	Enzymatic assays (where activity might be expressed per unit volume); Anti-drug antibody assays (where the readout is a titer or % bound).	-	+	+	+
Relative quantitative	Reference standard that is not well characterized, not available in a purified form, or not fully representative of the endogenous biomarker; The results of these assays are expressed in continuous numeric units of the relative reference standard; Most biomarker assays provide "relative" quantification of the analyte. The interest is in the temporal changes of the marker rather than the absolute concentration.	Quantitative -PCR, flow cytometry	+	+	+	+
Definitive quantitative	Reference standard is well defined or characterized and is representative of the endogenous biomarker; Assay results are absolute quantitation of the analyte and are expressed in continuous numerical units of the definitive reference standard.	Human insulin or a steroid assay, small molecule, LC-MS	+	+	+	+

exerted by hypertension cause endothelial damage and dysfunction, resulting in reduced production on nitric oxide (NO) [51]. Consequently, blood pressure-induced vasodilation is compromised. In addition, hypertension-mediated endothelial dysfunction promotes the development of atherosclerosis, which is associated with the accumulation of na atheromatous plaque. Atherosclerosis is primarily composed of oxidized low-density lipoprotein (LDL) and macrophages inside the artery walls [51,52]. It is a risk factor for coronary artery disease, myocardial infarction (MI), hypertension, stroke, and peripheral arterial disease. When blood pressure is low, endothelial cells secrete several vasoactive molecules, such as, endothelin- 1, angiotensin II, prostanoids, and ROS, which act on VSMCs to promote VSMC concentration and subsequent vasoconstriction [51,54]. In contrast, when blood pressure increases, vasodilating substances, such as NO, prostacyclin, and endothelium-derived hyperpolarization factor are produced by endothelial cells [62]. Arterial calcification is associated with atheroma progression and alters the mechanical properties of the vascular wall, thus increasing the risk of atherosclerotic plaque rupture [51,60,61].

5.2 Biomarkers and cancer

Cancer is caused by an excessive multiplication of cells in certain regions of the body [63].

It is the second leading cause of death in the worldwide, second only to cardiovascular disease [64]. The most common types of cancer are skin cancer [65], breast cancer [66], lung cancre [67], prostate cancer [68] among others. Strong evidence supports the concept that cancer is a genetic disease that involves the abnormal growth of transformed cells, and this abnormal growth is triggered in combination with extrinsic factors (lifestyle, environmental aspect) [63,65–68]. Characterizing early-stage tumor cells, researchers have reported important genetic markers for different cancers, such as breast cancer (RAD21, PCTAIRE, CDC25B, CENPF, VEGF, PGK1 and others) [69], prostate cancer (PSA, PSM) [68], pancreatic cancer (KRAS, TP53, DPC4) [70] and others [71] that allow for early diagnosis and monitoring (progression, regression) of such disease. Figure 3 shows intracellular signaling via the P13K-AKT-mTOR pathways that is dysregulated pathway in human cancers [7].

The P13K-AKT-mTOR signaling pathway is one of the main unregulated pathways in human cancers [7]. The P13K -AKT–mTOR pathway is triggered by the activation of various growth factor receptor tyrosine kinases or G protein-coupled receptors [72–74]. The class I PI3K proteins are recruited to the plasma membrane by adapter proteins, such as insulin receptor substrate (IRS) family members, that interact with these activated cell-surface receptors, leading to phosphorylation ofphosphatidylinositol 4,5-bisphosphate (PIP2) to generate phosphatidylinositol 3,4,5-trisphosphate

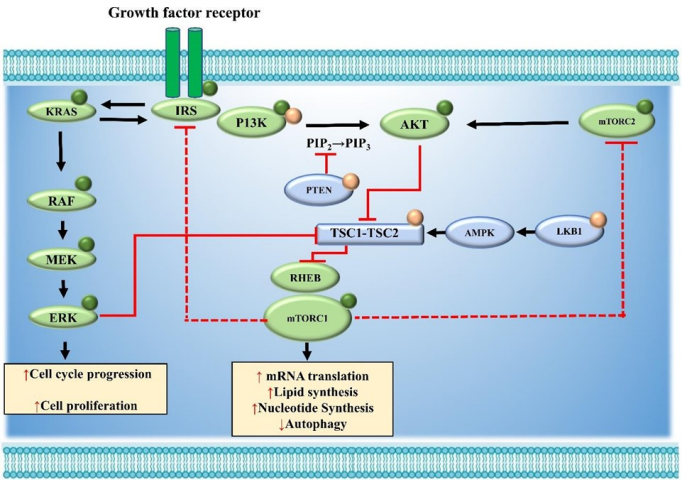

Figure 3 | Schematic of dysregulated pathway in human cancers involving intracellular signaling via the P13K-AKT-mTOR pathway.

(PIP3) [75–77]. PIP3 is a second messenger that activates the AKT kinases, which can phosphorylate tuberous sclerosis protein 1 (TSC1) and TSC2, dissociating, thus, the TSC1-TSC2 complex [77]. The TSC1–TSC2 complex down-regulates the mTOR kinase activity; and, therefore, AKT results in the activation of the mTOR complex 1 (mTORC1), increasing protein and lipid synthesis and decreased autophagy, which supports cell growth and proliferation [7,73,74,76,77]. Notably mTORC1 is involved in a negative feedback loop that serves to avoid overactivation of AKT (dashed red lines in the Figure 3) [78]. The PI3K- AKT-mTOR pathway can be upregulated by activating molecular changes in the PI3K subunits (such as PI3K catalytic subunit α isoform, encoded by PIK3CA), AKT, and mTOR (represented by green circles in the Figure 3) or by loss of function changes in regulatory subunits of PI3K (such as PI3K regulatory subunit-α, encoded by PIK3R1), PTEN, TSC1, TSC2, and LKB1 (also known as STK11) (depicted by orange circles in the Figure 3)[74,76]. In parallel, activation of the growth factor receptor tyrosine kinases and G protein-coupled receptors induces KRAS-RAF- MEK-ERK signaling, and ERK activation may further contribute to mTORC1 activation trough dissociation of the TSC1-TSC2 complex [78]. KRAS can also increase PI3K activation [7,74,76,77]. Notably, the KRAS-RAF-MEK-ERK pathway can also be constitutively activated by gain of function changes in components kinases or cell-surface receptors (green circles in the Figure 3) [7,73].

5.3 Biomarkers and infection diseases

Infectious diseases are categorized as diseases caused by pathogenic microorganism such as viruses [79]. These diseases have been major threat worldwide and have a great impact on public health and the world economy [80,81]. Among the different types of infectious diseases, human immunodeficiency virus (HIV) [82–84], tuberculosis [85,86], and malaria [87] are known as the leading causes of deaths globally [79]. In addition, several types of neglected tropical diseases, such as dengue [88,89], yellow fever [90,91], zika virus [92,93], and chikungunya [94] are also considered to be major global threats. Although such diseases emerge in tropical and subtropical regions, the risk of these infectious diseases may be worldwide due to the global economy and migration [95]. Recently, the WHO declared a pandemic state due to viral infection caused by a new virus of the Coronaviridae family: SRAS-CoV-2. Coronavirus disease (COVID-19) is currently one of the main causes of death in the world [79], causing about four million deaths in June 2021. In this sense, the search for the ideal biomarkers in infectious diseases (with high sensitivity, specificity, and predictive capacity) should be focused on the detection and identification of the infectious agent, monitoring the clinical response, and guiding the duration of treatment [95], as in the case of procalcitonin (PCT) assay, which can discriminate between a viral and a bacterial infection and has been approved by FAD [95]. Currently, the SARS-CoV-2 RNA genome serves as a major biomarker for direct viral detection and the primary COVID-19 diagnosis [96-98]. Viral proteins encoded by SARS-CoV-2 could theoretically serve as alternative biomarkers for viral detection, but due to the complexity of protein detection and the significantly greater number of biological samples required, they are often impractical targets [80,99]. Many of these viral proteins, however, can serve as potential targets for antiviral drugs or biomarkers of drug development for COVID-19 treatment [95,98]. Immune defense cells or T-lymphocytes (particularly CD4þ and CD8þ cells) are among the first human cells to respond to the threat of SARS-CoV-2 [98,100,101]. In a recent study, all twenty of the COVID-19 patients produced CD4þ T-cells and antibodies (IgG, IgM and IgA) targeting the viral S-protein, and 70% of cases produced measurable CD8þ T-cells [98] (Figure 4). These results confirm that the human imune system can mount a substantial and long-lasting response to the new coronavirus [79,99]. On the other hand, lymphopenia has also observed in COVID-19 patients, in which, levels of CD4þ and CD8þ T -cells were decreased in severe patients in comparison with mild cases but restored when the viral infection was cleared [79,95,98]. Cytokine release syndrome (CRS) is a major

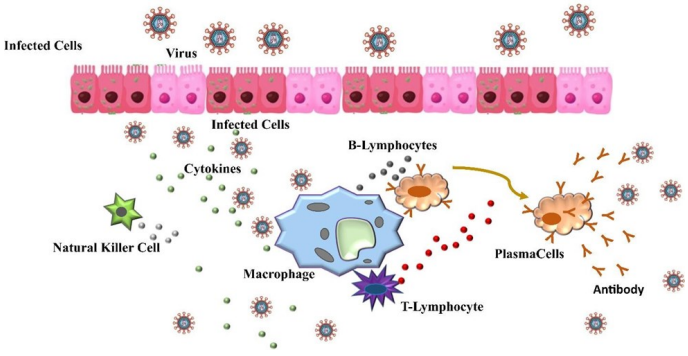

Figure 4 | Immune system scheme during a viral infection involving cytokines, macrophages, natural killer cells, and B and T cells and antibodies that constitute the lines of defense of the immune system.

cause of multiple organ injury and fatal outcome induced by SARS-CoV-2 infection in severe COVID-19 patients. Xiao et al. [102] investigated cytokines/chemokine profiles in the serum of healthy controls, patients with mild and severe COVID-19. Correlation analyses show close associations between metabolites and pro-inflammatory cytokines/ chemokines, such as IL-6, M-CSF, IL-1α, IL-1β, and imply a potential regulatory crosstalk between arginine, tryptophan, purine metabolismo and hyperinflammation [102]. In addition, it has been shown that targeting metabolism markedly modulates the release of pro-inflammatory cytokines by peripheral blood mononuclear cells isolated from SARS-CoV-2 rhesus monkeys infected ex vivo, suggesting that exploiting metabolic changes may be a potential strategy for the treatment of fatal CRS in COVID-19 [102]. In another study, Huang et al. [103] investigated cytokines in the serum of patients infected with the dengue virus during the Guangdong outbreak in 2014, in which more than 50,000 dengue cases were reported and 6 patients died. They found that the levels of CCL17 and CXCL5 were significantly lower than controls, while several pro-inflammatory cytokines such as CXCL9, IP-10, CXCL11, IL-8, and IL-10 were highly upregulated in patients after dengue infection. These results determine the association of clinical routine indices and inflammatory cytokines and would be useful to understand the interplay between the virus and the host responses during the acute stage of dengue infection [103].

6. Linking the biomarker to the clinical endpoint

The relationship of a biomarker to clinical outcome requires a firm promise of reliability for surrogate endpoints through a structured approach and a substantial body of evidence [5,104]. Three clinical features are required to obtain a surrogate marker of success: (i) efficiency, (ii) linkage and (iii) congruence [105]. To be considered efficient, the surrogate marker must exhibit superior accessibility in terms of technical and temporal acquisition allowing acquisition of accurate information in shorter time intervals and clinical trials with fewer resources and less subject participation. By linkage, it means that a plausible underlying relationship between surrogate marker and clinical endpoint must be demonstrated and substantiated by comprehensive scientific evidence. For congruence, the surrogate must produce parallel estimates of risk and benefit as endpoints. In addition, there must be a clear difference in surrogate marker measurements between individuals with and without the disease. In intervention studies, the expected clinical benefits must be deductible from the observed changes in the surrogate marker [5,104,105].

7. Outlook

As a perspective on biomarker validation of human disease, we highlight development of imaging technologies with accuracy, precision, specificity and sensitivity to validate a biomarker. For example, the use of both positron emission tomography and near-infrared imaging has been reported to quantitatively monitor the initial trafficking of vaccine to drain lymph nodes after intramuscular injection in non-human primates [106]. To do this, the researcher previously labeled a messenger RNA vaccine model a probe for both techniques. They used near-infrared fluorescence and flow cytometry to validate tissue extracted from sacrificed animals [106]. In another study, the intratumoral metabolic heterogeneity of breast cancer patients were determined via photoacoustic microscopy, measuring the oxygen consumption rates of single cells taken from the tissue after it was homogenized in a single-cell suspension and the cells deposited in microwell arrays [107]. Heterogeneity has also been observed in how tissues and even single cells metabolize glucose in cancer progression by using Raman spectroscopy and stimulated Raman scattering to trace deuterated glucose in living mice [108]. The research validated their quantitative measurements of Raman intensity using nuclear magnetic resonance spectra of tissue lipid extracts [108]. This optical multiplexing imaging of glucose metabolites may in the future, discover new metabolic biomarkers of disease or be used for the metabolic phenotyping of biopsied tissues from patients.

In addition, as the volume of biomarker data continues to grow, the scientific discovery of biomarkers based on artificial intelligence can be a complementary approach to classical strategies. In some cases, machine learning algorithms can be leveraged to classify biomolecules that accurately contribute to the prediction of a disease state, for example, or treatment, thus generating new and testable assays. However, additional information needs to be considered for better diagnostic alignment, such as standardization of protocols and equipment, large-scale cross- validation, multicenter trials controlling for age, gender and culture variables, definition of end points and use cases, cost-effective and easy to use.

8. Concluding remarks

Different types of human diseases are related to lifestyle and genetic susceptibility, which determine exposure to various risk factors and constitute a determining element in the emergence and development of human diseases. The risk of suffering health deterioration can be assessed using biomarkers that express the likelihood that an unwanted effect will occurs due to exposure to environmental aspects or gene dysfunction. Although their findings present great challenges, biomarkers are an important tool in the prediction, diagnosis, and treatment of human diseases, as they can reflect possible changes in the biological system, allowing the unequivocal specification of such disease. However, it is important to recognize the need for a detailed and a rigorous path to discovery and validation of human disease biomarkers in research and clinical settings,

considering important parameters, such as accuracy, precision, specificity, and sensitivity. Obviously, several challenges are observed during these processes.

Conflicts of interest

The authors declare no conflict of interest

Author contributions

Both authors (J.R.J and M.A.Z.A) contributed equally to the writing of this work.

Acknowledgements

The authors thank to Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, São Paulo, Brazil, grant numbers 2016/07384-7, 2018/00786-0, and 2020/02020-2), and Coordenação de Aperfeiçoamento de Pessoal de Nivel Superior – Brazil (CAPES – 88887.339545/2019-00) for financial support.

Abbreviation

AAPS - american Association of Pharmaceutical Scientists
AKT – protein kinase B
AT1 – angiotensin II type 1 receptor
CA-125 - carcinoembryonic antigen 125
cDNA – complementary DNA
CE-TOF-MS – capillary electrophoresis time of flight
COVID-19 - Coronavirus disease
CSF – cerebrospinal fluid
CVD – cardiovascular diseases
DTT dithiothreitol
ELISA – enzyme-linked immunosorbent assays
eNOS – endothelial nitric oxide synthase
ERK1/ - extracellular signal-regulated kinases 1and 2
ETA – endothelin receptor type A
FDA – Food and Drug Administration
FF – Freshly frozen
FFPE – formalin fixed paraffin embedded
HIV - human immunodeficiency virus
IAA – iodocetamide
IGNT - tolerable normal glycemia
IHC - Immunohistochemistry
IRS – insulin receptor substrate
LDL – low-density lipoprotein
MALDI – TOF MS – matrix-assisted laser desorption/ ionization mass spectrometry
MI - myocardial infarction
mRNAs - messenger RNAs
mTORC1 - mTOR complex 1
NIH - National Institute of Health
nLC-QTOF MS - nano-liquid chromatography quadrupole time of flight
NMR - nuclear magnetic resonance
NO – nitric oxide
PCT – procalcitonin
PDGF-R – platelet-derived growth factor receptor
PIP2 – phosphorylation of phosphatidylinositol 4,5-bisphosphate
PIP3 – phosphatidylinositol 3,4,5- trisphosphate
RNA-Seq – RNA sequencing
ROC – receiver operation characterization curve
ROS – reactive oxygen species ROS - reactive oxygen species
RT-qPCR - real-time reverse transcription polymerase chain reaction
SNP – single nucleotide
SRM/MRM – selected reaction monitoring/multiple reaction monitoring
TSC1 - tuberous sclerosis protein 1
VSMC – Vascular Smooth Muscle Cell
WHO – World Health Organization

References

[1] V.C. Sgarbieri, M.T.B. Pacheco, Brazilian Journal of Food Technology. 20 (2017) 1-23. https://doi.org/10.1590/1981-6723.00717.

[2] F.A.S. Addor, Clinical, Cosmetic and Investigational Dermatology. 11 (2018) 437–443. https://doi.org/10.2147/CCID.S177448.

[3] Y. Yoshida, N. Matsunaga, T. Nakao, K. Hamamura, H. Kondo, T. Ide, H. Tsutsui, A. Tsuruta, M. Kurogi, M. Nakaya, H. Kurose, S. Koyanagi, S. Ohdo, Nature Communications. 12 (2021) 1-17. https://doi.org/10.1038/s41467-021-23050-x.

[4] D.S. Wishart, B. Bartok, E. Oler, K.Y.H. Liang, Z. Budinski, M. Berjanskii, A. Guo, X. Cao, M. Wilson, Nucleic Acids Research. 49 (2021) 259–267. https://doi.org/10.1093/nar/gkaa1067.

[5] K.D. Davis, N. Aghaeepour, A.H. Ahn, M.S. Angst, D. Borsook, A. Brenton, M.E. Burczynski, C. Crean, R. Edwards, B. Gaudilliere, G.W. Hergenroeder, M.J. Iadarola, S. Iyengar, Y. Jiang, J.T. Kong, S. Mackey, C.Y. Saab, C.N. Sang, J. Scholz, M. Segerdahl, I. Tracey, C. Veasley, J. Wang, T.D. Wager, A.D. Wasan, M.A. Pelleymounter, Nature Reviews Neurology. 16 (2020) 381–400. https://doi.org/10.1038/s41582- 020-0362-2.

[6] G. Jung, E. Hernández-Illán, L. Moreira, F. Balaguer, A. Goel, Nature Reviews Gastroenterology and Hepatology. 17 (2020) 111–130. https://doi.org/10.1038/s41575-019-0230-y.

[7] F. Janku, T.A. Yap, F. Meric-Bernstam, Nature Reviews Clinical Oncology. 15 (2018) 273–291. https://doi.org/10.1038/nrclinonc.2018.28.

[8] T. Powles, D. Carroll, S. Chowdhury, G. Gravis, F. Joly, J. Carles, A. Fléchon, P. Maroto, D. Petrylak, F. Rolland, N. Cook, A. v. Balar, S.S. Sridhar, M.D. Galsky, P. Grivas, A. Ravaud, R. Jones, J. Cosaert, D. Hodgson, I. Kozarewa, R. Mather, R. McEwen, F. Mercier, D. Landers, Nature Medicine. (2021). https://doi.org/10.1038/s41591- 021-01317-6.

[9] T. Hirata, Y. Arai, S. Yuasa, Y. Abe, M. Takayama, T. Sasaki, A. Kunitomi, H. Inagaki, M. Endo, J. Morinaga, K. Yoshimura, T. Adachi, Y. Oike, T. Takebayashi, H. Okano, N. Hirose, Nature Communications. 11 (2020) 1-17. https://doi.org/10.1038/s41467-020-17636-0.

[10] F. Genre, J. Rueda-Gotor, S. Remuzgo-Martínez, V. Pulito-Cueto, A. Corrales, V. Mijares, L. Lera-Gómez, V. Portilla, R. Expósito, C. Mata, R. Blanco, J. Llorca, V. Hernández-Hernández, E. Vicente, C. Fernández-Carballido, M.P. Martínez-Vidal, D. Castro-Corredor, J. Anino- Fernández, C. Rodríguez-Lozano, O. Gualillo, J.C. Quevedo-Abeledo, S. Castañeda, I. Ferraz- Amaro, R. López-Mejías, M. González-Gay, Scientific Reports. 10 (2020) 9636. https://doi.org/10.1038/s41598-020-66816-x.

[11] J.R. Jesus, M.A.Z. Arruda, Rapid Communications in Mass Spectrometry. (2020) 1–10. https://doi.org/10.1002/rcm.8798.

[12] J.R. de Jesus, R.M. Galazzi, T.B. de Lima, C.E.M. Banzato, L.F. de Almeida Lima e Silva, C. de Rosalmeida Dantas, F.C. Gozzo, M.A.Z. Arruda, Clinical Biochemistry. (2017) 1118-1125. https://doi.org/10.1016/j.clinbiochem.2017.06.009.

[13] G.D.S. Pessôa, J.R. de Jesus, T.S. Balbuena, M. Aurélio, Z. Arruda, Rapid Communications in Mass Spectrometry 34 (2020) e8698. https://doi.org/10.1002/rcm.8698.

[14] J.R. de Jesus, (2019) 39– 58. https://doi.org/10.1007/978-3-030-29473-1_3.

[15] S. Paramasivan, S.S. Adav, S.F.C. Ngan, R. Dalan, M.K.S. Leow, H.H. Ho, S.K. Sze, Scientific Reports. 10 (2020) 1-12. https://doi.org/10.1038/s41598-020-62341-z.

[16] S. Park, S. Chon, S.Y. Park, S. Yun, S.H. Baik, J.T. Woo, S.Y. Rhee, Y.K. Pak, S.H. Kim, Scientific Reports. 11 (2021) 1-9. https://doi.org/10.1038/s41598-021-82794-0.

[17] J.R. Wang, W.N. Gao, R. Grimm, S. Jiang, Y. Liang, H. Ye, Z.G. Li, L.F. Yau, H. Huang, J. Liu, M. Jiang, Q. Meng, T.T. Tong, H.H. Huang, S. Lee, X. Zeng, L. Liu, Z.H. Jiang, Nature Communications. 8 (2017) 1-14. https://doi.org/10.1038/s41467-017-00662-w.

[18] N. Goossens, S. Nakagawa, X. Sun, Y. Hoshida, Translational Cancer Research. 4 (2015) 256–269. https://doi.org/10.3978/j.issn.2218- 676X.2015.06.04.

[19] V.B. Kraus, Nature Reviews Rheumatology. 14 (2018) 354–362. https://doi.org/10.1038/s41584-018-0005-9.

[20] M. v. Holmes, T.G. Richardson, B.A. Ference, N.M. Davies, G. Davey Smith, Nature Reviews Cardiology. 18 (2021) 1-19. https://doi.org/10.1038/s41569-020-00493-1.

[21] E.E. Redei, J.D. Ciolino, S.L. Wert, A. Yang, S. Kim, C. Clark, K.B. Zumpf, K.L. Wisner, Translational Psychiatry. 11 (2021) 1-9. https://doi.org/10.1038/s41398- 020-01188-4.

[22] T.L. Whiteside, Validation of plasma-derived small extracellular vesicles as cancer biomarkers, Nature Reviews Clinical Oncology. 17 (2020)719–720. https://doi.org/10.1038/s41571-020-00433-5.

[23] X. Zheng, T. Chen, A. Zhao, Z. Ning, J. Kuang, S. Wang, Y. You, Y. Bao, X. Ma, H. Yu, J. Zhou, M. Jiang, M. Li, J. Wang, X. Ma, S. Zhou, Y. Li, K. Ge, C. Rajani, G. Xie, C. Hu, Y. Guo, A. Lu, W. Jia, W. Jia, 9 Nature Communications. 12 (2021) 1487. https://doi.org/10.1038/s41467-021-21744-w. 10

[24] R.M. Califf, Experimental Biology and Medicine. 243 (2018) 213–221. https://doi.org/10.1177/1535370217750088.

[25] J.R. de Jesus, G. de S. Pessôa, A. Sussulini, J.L.C. Martínez, M.A.Z. Arruda, Journal of Proteomics. 145 (2016) 187–196. https://doi.org/10.1016/j.jprot.2016.04.034.

[26] J.R. de Jesus, B.K. de Campos, R.M. Galazzi, J. Luis, C. Martinez, Analytical and Bioanalytical Chemistry (2014) 661 –667. https://doi.org/10.1007/s00216-014-8341-3.

[27] T. Huo, S. Cai, X. Lu, Y. Sha, M. Yu, F. Li, Journal of Pharmaceutical and Biomedical Analysis. 49 (2009) 976–82. https://doi.org/10.1016/j.jpba.2009.01.008.

[28] A. aroon Lertkhachonsuk, S. Buranawongtrakoon, N. Lekskul, N. Rermluk, W.W. Wee- Stekly, C. Charakorn, Journal of Obstetrics and Gynaecology Research. 46(2020) 2287–2291. https://doi.org/10.1111/jog.14427.

[29] Y. ho Lee, S.G. Lee, M.H. Lee, J.H. Kim, B.W. Lee, E.S. Kang, H.C. Lee, B.S. Cha, Journal of the American Heart Association. 3 (2014) 1-13. https://doi.org/10.1161/JAHA.113.000650.

[30] L. Zou, S. Yan, X. Guan, Y. Pan, X. Qu, Journal of Diabetes Research. (2013) 1-5. https://doi.org/10.1155/2013/721493.

[31] L. Larsson, J. Frisén, J. Lundeberg, Nature Methods. 18 (2021) 15–18. https://doi.org/10.1038/s41592-020-01038-7.

[32] Y. Mo, R. Wan, Q. Zhang, Methods Mol. Bio. (2012) 99–112. https://doi.org/10.1007/978-1-62703-002-1_7.

[33] L. Zhang, H. Xiao, S. Karlan, H. Zhou, J. Gross, D. Elashoff, D. Akin, X. Yan, D. Chia, B. Karlan, D.T. Wongiomarkers for the non- invasive detection of breast cancer, PLoS ONE. 5 (2010) 1-7. https://doi.org/10.1371/journal.pone.0015573.

[34] S. Hu, M. Arellano, P. Boontheung, J. Wang, H. Zhou, J. Jiang, D. Elashoff, R. Wei, J.A. Loo, D.T. Wong, Clinical Cancer Research. 14 (2008) 6246–6252. https://doi.org/10.1158/1078-0432.CCR-07-5037.

[35] L. Perez de Souza, S. Alseekh, F. Scossa, A.R. Fernie, Nature Methods. (2021) 1-14. https://doi.org/10.1038/s41592-021-01116-4.

[36] M. Daimon, T. Soga, A. Hozawa, T. Oizumi, W. Kaino, K. Takase, S. Karasawa, Y. Jimbu, K. Wada, W. Kameda, S. Susa, T. Kayama, K. Saito, M. Tomita, T. Kato, Internal Medicine. 51 (2012) 545–551. https://doi.org/10.2169/internalmedicine.51.6612.

[37] Food and Agriculture Department. www.fda.gov/drugs/biomarker-qualification-program/list-qualified-biomarkers (accessed at 28th of may 2021)

[38] L. Fang, C. Su, Springer International Publishing, 2019. https://doi.org/10.1007/978-3-030-31503-0.

[39] K.K. Dobbin, A. Cesano, J. Alvarez, R. Hawtin, S. Janetzki, I. Kirsch, G. v. Masucci, P.B. Robbins, S.R. Selvan, H.Z. Streicher, J. Zhang, L.H. Butterfield, M. Thurin, Journal for Immuno Therapy of Cancer. 4 (2016) 1-14. https://doi.org/10.1186/s40425-016-0179-0.

[40] L. Antoury, N. Hu, L. Balaj, S. Das, S. Georghiou, B. Darras, T. Clark, X.O. Breakefield, T.M. Wheeler, Nature Communications. 9 (2018) 1-16. https://doi.org/10.1038/s41467-018-06206-0.

[41] Henry Erlich (Ed), PCR Technology, Palgrave Macmillan, UK, 1989, pp 1-246.

[42] M. del Campo, W. Jongbloed, H.A.M. Twaalfhoven, R. Veerhuis, M.A. Blankenstein, C.E. Teunissen, Frontiers in Neurology. 6 (2015) 1-10. https://doi.org/10.3389/fneur.2015.00202.

[43] E. Sheikhzadeh, S. Eissa, A. Ismail, M. Zourob, Talanta. 220 (2020) 1-17. https://doi.org/10.1016/j.talanta.2020.121392.

[44] Y.J. Jou, C. der Lin, C.H. Lai, C.H. Chen, J.Y. Kao, S.Y. Chen, M.H. Tsai, S.H. Huang, C.W. Lin, Analytica Chimica Acta. 681 (2010) 41–48. https://doi.org/10.1016/j.aca.2010.09.030.

[45] G. Mermelekas, A. Vlahou, J. Zoidakis, Expert Review of Molecular Diagnostics. 15 (2015) 1441–1454. https://doi.org/10.1586/14737159.2015.1093937.

[46] P. Picotti, R. Aebersold, Nature Methods. 9 (2012) 555–566. https://doi.org/10.1038/nmeth.2015.

[47] Y. Jin Kim, Z. Zaidi-Ainouch, S. Gallien, B. Domon, Nature Protocols. 7 (2012) 859–871. https://doi.org/10.1038/nprot.2012.023.

[48] Y.H. Ahn, E.S. Ji, N.R. Oh, Y.S. Kim, J.H. Ko, J.S. Yoocancer, Journal of Proteomics. 106 (2014) 221–229. https://doi.org/10.1016/j.jprot.2014.04.031.

[49] J.W. Lee, V. Devanarayan, Y.C. Barrett, R. Weiner, J. Allinson, S. Fountain, S. Keller, I. Weinryb, M. Green, L.

Duan, J.A. Rogers, R. Millham, P.J. O'Brien, J. Sailstad, M. Khan, C. Ray, J.A. Wagner, Pharmaceutical Research, 23 (2006) 312–328. https://doi.org/10.1007/s11095-005-9045-3.

[50] J. Cummings, F. Raynaud, L. Jones, R. Sugar, C. Dive, British Journal of Cancer. 103 (2010) 1313–1317. https://doi.org/10.1038/sj.bjc.6605910.

[51] C.M. Ghantous, L. Kamareddine, R. Farhat, F.A. Zouein, S. Mondello, F. Kobeissy, A. Zeidan, Biomedicines. 8 (2020) 1–19. https://doi.org/10.3390/biomedicines8120552.

[52] G. Wang, S.D. Grosse, M.W. Schooley, American Journal of Preventive Medicine. 53 (2017) S115–S117. https://doi.org/10.1016/j.amepre.2017.08.005.

[53] C.M. Ghantous, F.H. Kobeissy, N. Soudani, F.A. Rahman, M. Al-Hariri, H.A. Itani, R. Sabra, A. Zeidan, Frontiers in Pharmacology. 6 (2015) 1-14. https://doi.org/10.3389/fphar.2015.00240.

[54] Y. Zhou, M.J. Zhang, B.H. Li, L. Chen, Y. Pi, Y.W. Yin, C.Y. Long, X. Wang, M.J. Sun, X. Chen, C.Y. Gao, J.C. Li, L.L. Zhang, PLoS ONE. 11 (2016) 1-14. https://doi.org/10.1371/journal.pone.0154720.

[55] W. Nour-Eldine, C.M. Ghantous, K. Zibara, L. Dib, H. Issaa, H.A. Itani, N. El-Zein, A. Zeidan, Frontiers in Pharmacology. 7 (2016) 1-15. https://doi.org/10.3389/fphar.2016.00086.

[56] P. Liu, Y. Gu, J. Luo, P. Ye, Y. Zheng, W. Yu, S. Chen, Experimental Cell Research. 380 (2019) 36–46. https://doi.org/10.1016/j.yexcr.2019.02.022.

[57] M. Yoshizumi, Y. Kyotani, J. Zhao, K. Nakahira, Annals of Translational Medicine. 8 (2020) 157–157. https://doi.org/10.21037/atm.2019.12.145.

[58] X. Lian, C. Matthaeus, M. Kaßmann, O. Daumke, M. Gollasch, Pathophysiological Role of Caveolae in Hypertension, Frontiers in Medicine. 6 (2019) 1-9. https://doi.org/10.3389/fmed.2019.00153.

[59] S.J. Forrester, K.J. Elliott, T. Kawai, T. Obama, M.J. Boyer, K.J. Preston, Z. Yan, S. Eguchi, V. Rizzo, Hypertension. 69 (2017) 79–86. https://doi.org/10.1161/HYPERTENSIONAHA.116.08278.

[60] I.A.M. Brown, L. Diederich, M.E. Good, L.J. Delalio, S.A. Murphy, M.M. Cortese-Krott, J.L. Hall, T.H. Le, B.E. Isakson, Arterioscler Thromb Vasc Biol. 38 (2018) 1969–1985. https://doi.org/10.1161/ATVBAHA

[61] A. Jaminon, K. Reesink, A. Kroon, L. Schurgers, International Journal of Molecular Sciences. 20 (2019) 1-29. https://doi.org/10.3390/ijms20225694.

[62] G. Jia, W. Durante, J.R. Sowers, Diabetes. 65 (2016) 2118–2120. https://doi.org/10.2337/dbi16-0026.

[63] K. Pantel, R.H. Brakenhoff, Nature Reviews Cancer. 4 (2004) 448–456. https://doi.org/10.1038/nrc1370.

[64] World Health Organization – www.who.int/news-room/fact-sheets/detail/the-top-10- causes-of-death (accessed in 28th of may 2021).

[65] J. Ock, T. Kim, S. Lee, T.S. Yang, M. kim, W. Jeong, J. Choi, N. Kim, Scientific Reports. 11 (2021). https://doi.org/10.1038/s41598-021-88287-4.

[66] S.J. Rodgers, L.M. Ooms, V.M.J. Oorschot, R.B. Schittenhelm, E. v. Nguyen, S.A. Hamila, N. Rynkiewicz, R. Gurung, M.J. Eramo, A. Sriratana, C.G. Fedele, F. Caramia, S. Loi, G. Kerr, H.E. Abud, G. Ramm, A. Papa, A.M. Ellisdon, R.J. Daly, C.A. McLean, C.A. Mitchell, Nature Communications. 12 (2021) 3140. https://doi.org/10.1038/s41467-021-23241-6.

[67] J. Ju, A. Chen, Y. Deng, M. Liu, Y. Wang, Y. Wang, M. Nie, C. Wang, H. Ding, B. Yao, T. Gui, X. Li, Z. Xu, C. Ma, Y. Song, M. Kvansakul, K. Zen, C.Y. Zhang, C. Luo, M. Fang, D.C.S. Huang, C.D. Allis, R. Tan, C.K. Zeng, J. Wei, Q. Zhao, Nature Communications. 8 (2017) 1-14. https://doi.org/10.1038/s41467-017-00988-5.

[68] R. J. Rabello; C. Oing; K. E. Knudsen; S. Loeb; D.C. Johnson; R.E.Reither; S. Gillssen; T. Van der Kwast; R. Bristow, Nature Reviews Disease Primers. 7 (2021) 1-27. https://doi.org/10.1038/s41572-021-00249-2.

[69] L. J. van't Veer; H. Day; M. J. Van de Vijver; Y. D. He; A.A.M. Hart; M. Mao; H.L. Peterse et al, Nature. 415 (2002) 530-536. https://doi:10.1038/415530a.

[70] S. Dhara, S. Chhangawala, H. Chintalapudi, G. Askan, V. Aveson, A.L. Massa, L. Zhang, D. Torres, A.P. Makohon-Moore, N. Lecomte, J.P. Melchor, J. Bermeo, A. Cardenas, S. Sinha, D. Glassman, R. Nicolle, R. Moffitt, K.H. Yu, S. Leppanen, S. Laderman, B. Curry, J. Gui, V.P. Balachandran, C. Iacobuzio-Donahue, R. Chandwani, C.S. Leslie, S.D. Leach, Nature Communications. 12 (2021) 3044. https://doi.org/10.1038/s41467-021-23237-2.

[71] N. Kosaka, H. Iguchi, T. Ochiya, Cancer Science. 101 (2010) 2087–2092. https://doi.org/10.1111/j.1349-7006.2010.01650.x.

[72] F. Janku, Cancer Treatment Reviews. 59 (2017) 93–101. https://doi.org/10.1016/j.ctrv.2017.07.005.

[73] J. Polivka, F. Janku, Pharmacology and Therapeutics. 142 (2014) 164–175. https://doi.org/10.1016/j.pharmthera.2013.12.004.

[74] D.A. Fruman, C. Rommel, Nature Reviews Drug Discovery. 13 (2014) 140–156. https://doi.org/10.1038/nrd4204.

[75] D.M. Hyman, L.M. Smyth, M.T.A. Donoghue, S.N. Westin, P.L. Bedard, E.J. Dean, H. Bando, A.B. El-Khoueiry, J.A. Pérez-Fidalgo, A. Mita, J.H.M. Schellens, M.T. Chang, J.B. Reichel, N. Bouvier, S.D. Selcuklu, T.E. Soumerai, J. Torrisi, J.P. Erinjeri, H. Ambrose, J.C. Barrett, B. Dougherty, A. Foxley, J.P.O. Lindemann, R. Mcewen, M. Pass, G. Schiavon, M.F. Berger, S. Chandarlapaty, D.B. Solit, U. Banerji, J. Baselga, B.S. Taylor, Journal of Clinical Oncology AKT Inhibition in Solid Tumors With AKT1 Mutations, J Clin Oncol. 35 (2017) 2251– 2259. https://doi.org/10.1200/JCO.

[76] F. Janku, D.S. Hong, S. Fu, S.A. Piha-Paul, A. Naing, G.S. Falchook, A.M. Tsimberidou, V.M. Stepanek, S.L. Moulder, J.J. Lee, R. Luthra, R.G. Zinner, R.R. Broaddus, J.J. Wheler, R. Kurzrock, Cell Reports. 6 (2014) 377–387. https://doi.org/10.1016/j.celrep.2013.12.035.

[77] D.J. Kwiatkowski, T.K. Choueiri, A.P. Fay, B.I. Rini, A.R. Thorner, G. de Velasco, M.E. Tyburczy, L. Hamieh, L. Albiges, N. Agarwal, T.H. Ho, J. Song, J.C. Pignon, P.M. Barrios, M.D. Michaelson, E.M. van Allen, K.M. Krajewski, C. Porta, S.K. Pal, J. Bellmunt, D.F. McDermott, D.Y.C. Heng, K.P. Gray, S. Signoretti, Clinical Cancer Research. 22 (2016) 2445–2452. https://doi.org/10.1158/1078-0432.CCR-15-2631.

[78] H. Cheng, Y. Zou, J.S. Ross, K. Wang, X. Liu, B. Halmos, S.M. Ali, H. Liu, A. Verma, C. Montagna, A. Chachoua, S. Goel, E.L. Schwartz, C. Zhu, J. Shan, Y. Yu, K. Gritsman, R. Yelensky, D. Lipson, G. Otto, M. Hawryluk, P.J. Stephens, V.A. Miller, B. Piperdi, R. Perez- Soler, Cancer Discovery. 5 (2015) 1262–1270. https://doi.org/10.1158/2159-8290.CD-14-0971.

[79] J.R. de Jesus, T. de Araújo Andrade, Metallomics. 12 (2020) 1912–1930. https://doi.org/10.1039/d0mt00220h.

[80] X. Ou, Y. Liu, X. Lei, P. Li, D. Mi, L. Ren, L. Guo, R. Guo, T. Chen, J. Hu, Z. Xiang, Z. Mu, X. Chen, J. Chen, K. Hu, Q. Jin, J. Wang, Z. Qian, Nature Communications. 11 (2020) 1-12. https://doi.org/10.1038/s41467-020-15562-9.

[81] A. Tsatsakis, D. Petrakis, T.K. Nikolouzakis, A.O. Docea, D. Calina, M. Vinceti, M. Goumenou, R.N. Kostoff, C. Mamoulakis, M. Aschner, A.F., Food and Chemical Toxicology. 141 (2020) 1-15. https://doi.org/10.1016/

j.fct.2020.111418.

[82] N. Khan, X. Chen, J.D. Geiger, Viruses. 12 (2020) 1-25. https://doi.org/10.3390/v12040471.

[83] G.A. Heeren, J.B. Jemmott, L. Sidloyi, Z. Ngwane, J.C. Tyler, Vulnerable Children and Youth Studies. 7 (2012) 47–54. https://doi.org/10.1080/17450128.2012.656733.

[84] M. Li, W. Liu, T. Bauch, E.A. Graviss, R.C. Arduino, J.T. Kimata, M. Chen, J. Wang, Nature Communications. 11 (2020) 1-15. https://doi.org/10.1038/s41467-020-17753-w.

[85] B.E. Feleke, T.E. Feleke, D. Mekonnen, M.B. Beyene, Clinical Nutrition ESPEN. 31 (2019) 56–60. https://doi.org/10.1016/j.clnesp.2019.03.001.

[86] A. Tapley, N. Switz, C. Reber, J.L. Davis, C. Miller, J.B. Matovu, W. Worodria, L. Huang, D.A. Fletcher, Journal of Clinical Microbiology. 51 (2013) 1774–1778. https://doi.org/10.1128/JCM.03432-12.

[87] A. Ayanlade, I.J. Nwayor, C. Sergi, O.S. Ayanlade, P. di Carlo, O.D. Jeje, M.O. Jegede, Scientific Reports. 10 (2020) 1-13. https://doi.org/10.1038/s41598-020-71094-8.

[88] M.G. Guzman, D.J. Gubler, A. Izquierdo, E. Martinez, S.B. Halstead, Nature Reviews Disease Primers. 2 (2016) 1-25. https://doi.org/10.1038/nrdp.2016.55.

[89] J. Lee, Y.E. Kim, H.Y. Kim, M. Sinniah, C.K. Chong, H.O. Song, Scientific Reports. 5 (2015) 1-10. https://doi.org/10.1038/srep18077.

[90] J.M. Dietz, S.J. Hankerson, B.R. Alexandre, M.D. Henry, A.F. Martins, L.P. Ferraz, C.R. Ruiz-Miranda, Scientific Reports. 9 (2019) 1-13. https://doi.org/10.1038/s41598-019-49199-6.

[91] M.S. Cunha, R.M. Tubaki, R.M.T. de Menezes, M. Pereira, G.S. Caleiro, E. Coelho, L. del C. Saad, N.C.C. de A. Fernandes, J.M. Guerra, J.S. Nogueira, J.L. Summa, A.A.C. Coimbra, T. Zwarg, S.S. Witkin, L.F. Mucci, M. do C.S.T. Timenetsky, E.C. Sabino, J.T. de Deus, Scientific Reports. 10 (2020) 1-8. https://doi.org/10.1038/s41598-020-72794-x.

[92] A.F. Versiani, E.M.N. Martins, L.M. Andrade, L. Cox, G.C. Pereira, E.F. Barbosa- Stancioli, M.L. Nogueira, L.O. Ladeira, F.G. da Fonseca, Scientific Reports. 10 (2020) 1-17. https://doi.org/10.1038/s41598-020-68357-9.

[93] M.I. Giraldo, H. Xia, L. Aguilera-Aguirre, A. Hage, S. van Tol, C. Shan, X. Xie, G.L. Sturdevant, S.J. Robertson, K.L. McNally, K. Meade-White, S.R. Azar, S.L. Rossi, W. Maury, M. Woodson, H. Ramage, J.R. Johnson, N.J. Krogan, M.C. Morais, S.M. Best, P.Y. Shi, R. Rajsbaum, Nature. (2020) 1-23. https://doi.org/10.1038/s41586-020-2457-8.

[94] N.A. Prow, L. Liu, M.K. McCarthy, K. Walters, R. Kalkeri, J. Geiger, F. Koide, T.H. Cooper, P. Eldi, E. Nakayama, K.R. Diener, P.M. Howley, J.D. Hayball, T.E. Morrison, A. Suhrbier, Npj Vaccines. 5 (2020) 1-10. https://doi.org/10.1038/s41541-020-0191-8.

[95] H. Hwang, B.Y. Hwang, J. Bueno, Disease Markers. 2018 (2018) 1-2. https://doi.org/10.1155/2018/8509127.

[96] J.F.W. Chan, K.H. Kok, Z. Zhu, H. Chu, K.K.W. To, S. Yuan, K.Y. Yuen, Emerging Microbes and Infections. 9 (2020) 221–236. https://doi.org/10.1080/22221751.2020.1719902.

[97] Y.J. Tan, S.G. Lim, W. Hong, Antiviral Research. 65 (2005) 69–78. https://doi.org/10.1016/j.antiviral.2004.10.001.

[98] L. Zhang, H. Guo, Advances in Biomarker Sciences and Technology. 2 (2020) 1–23. https://doi.org/10.1016/j.abst.2020.08.001.

[99] N. le Bert, A.T. Tan, K. Kunasegaran, C.Y.L. Tham, M. Hafezi, A. Chia, M.H.Y. Chng, M. Lin, N. Tan, M. Linster, W.N. Chia, M.I.C. Chen, L.F. Wang, E.E. Ooi, S. Kalimuddin, P.A. Tambyah, J.G.H. Low, Y.J. Tan, A. Bertoletti, Nature. 584 (2020) 457–462. https://doi.org/10.1038/s41586-020-2550-z.

[100] A.C. Walls, Y.J. Park, M.A. Tortorici, A. Wall, A.T. McGuire, D. Veesler, Structure, Cell. 181 (2020) 281-292.e6. https://doi.org/10.1016/j.cell.2020.02.058.

[101] E. Andreano, R. Rappuoli, Nature Medicine. 27 (2021) 759–761. https://doi.org/10.1038/s41591-021-01347-0.

[102] N. Xiao, M. Nie, H. Pang, B. Wang, J. Hu, X. Meng, K. Li, X. Ran, Q. Long, H. Deng, N. Chen, S. Li, N. Tang, A. Huang, Z. Hu, Nature Communications. 12 (2021) 1-13. https://doi.org/10.1038/s41467-021-21907-9.

[103] J. Huang, W. Liang, S. Chen, Y. Zhu, H. Chen, C.K.P. Mok, Y. Zhou, Serum Cytokine Profiles in Patients with Dengue Fever at the Acute Infection Phase, Disease Markers. (2018) 1-8. https://doi.org/10.1155/2018/8403937.

[104] D.J. Hunter, E. Losina, A. Guermazi, D. Burstein, M.N. Lassere, V. Kraus, Current Drug Targets. 11 (2010) 536-545.

[105] V.O. Puntmann, Postgraduate Medical Journal. 85 (2009) 538–545. https://doi.org/10.1136/pgmj.2008.073759.

[106] K. E. Lindsay et al, Nature Biomedical Engineering. 3 (2019) 371-380. https://doi.org/10.1038/s41551-019-0378-3.

[107] A.J. Walsh, J.T. Sharick, M.C. Skala, Nature Biomedical Engineering. 3 (2019) 333–334. https://doi.org/10.1038/s41551-019-0398-z.

[108] L. Zhang, L. Shi, Y. Shen, Y. Miao, M. Wei, N. Qian, Y. Liu, W. Min, Nature Biomedical Engineering. 3 (2019) 402–413. https://doi.org/10.1038/s41551-019-0393-4.