IntechOpen

Gene Expression and Control

Edited by Fumiaki Uchiumi

Gene Expression and Control

Edited by Fumiaki Uchiumi

Published in London, United Kingdom







IntechOpen






Supporting open minds since 2005

Gene Expression and Control
http://dx.doi.org/10.5772/intechopen.76958
Edited by Fumiaki Uchiumi

Contributors
Edio Maldonado, Diego A. Rojas, Sandra Moreira-Ramos, Fabiola Urbina, Fumiaki Uchiumi, Masashi Asai, Fei Yan, Meng Du, Zhiyi Chen, Johan Osorio, Sonia Moisa, William E Bentley, Sally Wang, Naida Hadziabdic, Amina Kozaric, Martin Berger, Doaa Ali, Francesc Posas, Mariona Nadal, Carme Sole, Gerard Martinez-Cebrian, Eulàlia De Nadal, Ljubica Caldovic

First published in London, United Kingdom, 2019 by IntechOpen
eBook (PDF) Published by IntechOpen, 2019
IntechOpen is the global imprint of INTECHOPEN LIMITED, registered in England and Wales, registration number: 11086078, The Shard, 25th floor, 32 London Bridge Street
London, SE19SG – United Kingdom
Printed in Croatia

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Additional hard and PDF copies can be obtained from orders@intechopen.com

Gene Expression and Control
Edited by Fumiaki Uchiumi
p. cm.
Print ISBN 978-1-78985-517-3
Online ISBN 978-1-78985-518-0
eBook (PDF) ISBN 978-1-83962-113-0

Meet the editor

Fumiaki Uchiumi, Professor of Pharmaceutical Sciences, Tokyo University of Science, grew up in Japan and received his bachelor's degree (Chemistry) from Tokyo University of Science in 1987. In 1993, after obtaining his Ph.D. degree (Molecular Biology) from Tokyo University, he joined Professor S. Tanuma's Laboratory at Tokyo University of Science as an Assistant Professor. He obtained his second Ph.D. (Pharmaceutical Science) from Tokyo University of Science in 1999. In 2000, he was promoted to the position of Lecturer at Tokyo University of Science. He then went abroad as a post-doctoral researcher of The United States-Japan Cooperative Cancer Research Program at Professor E. Fanning's Laboratory in Vanderbilt University, 2000-2001. Fumiaki Uchiumi was promoted to Associate Professor and then Full Professor at Tokyo University of Science in 2010 and 2016, respectively.

Contents

Preface

In accordance with progress in genome informatics, developments in molecular biology have greatly changed medical sciences. For example, diagnosis of certain diseases, including cancer, could be carried out by genome analysis, which accurately shows every mutated nucleotide in individual genomes. Moreover, genome editing technologies are expected to be clinically applied on some diseases, such as Duchenne muscular dystrophy. Recent studies revealed that specific human diseases are frequently accompanied with dysregulation in transcription or gene expression. In this regard, understanding mechanisms of the transcription system is essential for development in medical and life sciences. This will direct us as to how we could apply gene expression control for practical uses, not only in diagnosis and therapeutics but also in industrial production of recombinant cytokines, antibodies, and hormones.

In Section 1 of this book, pathologically important roles of the transcriptional regulation are described. Inner cellular signalling pathway plays important roles in the regulation of various cellular behaviours.

In Section 2, Chapter 2 discusses in detail how MAPK induced signals affect transcriptional initiation, elongation, and termination. The studies in MAPK-mediated transcriptional controlling system will contribute to novel therapeutics for cancer, immunological and metabolic diseases. Not only cell surviving or proliferation-causing signals, but also dysregulation in autophagy could play causative roles in generation of cancer and neurodegenerative diseases. In Chapter 3, the possible relationships between expression of the autophagy associated factor-encoding genes and nicotinic acetylcholine receptors are discussed. In Chapter 4, we explain dysregulated transcription in odontogenic cysts. The author discusses if the over-expressed genes, including *PTCH*, or their encoding proteins could be the right targets for medical treatment.

In Section 3, we look over transcriptional control. Current advances in molecular biology enabled accumulation of data from genomic, transcriptomic, and proteomic analyses. In Chapter 5, we learn how to utilize and apply such data on the examination of the transcriptional control system. The author performed data-mining from ChIP-seq analyses to discuss the transcriptional control of urea-cycle factor encoding gene expression. If scientists find some key transcription factors, they might be applied on gene-therapy. In Chapter 6, the authors argue for the gene delivery system, which enables pin-point therapeutics on pathological tissue. The transcriptional control is not only applied for medical subjects, but also with food supply. In Chapter 7, the author presents a platform to discuss nutrigenomics, which will contribute to increase the production rate of meat and milk. The concept might be relevant to prevention of diseases, including diabetes and aging-associated syndromes.

Although applications of the transcriptional control on life sciences are making rapid progress, basic study is necessary for further progress toward a new era

of molecular biology. In Section 4, transcriptional control in microorganisms is reviewed. In Chapter 8, authors successfully review the transcription initiation in ribosomal protein-encoding genes in yeast cells. The concept encourages focus on transcription from TATA-less promoters to overview landscapes of the *Central Dogma*. In prokaryotic cells, quorum sensing plays a role in regulation of cell-cell communication system by affecting transcription. In Chapter 9, authors faithfully address the mechanisms of how environmental signals control gene expression in prokaryotic cells.

"Examine thoroughly not only enemies but also ourselves, and we will win every battle." This is a saying from a general in ancient China. Now, what are the enemies? Molecular biologists who are studying transcriptional systems might define them as human diseases or problems/difficulties in producing recombinant proteins. Now, based on the scientific discussion, we are reaching the point to look over life on earth. In the near future, artificial intelligence (AI) will be applied not only to basic studies but also to clinical purposes. We should always be aware what science is aiming for.

Fumiaki Uchiumi, Ph.D.
Professor, Department of Gene Regulation,
Tokyo University of Science,
Japan

Section 1

Introduction

Chapter 1

Introductory Chapter: Gene Expression Controlling System and Its Application to Medical Sciences

Fumiaki Uchiumi and Masashi Asai

1. Introduction

We have learned that genes in mammalian cells are transcribed into messenger RNAs (mRNAs), which are to be translated into polypeptides (proteins). This is known as "Central Dogma." Gene expression must be appropriately maintained to regulate development, differentiation, and proliferation of cells. Imbalances or disturbances in gene expression are sometimes deleterious for living things. For example, steroid and thyroid hormones directly bind to nuclear receptors, which induce expression of specific genes. Recent global analyses of gene transcripts revealed that specific transcription factors (TFs) and their networking systems physiologically correspond to the onset of human diseases, including cancer. In other words, expression of specific genes might have relevance to pathogenesis of diseases. Given that OKSM (Yamanaka) factors convert somatic cells into induced pluripotent stem (iPS) cells, alterations in transcriptional state could affect destiny of the cells. In this chapter, revisiting known TFs, we would argue if transcription controlling strategies could contribute for the novel therapies on human diseases.

2. Transcription factors (TFs) in mammalian cells

Transcription factors are divided into two groups. First, the general TFs (GTFs), including preinitiation complex components TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and THIIH, are the primary protein factors that are required for the initiation of transcription from the TATA box (or TATA element), then elongation is executed by RNA polymerase II (RNA pol II) [1]. The others are the site-specific TFs or the DNA sequence-specific binding proteins.

2.1 The GTFs and TATA-less promoters

Molecular mechanisms of the initiation of transcription from TATA box have been well known as the most essential nuclear events in mammalian cells. However, about 70% of mammalian gene promoters have no TATA or TATA-like elements, and they are referred to as TATA-less promoters [2]. In yeast cells, most of genes are regulated by the same general TF-dependent system independent of TATA or TATA-like sequences [3]. TATA-containing and TATA-less promoters

are different in the distribution of the A-tracts, G-quadruplex, and CpG islands [4]. Given that most of the housekeeping genes are controlled by TATA or TATA-like elements, TATA-less promoters would be sensitive to the environments in response to various stresses. For example, some of the TATA-less promoters have DTIE binding sites [5]. The duplicated GGAA (TTCC) elements are frequently found near the transcription start sites (TSSs) of the human TATA-less promoters [6], implying that GGAA (TTCC) binding TFs contribute to the initiation of transcription.

2.2 The site-specific transcription factors

Transcription mediator interplays with the transcription activator proteins and the GTFs to enable plasticity and flexibility of gene expression [7]. However, site-specific TF-mediated and TATA-independent transcription system should be investigated because it may control responses to stresses. The TFs are categorized by their amino acid sequences or domain structures according to TRANSFAC (www.edgar-wingender.de/huTF_classification.html) database [8]. Some proteins might have multiple characteristic motifs.

2.2.1 Basic helix-loop-helix (bHLH) proteins

The bHLH proteins contain a basic region and HLH motif, which is comprised of two α-helices separated by a loop [9]. The bHLH family, BMAL, C/EBP, CLOCK, c-MYC, MYOD, NPAS2, and SREBP1/2, regulates development of mesodermal and neural tissues.

2.2.2 Helix-turn-helix (HTH) proteins

The ETS, FOX, IRF, HOX, HSF, POU, and PAX proteins belong to the HTH protein group. The HTH domains are present in NANOG, OCT1, PDN, and FLI1 [10]. ETS family proteins [11–13], which bind to GGAA (TTCC)-core motifs, contain ETS domain [14]. They regulate oncogenesis, development, differentiation, and apoptosis. The homeodomain (HD) is classified into 16 major subclasses, including ANTP, HNF, POU, PRD, SIX/SO, and ZF [15]. The HD proteins, including OTFs [16], which consist about 15–30% of TFs in animals, regulate differentiation and development of organisms.

2.2.3 High-mobility group (HMG) proteins and heteromeric CCAAT-binding factors

This group includes SOX [17] and NF-Y [18] proteins. They are thought to be the key TFs to regulate differentiation of cells, development of organs, and aging.

2.2.4 Immunoglobulin-fold proteins

This group, including NF-κB, IκB, NFAT, STAT, and p53, plays roles in transduction of biologically important signals to nuclei in response to cytokine-induced stimulation, viral infection, DNA-damage, and nutrients.

2.2.5 Leucine zipper proteins (bZIP)

The bZIP family includes CREB/ATF, BACH1/2, and FOS/JUN. The FOS/JUN has been originally characterized as product of the proto-oncogene or immediate

early gene, which very quickly responds to neuronal stimulatory or cellular proliferation-inducing signals [19].

2.2.6 Zinc finger motif containing proteins

The zinc finger motifs, which could be separated into major two classes, (1) Cys_4 and (2) Cys_2His_2, are the major motifs in numbers of TFs and DNA binding proteins [20].

1. The Cys_4-type proteins, including GRs, ERs, RARs, CREBBP, IGHMBP2, GATA, PPAR, are classified into zinc fingers ZZ (ZZZ), zinc fingers AN1 (ZFAND), and GATA zinc-finger domain-containing (GATAD) types.

2. The Cys_2His_2, or the zinc fingers C2H2 (ZNF) type, is the typical motif in the mammalian TFs. SP1, KLF4, KLF5, EGR3, and numbers of ZNF proteins belong to this class.

2.2.7 β-Structure (β-scaffold, β-sheet, and β-barrel) containing transcription factors

This group includes DBP, NF-1, HMGA, SMAD, and TBP. The RNase II and DNA glycosylases carry common motif, TBP-domain. The origin is thought to trace back to before the divergence of the three domains of life, bacteria, archaea, and eukaryote [21].

The classification implies that TATA-dependent initiation is not the standard but one of the site-specific transcriptions. For correct understanding of the cellular responses to differentiation/development-inducing signals, it should be examined how the site-specific TFs could initiate transcription.

3. Stresses and signals that regulate gene expression

Water-insoluble factors, steroids, and vitamins will easily go through lipid bilayer to bind to nuclear receptors. On the other hand, water-soluble compounds bind to the membrane receptors to transfer signals into cells, causing prompt response to induce certain signal cascade [22] to enhance/suppress specific gene expression as follows:

1. cyclic AMP (protein kinase A) pathway

2. MAP kinase pathways

3. Wnt signaling

4. TGFβ signaling

5. JAK/STAT pathway

6. Toll-like receptor signaling

7. immunoreceptor signaling

Cells are continuously receiving stresses, which are, for example, temperature, lights and radiations, proton and ionic gradient, nutrients, and pathogens, including bacteria and viruses. Therefore, responses to these stresses are essential for life [23]. The stress-induced signals will be converted to cellular responses, including secretion, gating ion channels, cellular behavior, proliferation, differentiation, senescence, apoptosis, and the development of organs. Some of the signals cause epigenetic changes or affect transcription of specific genes. We should remember that DNA methylations are required for setting accurate TSSs [24]. The epigenetic modification of DNAs, such as methylation, acetylation, and phosphorylation, and poly(ADP-ribosyl)ation, is dependent on the substrate molecules, S-adenosyl methionine, acetyl-CoA, ATP, and NAD^+, respectively [25]. They are the metabolites that we have learned from textbooks in biological chemistry. These molecules are so essential and indispensable for living things that they must be obtained from metabolism of nucleic acids, amino acids, lipids, and carbohydrates. Not only the epigenetic regulation, but also transcription is dependent on nutrients or the metabolites. For example, AMPK-FOXO pathway plays a role as a nutrient sensor to affect gene expression [26]. The AMPK has been shown to regulate NADPH homeostasis during energy stress [27]. Increased NAD^+/NADH ratio serves as a metabolic switch for transcription of the *BRCA1* gene [28]. The transcriptional corepressor protein, CtBP1, which possesses an NADH binding domain [29, 30], is one of the candidates that play a role in the NAD^+/NADH ratio-dependent transcription system. Importantly, several human DNA-repair-factor encoding gene promoters respond to natural compound *trans*-resveratrol (Rsv), which upregulates NAD^+/NADH ratio in HeLa S3 cells [31], implying that expression of the DNA-repair associated genes is partly regulated by metabolic state.

4. Analyses of gene expression in human diseases

The next-generation sequencing (NGS) analysis greatly contributed for the identification of specific genetic errors [32]. Somatic mutations on cancer driver genes have been identified [33, 34] and the statistical data contribute for the prediction or diagnosis of cancer. In search of biomarkers and cancer-causing factors, transcriptome analyses have been applied on model animals and clinical samples from patients, gradually unveiling mechanisms of the cancer progression. Hot spot somatic mutations are observed in the promoter regions of the human cancer genomes [35]. Binding of the TFs on the promoter could regulate mutation rate to modulate both transcription and DNA-repair systems [36]. In melanomas, such mutations are highly found on the human *TERT* gene promoter [37, 38]. Thus, cancer-causing mutations are not only present in the protein-coding genes, but also in the gene expression regulatory regions.

Transcriptome analysis does not only contribute to the diagnosis or prognosis of cancer but also other human diseases. For example, different gene expression patterns have been shown in autism [39], type 2 diabetes (T2DB) [40], schizophrenia [41], and neurodegenerative disease, including Alzheimer's disease (AD) [42, 43]. Accumulation of repeat containing RNAs in aberrant foci in nucleus has been observed in Huntington disease, muscular dystrophy, and amyotrophic lateral sclerosis (ALS) [44]. The noncoding RNAs (ncRNAs), the long ncRNAs (lncRNAs), or long intergenic ncRNAs (lincRNAs) play essential roles in epigenetic regulation and in conversion of the nuclear structures [45]. Therefore, not only the protein encoding genes but also the ncRNAs are thought to be involved in the pathogenesis [46–48]. The lncRNAs control chromatin structure and nuclear architecture to regulate transcription in eukaryotic cells [49]. Thus, it is very important to explore how the lncRNAs are being regulated. Conversely, introduction of lncRNAs could be applied for treatment of specific diseases.

5. Application of transcriptional control on medical sciences

Varieties of human diseases could be caused by the dysregulation in transcription of either or both protein-encoding genes and ncRNAs. Although, it has not been examined if alteration in transcription is applicable for treatment of the presently intractable diseases, it will be possible in principle. A number of natural and chemical compounds could affect gene expression. Alternatively, transcription could be modulated by the introduction of expression vectors, which express specific TFs in the target cells. In an effort to reach that goal, we have to develop reliable methods to deliver genes into abnormal or lesion cells.

5.1 Natural and the derivative compounds that regulate transcription

Natural compounds or phytochemicals could be applied on treatment of specific diseases. Some of them affect the JAK/STAT signaling pathway that activates cell cycle/proliferation-controlling factor-encoding gene expression [50]. Sulforaphane targets Nrf2 [51], which responds to oxidative stress, implying that it can be applied for the treatment of metabolic syndromes [52]. Tannic acid, coumarins, and chalcones suppress 12-*O*-tetra decanoylphorbol-13-acetate (TPA)-induced HIV promoter activity in human Jurkat cells [53]. Rsv induces promoter activities of DNA-repair factor-encoding genes, including *TP53*, *WRN*, *TERT*, and *HELB* [54–56]. The GGAA-containing motif and the GC-box have been suggested to play an essential role in the response. Vitamin E and the related compounds, including tocotrienols, activate transcription by binding to estrogen receptors [57] but inactivate NF-κB activity [58]. Curcumin, which is a diferuloylmethane from the Indian spice turmeric, also targets NF-κB and AP1 to affect survival and proliferation of cells [59], might be applied on the suppression of the *APP* and the beta-site APP cleaving enzyme 1-encoding *BACE1* gene expression [60].

These observations imply that phytochemicals could be used for therapeutic use. To minimize side effects, how the molecular mechanisms affect transcription should be elucidated. Next-generation therapies for intractable human diseases might be enabled by introduction of TF-expression vectors or by editing specific site(s) of the genome. Additionally, nucleotides or nucleic acid-based pharmaceutical compounds might be developed to ameliorate profile or status of gene expression in dysregulated cells.

5.2 Dysregulation in transcription might cause aging-related diseases

Indeed, aging is not a disease in principle. However, incidences of specific diseases are increased according to the process, for example, arteriosclerosis, cancer, sarcopenia, and neurodegenerative diseases, including AD. The causations of aging have been studied and discussed at the cellular level and they could be classified as follows:

5.2.1 Genome instability, including telomere shortening

Aging is thought to be accelerated by the accumulation of damage on chromosomal DNAs. This hypothesis is supported by the identification of responsive genes for premature aging, for example, LMNA, WRN, ATM, and SRPTN [61, 62]. Therefore, the aberrances in the nuclear architecture and DNA repair systems may cause premature aging. Cellular senescence is generally accompanied with the shortening of telomeres [63]. This is consistent with the observations that shortened telomeres cause genomic instabilities.

5.2.2 Accumulation of damage from oxidative stresses

Neurotoxicity could be induced by numbers of factors. The oxidative stress or the oxidative damage would cause neurodegenerative diseases, including the AD. The oxidative stresses impact on the Keap1-NRF2-ARE signaling pathway to upregulate expression of stress response genes, such as *NRF2*, *KEAP1*, *SOD1*, and *ETS1* [64]. The observation is consistent with the theory that the incidence of the AD increases with the aging [65].

5.2.3 Mitochondrial dysregulations and dysfunctions

Morphology of mitochondria and the respiratory functions are declined by aging process, accompanied with the accumulation of mitochondrial DNAs (mtDNAs) [66, 67]. The mutations on the mtDNAs are thought to be partly caused by the oxidative stresses. Nuclear-mitochondrial communication is thought to play a role in the regulation of longevity [68]. Moreover, telomere dysfunctions induce mitochondrial compromise [69]. In summary, interplay between mitochondria and nuclei would affect aging.

5.2.4 Decrease in the cellular NAD^+ level

Nicotine amide adenine dinucleotide (NAD^+), which is required for the respiratory functions of mitochondria, regulating enzymes in TCA cycle, and oxidative phosphorylation (OXPHOS), is declined with aging [70, 71]. The NAD^+ activates protein deacetylase Sirtuins, which regulate health span and life span. Moreover, the NAD^+ is a substrate for poly(ADP-ribose) (PARP), which plays an essential role in the DNA damage response [31].

5.2.5 Alteration in the epigenetic regulations of the genomes

The profiles of the methylation of DNAs and the modifications, including acetylation and methylation of histones, are altered in the aging process [72, 73]. These epigenetic changes will enhance or reduce transcription of specific genes.

5.2.6 Alterations in immune response and cytokine dysregulation

Cytokine dysregulation or prolonged inflammation is observed in aged person [74]. Numbers of proteins are involved in the inflammation process, suggesting that expression and degradation of proteins could accelerate or decelerate cellular senescence.

Aging might start from subtle changes in the TFs profile that regulate expression of genes encoding DNA-repair/mitochondrial/immune response-associated protein factors [31, 75]. Decline in metabolites, which target DNA-binding proteins and ncRNAs [46–48] to modulate epigenetic systems [24–30], would accelerate aging. Given that alterations in the transcription could cause the aging, it could be slowed by manipulating TFs profile.

6. Concluding remarks and future prospects

Homeostasis or negative feedback is not only required for endocrine system but also for biochemical reactions. Metabolites, including sugars, lipids and amino acids, and ionic substrates need to be maintained within biologically significant

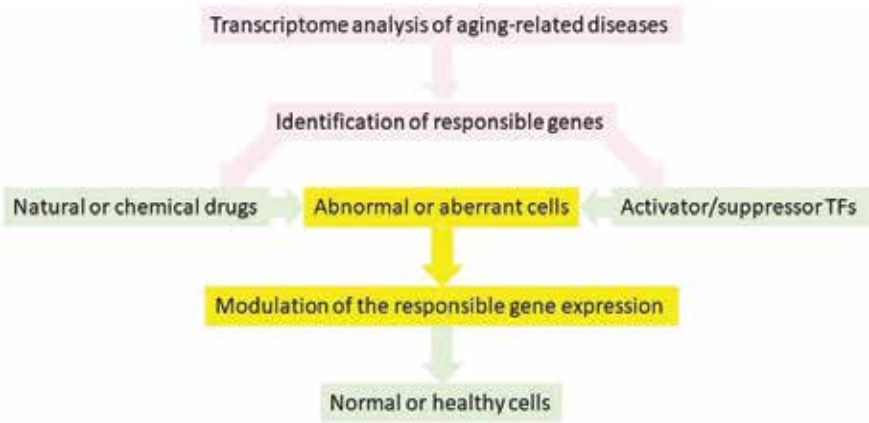

Figure 1.
Concept of the next-generation transcriptome-based therapeutics.

level. Similarly, nucleic acids and proteins should be appropriately managed being continuously synthesized and degraded. The equilibrium will gradually change in accordance with aging. Homeostasis in organs would play a role in the cancer generation [76]. Senescence can be metaphorically expressed as a walk on a balance beam that is narrowed afterward. The younger cells proliferate faster and respond to stresses with more accuracy than the older ones. After repeated proliferation, cells will reach the point where they cannot go forward but initiate aging. Somatic cells can acquire a pluripotency by incorporation of TFs. So, is it not possible to reverse aged or abnormal cells into younger or healthy cells? Then, what molecules (nucleic acids, proteins, etc.) are required and how they are introduced (by specific vectors, genome-editing systems, etc.) into cells? That should be answered for the next-generation therapeutics against aging-related diseases (**Figure 1**).

Acknowledgements

The authors are grateful to Hiroshi Hamada, Sakiko Kawahara, Natsuki Mizuno, Yui Suzuki, Yuki Nakano, Monami Kusaka, Kaori Orino, and Daisuke Sudo for discussion and their outstanding technical assistance.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

Abbreviations

ALS	amyotrophic lateral sclerosis
AD	Alzheimer's disease
GTF	general transcription factor
HD	homeodomain
NAD^+	nicotine amide adenine dinucleotide
bHLH	helix-loop-helix

HTH	helix-turn-helix
lncRNA	long noncoding RNA
mtDNA	mitochondrial DNA
ncRNA	noncoding RNA
TF	transcription factor
TSS	transcription start site

Author details

Fumiaki Uchiumi* and Masashi Asai
Department of Gene Regulation, Faculty of Pharmaceutical Sciences,
Tokyo University of Science, Noda, Chiba-ken, Japan

*Address all correspondence to: uchiumi@rs.noda.tus.ac.jp

References

[1] Carey MF, Peterson CL, Smale ST. A primer on transcriptional regulation in mammalian cells. In: Carey MF, Peterson CL, Smale ST, editors. Transcriptional Regulation in Eukaryotes: Concepts, Strategies, and Techniques. 2nd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 2009. pp. 1-45

[2] Yang C, Bolotin E, Jiang T, Sladek FM, Martinez E. Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters. Gene. 2007;**389**(1):52-65

[3] Hahn S, Donczew R. Mechanistic differences in transcription initiation at TATA-less and TATA-containing promoters. Molecular and Cellular Biology. 2018;**38**(1):e00448-e00417

[4] Yella VR, Bansal M. DNA structural features of eukaryotic TATA-containing and TATA-less promoters. FEBS Open Bio. 2017;**7**:324-334

[5] Marbach-Bar N, Bahat A, Ashkenazi S, Golan-Mashiach M, Haimov O, Wu SY, et al. DTIE, a novel core promoter element that directs start site selection in TATA-less genes. Nucleic Acids Research. 2016;**44**(3):1080-1094

[6] Uchiumi F, Miyazaki S, Tanuma S. The possible functions of duplicated ETS (GGAA) motifs located near transcription start sites of various human genes. Cellular and Molecular Life Sciences. 2011;**68**(12):2039-2051

[7] Soutourina J. Transcription regulation by the mediator complex. Nature Reviews. Molecular Cell Biology. 2018;**19**(4):262-274

[8] Carlberg C, Molnár F. Transcription factors. In: Carsten C, Molnar F, editors. Mechanisms of Gene Regulation. 2nd ed. Dordrecht: Springer Science+Business Media; 2016. pp. 57-73

[9] Gyoja F. Basic helix-loop-helix transcription factors in evolution: Roles in development of mesoderm and neural tissues. Genesis. 2017;**55**(9):e23051

[10] Wintjens R, Rooman M. Structural classification of HTH DNA-binding domains and protein-DNA interaction modes. Journal of Molecular Biology. 1996;**262**:294-313

[11] Oikawa T, Yamada T. Molecular biology of the ETS family of transcription factors. Gene. 2003;**303**:11-34

[12] Hsu T, Trojanowska M, Watson DK. ETS proteins in biological control and cancer. Journal of Cellular Biochemistry. 2004;**91**(5):896-903

[13] Wei GH, Badis G, Berger MF, Kivioja T, Palin K, Enge M, et al. Genome-wide analysis of ETS-family DNA-binding in vitro and in vivo. The EMBO Journal. 2010;**29**:2147-2160

[14] Karim FD, Urness LD, Thummel CS, Klemsz MJ, McKercher SR, Celada A, et al. The ETS-domain: A new DNA-binding motif that recognizes a purine-rich core DNA sequence. Genes & Development. 1990;**4**:1451-1453

[15] Bürglin TR, Affolter M. Homeodomain proteins: An update. Chromosoma. 2016;**125**(3):497-521

[16] Sturm RA, Herr W. The POU domain is a bipartite DNA-binding structure. Nature. 1988;**336**(6199):601-604

[17] Kamachi Y, Kondoh H. Sox proteins: Regulators of cell fate specification and differentiation. Development. 2013;**140**(20):4129-4144

[18] Dolfini D, Gatta R, Mantovani R. NF-Y and the transcriptional activation of CCAAT promoters. Critical Reviews in

Biochemistry and Molecular Biology. 2012;**47**(1):29-49

[19] Lamph WW, Wamsley P, Sassone-Corsi P, Verma IM. Induction of proto-oncogene *JUN/AP1* by serum and TPA. Nature. 1988;**334**(6183):474-476

[20] Cassandri M, Smirnov A, Novelli F, Pitolli C, Agostini M, Malewicz M, et al. Zinc-finger proteins in health and disease. Cell Death Discovery. 2017;**3**:17071

[21] Brindefalk B, Dessailly BH, Yeats C, Orengo C, Werner F, Poole AM. Evolutionary history of the TBP-domain superfamily. Nucleic Acids Research. 2013;**41**:2832-2845

[22] Heldin CH, Lu B, Evans R, Gutkind S. Signals and receptors. In: Cantley LC, Hunter T, Sever R, Thorner J, editors. Signal Transduction, Principles, Pathways, and Processes. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 2014. pp. 3-29

[23] Hotamisligil GS, Davis RJ. Cell signaling and stress responses. In: Cantley LC, Hunter T, Sever R, Thorner J, editors. Signal Transduction, Principles, Pathways, and Processes. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 2014. pp. 345-364

[24] Neri F, Rapelli S, Krepelova A, Incarnato D, Parlato C, Basile G, et al. Intrageneic DNA methylation prevents spurious transcription initiation. Nature. 2017;**543**(7643):72-77

[25] Gut P, Verdin E. The nexus of chromatin regulation and intermediary metabolism. Nature. 2013;**502**(7472):489-498

[26] Shin HJ, Kim H, Oh S, Lee JG, Kee M, Ko HJ, et al. AMPK-SKP2-CARM1 signalling cascade in transcriptional regulation of autophagy. Nature. 2016;**534**(7608):553-557

[27] Jeon SM, Chandel NS, Hay N. AMPK regulates NADPH homeostasis to promote tumor cell survival during energy stress. Nature. 2012;**485**(7400):661-665

[28] Di LJ, Fernandez AG, De Siervi A, Longo DL, Gardner K. Transcriptional regulation of *BRCA1* expression by a metabolic switch. Nature Structural & Molecular Biology. 2010;**17**(12):1406-1413

[29] Chinnadurai G. CtBP, an unconventional transcriptional corepressor in development and oncogenesis. Molecular Cell. 2002;**9**(2):213-224

[30] Blevins MA, Huang M, Zhao R. The role of CtBP1 in oncogenic processes and its potential as a therapeutic target. Molecular Cancer Therapeutics. 2017;**16**(6):981-990

[31] Takihara Y, Sudo D, Arakawa J, Takahashi M, Sato A, Tanuma S, et al. Nicotinamide adenine dinucleotide (NAD^+) and cell aging. In: Strakoš R, Lorens B, editors. New Research on Cell Aging and Death. Hauppauge, NY: Nova Science Publishers; 2018. pp. 131-158

[32] Chang F, Liu GL, Liu CJ, Li MM. Somatic diseases (cancer): Amplification-based next-generation sequencing. In: Kulkarni S, Pfeifer J, editors. Clinical Genomics: A Guide to Clinical Next Generation Sequencing. London, UK: Academic Press; 2015. pp. 297-319

[33] Ku CS, Cooper DN, Ziogas DE, Halkia E, Tzaphlidou M, Roukos DH. Research and clinical applications of cancer genome sequencing. Current Opinion in Obstetrics & Gynecology. 2013;**25**(1):3-10

[34] Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA Jr, Kinzler KW. Cancer genome landscapes. Science. 2013;**339**(6127):1546-1558

[35] Perera D, Poulos RC, Shah A, Beck D, Pimanda JE, Wong JW. Differential DNA repair underlies mutation hotspots at active promoters in cancer genomes. Nature. 2016;**532**(7598):259-263

[36] Sabarinathan R, Mularoni L, Deu-Pons J, Gonzalez-Perez A, López-Bigas N. Nucleotide excision repair is impaired by binding of transcription factors to DNA. Nature. 2016;**532**(7598):264-267

[37] Huang FW, Hodis E, Xu MJ, Kryukov GV, Chin L, Garraway LA. Highly recurrent TERT promoter mutations in human melanoma. Science. 2013;**339**(6122):957-959

[38] Horn S, Figl A, Sivaramakrishna Rachakonda P, Fischer C, Sucker A, Gast A, et al. *TERT* promoter mutations in familial and sporadic melanoma. Science. 2013;**339**(6122):959-961

[39] Parikshak NN, Swarup V, Belgard TG, Irimia M, Ramaswami G, Gandal MJ, et al. Genome-wide changes in lncRNA, splicing, and regional gene expression patterns in autism. Nature. 2016;**540**(7633):423-427

[40] Wu C, Chen X, Shu J, Lee CT. Whole-genome expression analyses of type 2 diabetes in human skin reveal altered immune function and burden of infection. Oncotarget. 2017;**8**(21):34601-34609

[41] Gusev A, Mancuso N, Won H, Kousi M, Finucane HK, Reshef Y, et al. Transcriptome-wide association study of schizophrenia and chromatin activity yields mechanistic disease insights. Nature Genetics. 2018;**50**(4):538-548

[42] Zou F, Chai HS, Younkin CS, Allen M, Crook J, Pankratz VS, et al. Brain expression genome-wide association study (eGWAS) identifies human disease-associated variants. PLoS Genetics. 2012;**8**(6):e1002707

[43] Theuns J, Van Broeckhoven C. Transcriptional regulation of Alzheimer's disease genes: Implications for susceptibility. Human Molecular Genetics. 2000;**16**:2383-2394

[44] Jain A, Vale RD. RNA phase transitions in repeat expansion disorders. Nature. 2017;**546**(7657):243-247

[45] Schein A, Carninci P. Complexity of mammalian transcriptome analyzed by RNA deep sequencing. In: Kurokawa R, editor. Long Noncoding RNAs. Dordrecht: Springer Science+Business Media; 2015. pp. 3-22

[46] Takayama K, Inoue S. The role of androgen-regulated long noncoding RNAs in prostate cancer. In: Kurokawa R, editor. Long Noncoding RNAs. Dordrecht: Springer Science+Business Media; 2015. pp. 191-210

[47] Tano K, Onoguchi-Mizutani R, Yeasmin F, Uchiumi F, Suzuki Y, Yada T, et al. Identification of minimal p53 promoter region regulated by MALAT1 in human lung adenocarcinoma cells. Frontiers in Genetics. 2018;**8**:00208

[48] Pandian GN, Syed J, Sugiyama H. Synthetic strategies to identify and regulate noncoding RNAs. In: Kurokawa R, editor. Long Noncoding RNAs. Dordrecht: Springer Science+Business Media; 2015. pp. 23-43

[49] Ransohoff JD, Khavari PA. The functions and unique features of long intergenic non-coding RNA. Nature Reviews. Molecular Cell Biology. 2018;**19**(3):143-157

[50] Arumuggam N, Bhowmick NA, Vasantha Rupasinghe HP. A review: Phytochemicals targeting JAK/STAT signaling and IDO expression in cancer. Phytotherapy Research. 2015;**29**:805-817

[51] Cuadrado A, Manda G, Hassan A, Alcaraz MJ, Barbas C, Daiber A, et al. Transcription factor NRF2 as a therapeutic target for chronic diseases: A systems medicine approach. Pharmacological Reviews. 2018;**70**:348-383

[52] Matzinger M, Fischhuber K, Heiss EH. Activation of Nrf2 signaling by natural products-can it alleviate diabetes? Biotechnology Advances. 2018;**36**:1738-1767

[53] Uchiumi F, Hatano T, Ito H, Yoshida T, Tanuma S. Transcriptional suppression of the HIV promoter by natural compounds. Antiviral Research. 2003;**58**:89-98

[54] Uchiumi F, Shoji K, Sasaki Y, Sasaki M, Sasaki Y, Oyama T, et al. Characterization of the 5′-flanking region of the human *TP53* gene and its response to the natural compound, resveratrol. Journal of Biochemistry. 2016;**159**:437-447

[55] Uchiumi F, Watanabe T, Hasegawa S, Hoshi T, Higami Y, Tanuma S. The effect of resveratrol on the Werner syndrome RecQ helicase gene and telomerase activity. Current Aging Science. 2011;**4**:1-7

[56] Uchiumi F, Arakawa J, Iwakoshi K, Ishibashi S, Tanuma S. Characterization of the 5′-flanking region of the human DNA helicase B (*HELB*) gene and its response to *trans*-resveratrol. Scientific Reports. 2016;**6**:24510

[57] Khallouki F, de Medina P, Caze-Subra S, Bystricky K, Balaguer P, Poirot M, et al. Molecular and biochemical analysis of the estrogenic and proliferative properties of vitamin E. Frontiers in Oncology. 2016;**5**:287

[58] Husain K, Francois RA, Yamauchi T, Perez M, Sebti SM, Malafa MP. Vitamin E δ-tocotrienol augments the anticancer activity of gemcitabine and suppresses constitutive NF-κB activation in pancreatic cancer. Molecular Cancer Therapeutics. 2011;**10**(12):2363-2372

[59] Fong D, Chan MM. Dietary phytochemicals target cancer stem cells for cancer chemoprevention. In: Chandra D, editor. Mitochondria as Target for Phytochemicals in Cancer Prevention and Therapy. Dordrecht: Springer Science+Business Media; 2013. pp. 85-125

[60] Lin R, Chen X, Li W, Han Y, Liu P, Pi R. Exposure to metal ions regulates mRNA levels of APP and BACE1 in PC12 cells: Blockage by curcumin. Neuroscience Letters. 2008;**440**:344-347

[61] Vijg J. Genome instability and accelerated aging. In: Vijg J, editor. Aging of the Genome. Oxford: Oxford University Press; 2007. pp. 151-180

[62] Lidzbarsky G, Gutman D, Shekhidem HA, Sharvit L, Atzmon G. Genomic instabilities, cellular senescence, and aging: In vitro, in vivo and aging-like human syndromes. Frontiers in Medicine. 2018;**5**:104

[63] Blackburn EH, Epel ES, Lin J. Human telomere biology: A contributory and interactive factor in aging, disease risks, and protection. Science. 2015;**350**(6265):1193-1198

[64] Ramkissoon A, Shapiro AM, Loniewska MM, Wells PG. Neurodegeneration from drugs and aging-derived free radicals. In: Villamena FA, editor. Molecular Basis of Oxidative Stress. Hoboken: John Wiley & Sons; 2013. pp. 237-309

[65] Toral-Rios D, Carvajal K, Phillips-Farfán B, Camacho-Castillo LC, Campos-Peña V. Oxidative stress, metabolic syndrome and Alzheimer's disease. In: Gelpi RJ, Boveris A, Poderoso JJ, editors. Biochemistry of Oxidative Stress. Physiopathology and

Clinical Aspects. Dordrecht: Springer International Publishing Switzerland; 2016. pp. 361-374

[66] Kauppila TES, Kauppila JHK, Larsson NG. Mammalian mitochondria and aging: An update. Cell Metabolism. 2017;**25**(1):57-71

[67] Cedikova M, Pitule P, Kripnerova M, Markova M, Kuncova J. Multiple roles of mitochondria in aging processes. Physiological Research. 2016;**65**(Supplementum 5):S519-S531

[68] Houtkooper RH, Mouchiroud L, Ryu D, Moullan N, Katsyuba E, Knott G, et al. Mitonuclear protein imbalance as a conserved longevity mechanism. Nature. 2013;**497**(7450):451-457

[69] Sahin E, Colla S, Liesa M, Moslehi J, Müller FL, Guo M, et al. Telomere dysfunction induces metabolic and mitochondrial compromise. Nature. 2011;**470**(7334):359-365

[70] Katsyuba E, Auwerx J. Modulating NAD^+ metabolism, from bench to bedside. The EMBO Journal. 2017;**36**(18):2670-2683

[71] Verdin E. NAD^+ in aging, metabolism, and neurodegeneration. Science. 2015;**350**(6265):1208-1213

[72] Johnson KC, Christesen BC. Genome-wide DNA methylation changes during aging. In: Fraga M, Fernández AF, editors. Epigenomics in Health and Disease. San Diego, CA: Academic Press; 2016. pp. 127-144

[73] Bainor AJ, David G. The dynamics of histone modifications during aging. In: Fraga M, Fernández AF, editors. Epigenomics in Health and Disease. San Diego, CA: Academic Press; 2016. pp. 145-162

[74] Rea IM, Gibson DS, McGilligan V, McNerlan SE, Alexander HD, Ross OA. Age and age-related diseases: Role of inflammation triggers and cytokines. Frontiers in Immunology. 2018;**9**:586

[75] Uchiumi F, Larsen S, Tanuma S. Possible roles of a duplicated GGAA motif as a driver cis-element for cancer-associated genes. In: iConcept, editor. Understand Cancer – Research and Treatment. Hong Kong: iConcept Press; 2016. pp. 1-25

[76] Basanta D, Anderson ARA. Homeostasis back and forth: An ecoevolutionary perspective of cancer. In: Swanton C, Bardelli A, Polyak K, Shah SP, Graham TA, editors. Cancer Evolution. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 2017. pp. 149-167

Section 2

Transcriptional Control in Human Diseases

Chapter 2

Shaping the Transcriptional Landscape through MAPK Signaling

Mariona Nadal-Ribelles, Carme Solé, Gerard Martínez-Cebrián, Francesc Posas and Eulàlia de Nadal

Abstract

A change in the transcriptional landscape is an equilibrium-breaking event important for many biological processes. Mitogen-activated protein kinase (MAPK) signaling pathways are dedicated to sensing extracellular cues and are highly conserved across eukaryotes. Modulation of gene expression in response to the extracellular environment is one of the main mechanisms by which MAPK regulates proteome homeostasis to orchestrate adaptive responses that determine cell fate. A massive body of knowledge generated from population and single-cell analyses has led to an understanding of how MAPK pathways operate. MAPKs have thus emerged as fundamental transcriptome regulators that function through a multi-layered control of gene expression, a process often deregulated in disease, which therefore provides an attractive target for therapeutic strategies. Here, we summarize the current understanding of the mechanisms underlying MAPK-mediated gene expression in organisms ranging from yeast to mammals.

Keywords: MAP kinases, signal transduction, transcription, gene expression, chromatin

1. Introduction

The intracellular matrix is physically separated from the dynamic extracellular environment; however, their functions are intimately coordinated in order to ensure cell adaptation and survival. Mitogen-activated protein kinase (MAPK) cascades sense and integrate extracellular cues through sequential activation of protein kinases. These highly conserved transduction pathways are involved in a myriad of fundamental cellular processes and determine cell fate. Misregulation of these signaling cascades has major consequences for numerous diseases such as cancer, diabetes, inflammatory, and immune response diseases.

About 300 genes encode signaling proteins directly involved in signal transduction, including their positive and negative regulators as well [1]. Upon cell stimulation, in order to adapt to an extracellular insult, these seemingly simple linear signaling pathways harbor the potential to target a large number of substrates of which many are involved in gene expression. In fact, MAPKs control every step studied to date of the highly dynamic process of gene expression. The overall

picture of MAPK pathway substrates and interactors is still far from complete; however, the knowledge generated over the last 20 years has allowed a more holistic understanding of the underlying mechanisms of MAPK-regulated transcription. Due to the growing interest in MAPK-biology and the sheer volume of literature available, in this chapter, we not only mainly focus on the main mammalian MAPK cascades in humans (ERK1/2, JNK, p38, and ERK5), but we also discuss the main findings regarding MAPK cascades in the model organism *Saccharomyces cerevisiae*.

2. MAP kinase pathways

MAPKs mediate the transmission of extracellular information through a series of consecutive chemical reactions that lead to the activation of a terminal MAPK to orchestrate the appropriate gene expression pattern. To date, four major MAPK signaling cascades have been characterized in mammals, which are named according to their MAPK components: extracellular signal-regulated kinase 1 and 2 (ERK1/2), c-Jun N-terminal kinase 1 to 3 (JNK), p38 α/β/γ/δ (p38), and ERK5. Apart from these main MAPKs, several atypical MAPKs have also been described (ERK 3/4, ERK 7/8, and NLK, among others) with less well-defined functions and distinct modes of activation [2].

2.1 MAPK activators

MAPKs can be activated in two different ways: (1) ligand-dependent that requires the physical interaction of a ligand (e.g., growth factors, hormones, or cytokines) with a receptor or (2) ligand-independent that mediates the signaling of physical stressors (e.g., radiation, injury, and osmotic pressure). In general, ERK1/2 responds to proliferative and survival stimuli such as growth factors, serum, and phorbol esters and, to a lesser extent, to ligands of G protein-coupled receptors (GPCRs) or cytokines, or to osmotic stress and microtubule disorganization. ERK5 is activated by growth factors [e.g., EGF, NGF, FGF-2, and brain-derived neurotrophic factor (BDNF)] and cytokines (e.g., Leukemia inhibitory factor—LIF) as well as by some stresses such as osmotic stress and hydrogen peroxide [3]. JNKs and p38 MAPKs are functionally related and are collectively named stress-activated protein kinases (SAPKs). The JNK pathway strongly responds to cytokines, growth factor deprivation, intracellular stimuli (e.g., DNA damage, cytoskeletal changes, oxidative, and ER stress), and extracellular stresses (e.g., UV radiation and osmotic stress) and less efficiently responds to stimulation by some GPCRs, serum, and growth factors [4]. Finally, p38 signaling has been shown to be consistently activated by a wide variety of environmental stresses and inflammation but to be inconsistently activated by insulin and growth factors in certain cell types [1].

2.2 Modular architecture of the MAPK signaling cascades

MAPK signaling is triggered by the stimulation of different membrane receptor families (e.g., receptor tyrosine kinases (RTKs), GPCRs, cytokine receptors, Ser/Thr kinase receptors, and membrane-bound stress sensors) that are coupled to the MAPK signaling cascades. Depending on the stimulus, the signal is transmitted downstream through small G proteins, kinases, or adaptor proteins that are the immediate upstream activators of the conventional MAPK signaling cascades.

A major and highly conserved feature of MAPK pathways is their central three-tiered core signaling module of sequentially activating kinases (**Figure 1**). In the first tier, a Ser/Thr kinase MAPKKK (MAP3K) is activated by the effectors

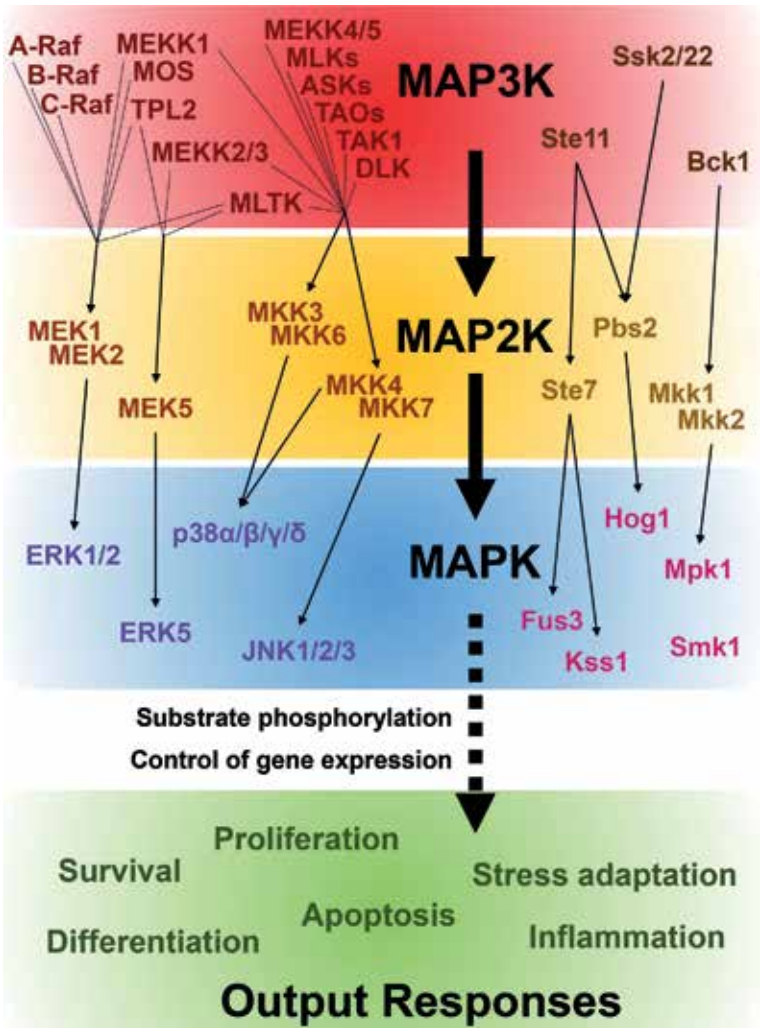

Figure 1.
Conceptual representation of the three core components of a MAPK signaling cascade. A typical MAPK cascade is composed of three consecutively activated tiers of kinases: MAP3Ks, MAP2Ks, and MAPKs. Kinases are grouped by layers according to their position in the signaling cascade. Arrows link components of different layers representing activation pathways. The core modules of mammalian (left) and yeast (right) signaling pathways are shown. Output responses resulting from MAPK activation through substrate phosphorylation and control of gene expression are indicated.

mentioned above. This MAPKKK then phosphorylates and activates a MAPKK (MAP2K) in the second tier; the MAPKK is a dual specificity kinase that phosphorylates both threonine and tyrosine within a conserved Thr-Xaa-Tyr motif in its substrate. Finally, there is a terminal Ser/Thr MAPK in the third tier, which, upon activation, phosphorylates a huge number of cytoplasmic and nuclear substrates on consensus Ser/Thr-Pro sites. Although not always present, other kinases are involved in MAPK signal transduction. One such kinase is the MAP4K that phosphorylates and activates the MAP3K; downstream MAPK-activated protein kinases (MAPKAPKs) contribute to the spread of the signal transduction.

The first MAPK pathway identified was ERK1 whose activation depends on the dimerization and autophosphorylation of the ligand-activated tyrosine kinase receptors (RTKs and GPCRs). These ligand-induced chemical and conformational receptor changes trigger recruitment of the adaptor proteins Shc and Grb2, a guanine exchange factor (SOS), and the small GTPase (Ras) to the plasma membrane.

The interaction of these four elements leads to the homo- and hetero-dimerization of the Raf family of kinases (B- or C-Raf) that activate the MAP3K module. The MAP3K then phosphorylates MEKK1/2 (MAP2K) at two serine within their activation loop (Ser-Met-Ala-Asn-Ser). Activated MEKK1/2 in turn phosphorylates ERK1/2 (MAPK) on the tyrosine and threonine residues of the Thr-Glu-Tyr motif in their activating loop. Additionally, MAPKAPKs have been identified that propagate ERK signaling (RSKs, MSKs, MNKs, and in some cases MK3/5) [1, 5].

The least studied of the four MAPK cascades is ERK5, whose mechanisms of upstream activation may include activation of tyrosine kinase receptors, the protein tyrosine kinase c-Src, the small GTPase Ras, the adaptor protein Lad1, and the protein Ser/Thr kinase WNK1, which acts as a MAP4K [1, 3]. Activation of these signaling molecules leads to activation of the MAP3Ks (not only MEKK2/3 but also TPL2 and MLTK) to phosphorylate the two alternatively spliced MEK5 isoforms (MEK5a and MEK5b, MAP2K) at the Ser-Xaa-Ala-Xaa-Thr activation motif, leading to ERK5 activation at the Thr-Glu-Tyr motif. The ERK5 pathway also involves downstream MAPKAPKs such as the serum and glucocorticoid-activated kinase (SGK) and p90 ribosomal S6 kinases (RSKs) [2].

The signal through the JNK cascade is transmitted through adaptor proteins (TRAFs), small GTPases (Rac1, Cdc42), or Ste20-like kinases that act as MAP4Ks [6]. A large number of MAP3Ks convey the signal to the main MAP2Ks (MKK4/7) by phosphorylating the sequence Ser-Xaa-Ala-Xaa-Ser/Thr in their activation loop [4]. Ultimately, the three components of the MAPK level (JNK1–3) are activated by dual Thr/Tyr phosphorylation at the Thr-Pro-Tyr motif. As for other kinases in the JNK cascade, MAPKAPKs such as MST1 are well-defined JNK substrates that can act as both upstream and downstream of JNK [7].

Finally, p38 operates through different receptors from apoptosis-related receptors to physical sensors. The initial signal is transferred using Cdc42, Rac1, and Ste20-like kinases (shared with JNK) and results in phosphorylation of the activation loop (Ser-Xaa-Ala-Xaa-Ser/Thr) of the MAP2Ks MKK3/6 that uniquely target p38. The differences between the p38 and JNK pathways lie within the specific scaffold proteins and substrates. All p38 isoforms, either the major isoforms (p38α,β,γ,δ) or the minor isoforms generated through alternative splicing, are activated through dual phosphorylation at the Thr-Gly-Tyr motif [1]. The main p38 isoform (p38α) is constitutively expressed, while the remaining isoforms are tissue-restricted. Uniquely for a MAP kinase, p38 can be activated through MAP2K-independent mechanisms that involve adaptors that promote p38 autophosphorylation [6]. Finally, the downstream MAPKAPK layer is partially shared with ERK and includes MAPKAPK2,3,5, MNKs, and MSKs [1].

2.2.1 Specificity of signaling cascades

The signaling proteome is composed of a limited number of genes that specifically integrate a virtually endless number of extracellular stimuli. Several strategies have evolved in order to maintain the signaling fidelity. For instance, this is achieved by the interaction of MAPKs with other components of the pathway and with substrates through docking sites composed of specific consensus motifs. Two types of docking motifs have been reported: D-motif and docking site for ERK (FXF)-motif, which ensure fidelity of signaling. D-motifs contain at least two basic residues flanking hydrophobic residues and are located opposite to the catalytic pocket in MAPKs [8]. The FXF-motif is composed of two Phe residues separated by one residue [9]. Another mechanism to gain specificity of signaling is the use of MAPK-scaffold proteins, which were first described in yeast (Ste5 and Pbs2) [10, 11]. Scaffolds are crucial for maintenance of signaling specificity

as they sequester multiprotein interactions to prevent crosstalk by controlling stability and subcellular localization.

2.2.2 Regulators of signaling cascades

The amplitude, frequency, and localization of activated MAPK-activity is tightly controlled, not only through positive and negative feedback mechanisms at the post-translational level mediated by regulatory proteins (e.g., phosphatases and kinases) but also through post-transcriptional control mediated by RNA-binding proteins and microRNAs (miRNAs).

The fastest mechanism of ablating MAPK activity is to remove one of the two activating phosphates through the activity of specific phosphatases. Their role in regulating the terminal MAPK has been extensively studied, but little is known about their effect on upstream signaling components. Phosphatase activity is mainly derived from Ser/Thr phosphatases, Tyr phosphatases, and the dual specificity phosphatases (DUSP) known as MAPK phosphatases (MKP) [1]. Based on sequence homology, substrate specificity, and subcellular localization, DUSPs can be divided into three groups: nuclear inducible (DUSP1/2/4/5), cytoplasmic and ERK-specific (DUSP6/7/9), and DUSPs with no specific cellular localization that targets JNK and p38 SAPKs (DUSP8/10/16) [4, 12]. MAPKs also exert a transcriptional control of regulatory elements such as these phosphatases and thereby generate a negative feedback loop. Another relevant type of negative feedback regulation is driven by the direct phosphorylation of different upstream components of the MAPK cascade by the MAPK itself to modulate basal [13] and stimuli-dependent signaling dynamics [5]. Additionally, scaffold proteins and other enzymatic activities either positively or negatively regulate different levels of MAPK signal transduction such as, for example, the formation of the ligand-receptor signaling complex, the intracellular modular interactions, and the degradation of the components [14]. Post-transcriptional regulation of MAPKs can also be achieved at the RNA level. RNA-binding proteins and miRNA negatively regulate MAPK gene expression by directly cleaving their mRNAs or through complementary pairing [15].

2.3 Yeast MAPK cascades

Five MAPK pathways have been well characterized in the budding yeast, *S. cerevisiae*. In vegetative cells, the four MAPKs Fus3, Kss1, Hog1, and Slt2/Mpk1 are involved in the mating-pheromone response, the filamentous-invasion pathway, the high osmolarity growth, and the cell integrity pathway, respectively. The fifth MAPK, Smk1, is believed to play a role in spore wall assembly [16, 17].

Haploid yeast cells sense the reciprocal mating pheromones (α-factor or a-factor) through Ste2 and Ste3 GPCRs. The signal is then transmitted by GTPases to the p21-activated kinase (PAK)-like kinase Ste20, the MAPK scaffold Ste5, Cdc42, a guanine-nucleotide exchange factor (GEF), and Far1. Ste5 signals and serves as a scaffold that links the MAP4K and the MAPK signalosome (Ste11 → Ste7 → Fus3; the latter is the ERK1 homolog) [18].

The high osmolarity glycerol (Hog1) MAPK, the yeast homolog of p38, is activated in response to osmotic stress as a consequence of signaling elicited from two upstream-independent mechanisms (Sln1/Sho1). The Sln1 sensor is the primary osmosensor and is a complex variation of the well-known bacterial two-component system. Upon osmostress, inactivation of the transmembrane histidine kinase Sln1 leads to the derepression and activation of the MAP3K (Ssk2/22) via Ypd1/Skk1. The Sho1 osmosensing branch is mediated by mucin-like proteins (Hkr1 and Msb2) and ultimately activates the MAP3K Ste11 through the integral transmembrane

protein Opy2, the GTPase Cdc42 and the MAP4K Ste20. These two osmosensing branches converge at the MAP2K (Pbs2) that acts as a scaffold protein for phosphorylation of the MAPK Hog1 [19].

The filamentous/invasive growth pathway leads to the activation of Kss1 (an ortholog of mammalian Erk2) under nutrient limiting conditions and, to a much lesser extent, to pheromone stimulation. Remarkably, it relies on proteins involved in the HOG pathway and the pheromone pathway (Mep2, Gpr1, Msb2, Sho1, Ste20, Ste11, and Ste7). In this case, specific activation of Kss1 is achieved by the absence of the Ste5 scaffold that liberates Ste7 allowing its interaction with Kss1 [20].

Cell wall instability is sensed through the cell wall integrity pathway (CWI) (Mpk1 MAPK) and is detected by five mechanosensors (Wsc1–3, Mid2, and Mtl1) that interact with the guanine nucleotide exchange factor (GED) Rom2 to activate Rho1 GTPase leading to protein kinase C (Pkc1) phosphorylation. Yeast Pkc1 serves as a MAP4K that phosphorylates the MAP3K Bck1, which leads to activation of Mpk1 through activation of the redundant MAP2K (Mkk1/2). Despite the absence of a cell wall in higher eukaryotes, mammalian ERK5 has been characterized as a functional ortholog of the CWI pathway [21].

Finally, the meiosis-specific MAPK Smk1 controls the postmeiotic program in diploid cells subjected to nutrient starvation. Activation of Smk1 differs from activation of MAPKs in the classical three-tiered MAPK cascade in which a CDK-activating kinase (CAK1) phosphorylates Smk1 and induces its auto-phosphorylation [22].

2.4 Dynamics of signal transduction

According to the nature of its input signal, MAPK activation can range from minutes (transient) to hours (sustained). The dynamics of MAPK activation results from the interplay between the extracellular environment and a myriad of intracellular feedforward/feedback regulators that give rise to cell fate decisions during cancer progression or development. For example, pulses of or continuous high EGF administration induce transient ERK activation and cell proliferation in rat adrenal cells, whereas repeated pulses of low EGF induce ERK-mediated differentiation into sympathetic-like neurons [23]. Similarly, different dynamics of JNK can generate opposing behaviors as persistent JNK activation has been shown to trigger apoptosis while its transient activation promotes cell survival [24]. Despite different signaling dynamics can determine cell fate, the underlying molecular mechanisms are not well understood.

2.5 Output responses to MAPK activation

The first response to extracellular insults is the immediate arrest of cell growth and hence a blockage of or a delay in cell cycle progression. Once activated, the different MAPKs phosphorylate a large number of substrates that distribute over many cellular compartments. In general terms, the ERK1/2 pathway is mainly associated with the promotion of growth in most cell types and is often linked with differentiation processes, although it can occasionally suppress cell survival [25]. Similarly, ERK5 also promotes proliferation during normal cell growth and differentiation [3]. On the other hand, JNK and p38 pathways have a well-established role in apoptosis, although they have also been shown to contribute to survival, immunity, development, and differentiation [4, 26–29]. One of the main mechanisms by which MAPKs modulate the abovementioned cellular processes is by controlling gene expression, mainly through regulation of the transcriptional machinery, chromatin structure, and post-transcriptional modifications.

3. Nuclear localization and function of MAPKs

In resting cells, MAPK components are usually located in the cytoplasm through their interaction with different anchor proteins, scaffolds, or phosphatases. Upon stimulation, MAPK signaling cascades rapidly transmit information into the nucleus to ensure the appropriate transcriptional response (**Figure 2A**). Across eukaryotes, this process is often initiated by transient accumulation of the MAPKs within the nucleus. The duration and the type of stimuli affect the nuclear localization of MAPK signaling proteins and play an important role in determination of the transcriptional output. Translocation of MAPK molecules requires specialized transport elements to travel through the nuclear pore complex (NPC).

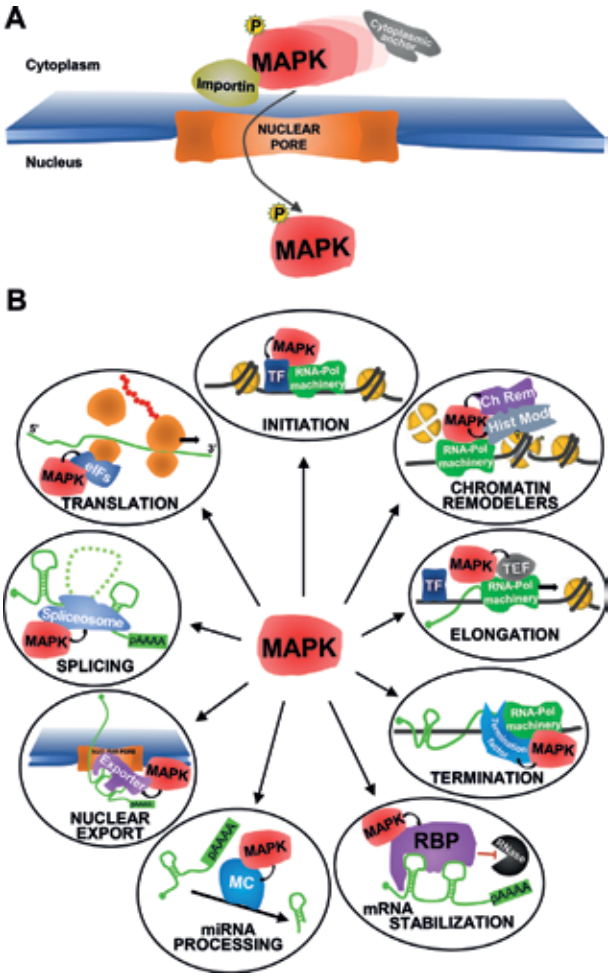

Figure 2.
*MAPK regulatory roles on gene expression: From transcription initiation to translation. (***A***) Activated MAPK is released from its cytoplasmic anchor and translocated to the nucleus. (***B***) From top and clockwise, MAPK regulation on different targets is represented by a black arrow; MAPKs are known to activate transcription factors (TFs) through phosphorylation and to recruit PolII to initiate transcription. Moreover, MAPKs also target several chromatin remodelers (Ch Rem) and histone modifiers (Hist mod) to regulate chromatin structure and histone eviction. MAPK interacts with transcription elongating (TEF) and termination factors to enhance transcription rate. mRNA is shown as a green line with a 5′ cap (green dot) and the polyA at the 3′ end. MAPKs also regulate several stabilizing RNA-binding proteins (RBPs), target miRNA processing through the microprocessor complex (MC), nuclear exporters, and splicing factors. Finally, MAPKs are also known to regulate translation elongation initiation factors (eIFs) to stimulate rapid mRNA translation. Overall, all these mechanisms aim to promote a rapid and efficient response for maximal cell adaptation.*

Nucleoplasmic shuttling of active MAPK can be mediated mainly by three strategies: (1) active regulation of import-export through the NPC; (2) escape from cytoplasmic anchors and/or sequestration by nuclear components; and (3) passive diffusion.

Canonical nuclear localization is an active process during which nuclear α/β importin complexes deliver cargo containing mono- or bi-partite stretches of basic residues (nuclear localization signals—NLSs) to the nucleus. Once in the nucleus, targeted proteins dissociate from importins by interacting with RanGTP. For example, in its inactive state, a nuclear export signal (NES) is exposed in ERK2, which confines it to the cytoplasm. Upon activation of ERK2, a conformational change disrupts its N- and C-terminal interactions, thereby exposing a NLS that sends the kinase into the nucleus. Similarly, activation of ERK1/2 allows their interaction with importin-7 and their nuclear accumulation [30]. Not only ERK1/2 nuclear accumulation is mediated through Ran as direct interaction, but also phosphorylation of nucleoporins (NUP50) facilitates translocation through importin-β [31]. The mechanisms by which p38 and JNK translocate into the nucleus are far less well understood. Recently, the motifs for interaction of both p38 and JNK with importins 3, 7, and 9 have been mapped at their N-terminal region. Ablated interaction of p38/JNK with their importins selectively impairs their nuclear accumulation and phosphorylation of their nuclear but not their cytosolic targets [32]. Nuclear translocation of the budding yeast Hog1 requires both Ran (GSP1) and importin-β (NMD5). Phosphorylation of Hog1 by its MAP2K Pbs2 is essential for its translocation, while MAPK activity is dispensable for its import. Similar to mammalian MAPKs, transcription factors such as Msn2/4, Hot1, Sko1, and the nuclear phosphatase Ptp2 contribute to the nuclear retention of Hog1. Dephosphorylated Hog1 is exported out of the nucleus through an importin-β homolog, XPO1. Blocking its nuclear export traps Hog1 in the nucleus but does not prevent its dephosphorylation [33].

An increasing body of knowledge supports the presence of other upstream kinases: MEK1/2, MEK5, MEKK2/3, and MKK6 in the nucleus [3, 6, 34, 35]. The role of upstream signaling components in transcriptional regulation has not received much attention and requires deeper understanding.

4. MAPK-regulated gene expression

Nuclear localized MAPKs have the capacity to rewire the transcriptional architecture by controlling several layers of mRNA biogenesis (**Figure 2B**). Nuclear localized MAPKs are competent to govern the transcription cycle by acting on several layers of the process. Temporal integration of MAPK signaling into transcription is generally mediated by the phosphorylation of hundreds of transcription-related targets. How this transcriptional control is achieved will be discussed in this section.

4.1 Genes regulated by MAPK activation: global induction/repression patterns

MAPK activation overrides the homeostatic transcriptional program by transiently governing the simultaneous upregulation and downregulation of gene expression. Activation of different MAPK cascades leads to a pathway-specific transcriptional landscape. This stimuli-specific response is required to redefine the demands of each condition and involves the regulation of all RNA species. Unbiased approaches such as tiling arrays and RNA-seq have further extended the type of MAPK-regulated transcripts to noncoding RNA (ncRNA), long noncoding RNA (lncRNA), and, specifically in higher eukaryotes, the expression of miRNA.

The MAPK-induced transcriptional response encompasses not only stimuli-specific genes but also a set of well-defined genes that respond to multiple signals, providing coping mechanisms for adaptation. The transcriptional program induced by MAPK activation is classically described in two stages: A primary response is independent of protein synthesis and triggers the expression of immediate and delayed early genes (IEG and DEG, respectively). Then, the secondary response follows in a protein synthesis-dependent manner to induce the expression of secondary response genes [36]. Here, we will focus our attention on the mechanisms that promote gene induction.

During this early or primary response, cells have to be able to repress cell cycle and growth genes while upregulating several transcription factor genes, which, once translated, will amplify the signal to generate a secondary or late response [37]. Thus, while a selected group of genes are upregulated, the rest of the transcriptome is transiently downregulated. Understanding of the mechanisms of MAPK-mediated gene repression has lagged behind when compared to the activating mechanisms, but some well understood prominent targets of repression are cell cycle- and growth-related genes (cyclins, tRNAs, and rRNAs).

4.2 MAPK as components of the transcriptional machinery

MAPKs localize and interact with all of the regulatory regions of their target genes to control gene expression through similar principles but through distinct molecular mechanisms. These mechanisms include the coordinated control of transcription initiation, elongation, and termination together with modulation of chromatin architecture to ensure proper transcription through its target genes. MAPK phosphorylation of chromatin-related factors alters their activity by regulating their nuclear localization, protein stability, or affinity to DNA [38].

4.2.1 Transcription initiation: transcription factor modification

Transcription initiation is the first step in governing gene expression and can be either directly or indirectly regulated by MAPKs. The most common regulatory mechanism involves the control of promoters by the regulation of an intricate network of transcription factors usually through direct phosphorylation and/or by induction of their expression [39]. Transcription factors serve as anchoring platforms for the recruitment of MAPKs to chromatin. Chromatin-tethered MAPK nucleates the key signaling components to promoters and other regulatory elements to form a competent pre-initiation complex (PIC). Examples of "hubs" in the transcription factor network that facilitate the recruitment of active MAPK to chromatin are Elk-1, c-Jun and c-Fos for p38, ERKs, and JNKs. ERK5 is a rare MAPK that contains a transcriptional coactivator domain and has the capability of stimulating transcription through transcription factors or by direct binding to DNA through its noncatalytic region [3].

One of the best characterized transcription factors is c-Jun upon which stimulation is phosphorylated by JNK in its transactivator domain, which is required for induction of its maximal transcriptional activity and increased protein stability [39]. A single-transcription factor can serve to integrate signals from different MAPKs, or several MAPKs can cooperate in regulation of the same target. In response to UV light, both p38 and ERK contribute to the activation of c-Fos. On the other hand, efficient Elk1 phosphorylation is achieved by its differential interaction with ERK1/2, p38, and JNK. Activated Elk1 induces the expression of c-Fos and c-Jun transcription factors that will subsequently regulate a second transcriptional wave that includes other transcription factors and phosphatases [38]. Alternatively,

a more indirect method to promote transcription is to activate downstream kinases that will themselves activate other transcription factors. For example, p38 activates two downstream kinases, mitogen- and stress-activated kinase 1/2 (MSK1–2), that activate another set of transcription factors STAT1/3, CREB, ATF1, and NF-κB [40].

In yeast, Fus3 and Kss1 MAPKs activate the transcription factor Ste12 that induces the expression of over 200 genes, including its own gene [41]. For example, in yeast, the combination of deletions of transcription factors and genome-wide analyses has been especially useful in providing a detailed view of the circuitry activated by the Hog1 or Fus3 MAPKs [42, 43].

The interrelationship between transcription factors and MAPKs is conserved throughout evolution, although the number of players and their functions has increased over time. MEF2 family transcription factors are substrates for several ERKs and in particular for p38 [44]. In yeast, it has been widely reported that the different transcription factors relevant for osmoresponsive gene expression are phosphorylated and recruited to target genes in a Hog1-dependent manner [45, 46]. Targeted recruitment of the MAPK activation machinery can also include recruitment of upstream MAPK-regulatory kinases. Examples of such in mammals are the recruitment of MEK1/2 to ERK-dependent genes [47] and the recruitment of MKK6 to p38 targeted regions in a MAPK-dependent manner [35]. Yeast upstream MAPK components have received far less attention than those of mammals, although Ste5 also associates with chromatin upon pheromone stimulation [48].

Besides controlling transcription factors, MAPKs control several other enzymatic activities, protein complexes, and targets that contribute to the formation of a transcriptionally competent Pre-Initiation Complex (PIC) (SAGA, Mediator, Ubp3) [49, 50]. A critical downstream node for MEK1/2 and ERK1/2 signaling upon the induction of EGF responsive genes is the integrator complex, a transcriptional coactivator. The binding of integrator to chromatin depends on catalytically active ERK1/2. Indeed, inhibition of the MAPK resulted in diminished association of integrator and RNA Pol II to chromatin [51].

4.2.2 Transcription elongation

Our knowledge of MAPK-regulated transcriptional control extends far beyond its control of transcription initiation and mainly originates from analysis of yeast MAPKs. The detection of MAPKs at the coding regions of their target genes suggested a far more extensive role for MAPKs as crucial components of the transcription regulatory complex. Seminal work regarding this phenomenon has been done in *S. cerevisiae* in which the association of Hog1, Fus3, and Mpk1 MAPKs with the coding regions of their target genes has been reported. Mpk1 elicits elongation of stress-responsive genes in a catalytic-independent manner by its interaction with the Paf1 elongation complex. Mpk1 is tethered to its target genes through binding to Paf1 that serves as a scaffold to escort Mpk1 into the elongating RNA Pol II. This binding requires the presence of the cell cycle transcriptional regulator SBF. The loss of this interaction restricts Mpk1 to the promoter region, which impairs both transcription and cell viability upon stress [52]. In response to osmotic stress, Hog1 and Paf1 interact through an unknown region, but the function and outcome of the Paf1 complex are kinase-specific.

The majority of genes targeted by Hog1 display an enrichment of the MAPK throughout the coding region [53, 54] that is mediated by the 3'UTR and is independent of promoter association. ORF-bound Hog1 behaves as a selective elongation factor by traveling and interacting with phosphorylated RNA Pol II (Rpb1). As RNA Pol II moves across the gene, it regulates chromatin structure through the recruitment of chromatin remodelers and chromatin-modifying enzymes (Section 4.3).

Moreover, Hog1 phosphorylates the Spt4 elongation factor to regulate RNA Pol II processivity to stimulate elongation efficiency at stress-responsive genes [55]. As happens during initiation, Hog1 recruits other protein complexes with specific enzymatic activities such as deubiquitinase (Ubp3) to ensure the proper production of stress-responsive genes [50]. Further studies in mammalian cells also corroborated p38 binding to coding regions of genes not only in response to osmotic stress but also during skeletal muscle differentiation, suggesting that the mechanism and purposes of Hog1/p38 transcriptional regulation are conserved throughout evolution.

Transcription elongation rates for many genes depend on the entangled interplay of factors and complexes that regulate RNA Pol II. During elongation, a number of positive and negative elongation factors (P-TEFs and N-TEFs, respectively) have been shown to accelerate or attenuate Pol II, and, not surprisingly, these factors are targeted by MAPKs at stress-responsive genes. In response to hormone stimulation, MEK1 and ERK1/2 promote elongation and abolish pausing of RNA Pol II [56].

4.2.3 Termination

Unlike initiation and elongation, transcription termination can be carried out through different pathways depending on the coding or noncoding nature of the transcript. The two best defined termination pathways that are also highly conserved are the polyA-dependent pathway for protein coding and the Sen1-dependent pathway for noncoding transcripts.

One of the best studied examples of the involvement of MAPKs in the control of transcription termination is that of the role of Mpk1 in transcription termination during heat stress in yeast. As mentioned before, Paf1 and Mpk1 interact at heat responsive genes; this association prevents the recruitment of the Sen1-Nrd1-Nab3 termination machinery (NNS). Interestingly, the same study showed that human ERK5 and human Paf1 complex expressed in yeast also regulated termination in response to cell wall stress [52]. Mpk1 has recently been shown to directly phosphorylate Tyr1 in the RNA Pol II CTD as it traverses the coding region with the elongating machinery. This phosphorylation occurs in a stress-dependent manner and prevents early termination through the NNS pathway [57]. Deep sequencing of osmotically stressed neuronal cell lines identified a new set of transcripts termed downstream of gene-containing transcripts (DoGs). These noncoding transcripts span large region downstream of annotated gene features (>45 Kb) and are actively regulated through IP3 signaling [58].

4.3 MAP kinases and their effects on chromatin

MAPKs facilitate the abovementioned transcription activity by also regulating several chromatin remodelers to generate the proper chromatin environment for the transcription machinery. For induction of gene expression, chromatin must be accessible to allow the assembly of transcription factors, RNA Pol II and other factors, during initiation, elongation, and termination. These chromatin remodelers have been studied in both yeast and mammalian models as has been extensively reviewed in [38].

There are numerous examples of MAPKs interacting with chromatin remodelers. For instance, both Hog1 and p38 govern the recruitment of the remodeling complex SWI/SNF to target genes [38]. On the other hand, MAPK regulation goes beyond the substrate phosphorylation. As described in previous sections, MAPKs can also regulate chromatin remodeling through direct protein-protein interactions. This is the case with ERK2, which contacts PolyADP-ribose polymerase (PARP1), thereby

increasing its activating activity on chromatin remodelers [59]. Apart from chromatin remodelers, MAPKs govern a cohort of histone modifiers that not only destabilize nucleosomes but also, in a more complex manner, generate selective marks that dictate nucleosome dynamics. An example of this type of regulation is the Hog1-dependent gene recruitment of Rpd3, a histone deacetylase, that induces gene expression by promoting the eviction of histones at osmoresponsive genes [60] and the regulation of H3K4 monomethylation to dictate specificity of chromatin remodelers [61]. During elongation, as Hog1 travels with the elongating RNA polymerase, it recruits the RSC remodeling complex, thereby facilitating transcription along the gene body [62]. In mammals, ERKs, p38, and JNK promote the phosphorylation of H3S10 either directly or through their downstream kinases [38, 63]. p38 also phosphorylates the transcription factor MEF2D, which, in turn, leads to recruitment of the Ash2L-containing methyltransferase complex that generates an increase in the activating mark H3K4me3 [64]. These examples highlight the relevance of MAPK-mediated histone modification to generate an efficient chromatin remodeling robustly achieved through different mechanisms.

MAPKs also regulate gene silencing through chromatin remodelers. ERK1/2 directly interacts with the histone deacetylase 4 (HDAC4) that removes acetyl groups leading to chromatin condensation [6]. Similarly, Hog1 promotes the transcription of *PNC1*, which encodes an activator of Sir2, a histone acetyltransferase that protects sensible rRNA-coding regions from DNA damage [65]. In these two cases, MAPKs act as repressing elements of chromatin remodeling.

5. Role of MAPK signaling in post-transcriptional regulation

The ultimate goal of MAPK-mediated transcriptional reprogramming is to change the proteome composition. This change becomes especially important upon extracellular challenge when a massive pool of previously low-abundant RNAs needs to be expressed. Activated MAPKs target mRNA-binding proteins to downregulate unnecessary mRNAs and favor expression of the required genes in order to adapt to the new conditions [31, 45].

5.1 Transcript/RNA stability

Genomic run-on (GRO) experiments that have revealed global changes of gene expression in response to stress are also achieved through the regulation of mRNA stability and decay [66]. In yeast, upon osmotic stress, there is a broad mRNA destabilization, while Hog1 plays a role on the stabilization of osmo-induced mRNAs [67]. The p38 MAPK pathway is also a key regulator of the mRNA stability of both TTP (tristetraprolin), a protein that shortens the half-lives of adenine-uracil rich element (ARE)-containing mRNA, and HuR (human antigen R), a protein that stabilizes such mRNA. The role of p38 turns out to be opposed depending on the cell type [68]. Like p38, ERK and possibly JNK are thought to target HuR, changing its localization to the cytosol, where it stabilizes ARE-containing mRNA [69]. As a further example of the role of p38 in regulating mRNA stability, p38-mediated phosphorylation of ADAR1p110, another mRNA-binding protein, suppresses apoptosis in stressed cells by protecting many antiapoptotic gene transcripts from mRNA decay [70].

Another layer of transcriptional regulation coordinated by MAPKs, which has gained importance over the years, is the regulation of miRNA biogenesis. A relevant example of the coordination of the regulation of miRNA by different MAPK cascades is the regulation that takes place in the early stages of the inflammatory response. JNK and p38 trigger transcription of the miRNA let-7f, which downregulates the

expression of Blimp1 and PRDM1, two transcriptional repressors of inflammatory genes. Since a sustained expression of inflammatory genes is detrimental, later activation of ERK promotes the transcription of Lin28, an inhibitor of let-7f biogenesis, thereby increasing the expression of the Blimp1 and PRDM1 repressors [71]. Here, the same stimuli generate a time-dependent regulated activation of multiple signaling pathways to achieve a finely tuned transcriptional response.

5.2 mRNA export

There are numerous examples of interactions between MAPK pathways and different components of the mRNA exporting machinery in biological systems ranging from yeast to mammalian cells. In yeast, it has been reported that, in response to osmotic or heat stress, Hog1 and Mpk1, respectively, phosphorylate components of the nuclear pore complex to increase the export efficiency of stress-responsive mRNAs [72, 73]. Similarly, in mammals, both p38 and ERK pathways regulate RNA-binding proteins such as eIF4E or hDl1 that facilitate mRNA export [74, 75]. In the event of stress, the export of the newly transcribed mRNAs is prioritized to maximize the transcriptional response.

5.3 mRNA splicing

The transcriptional response to external stimuli generates an outburst of mRNAs that have to be spliced. The associations between splicing events that modulate MAPK genes are becoming increasingly relevant in human disease [76]. One strategy to regulate splicing under stress is phosphorylation of the splicing factor TDP-43 by MEK1/2, which prevents TDP-43 aggregation [77]. Another mechanism of regulating splicing is by interfering with the localization of splicing factors such as RNM4, hn-RNPA1, or hSlut7 [78–80].

5.4 Translation

Translation plays a pivotal role in the control of gene expression and is tightly regulated by MAPK pathways that modulate the activity of several components within the translational machinery [81]. In yeast, Hog1 promotes Rck2-mediated attenuation of protein synthesis in response to osmotic stress by phosphorylation of the translation elongation factor 2 (EF-2) [82]. ERK- and p38-activated MNKs phosphorylate the elongation initiation factor eIF4E to enhance translation initiation [83]. Another example is the activation of RSK, a downstream kinase of ERK. RSK phosphorylates S6, a component of the 40S ribosomal subunit, as well as the elongation initiation factor eIF4B, which facilitates their binding to eIF3 to promote mRNA translation [84, 85]. Besides the targeting of newly transcribed mRNAs, translation regulation can also target mRNAs that have not been transcriptionally induced, a type of regulation found in yeast and mammals [86].

6. Future perspectives and challenges

Due to their master regulatory role, MAPKs have sparked a lot of interest and have been the main focus of multiple researchers worldwide over the last 30 years. As MAPK knowledge advances, it has become obvious that the external control of MAPK activity has the potential to modulate cellular behavior and survival. Moreover, MAPK signaling has been found to be altered or defective in many human diseases such as cancer; therefore, achievement of the control of

MAPK activity could provide an attractive intervention point for new therapeutic approaches. However, despite the tremendous amount of knowledge generated, there are fundamental questions that remain to be elucidated in order to transform the biomedical potential of MAPKs into a reality.

While the central core of MAPK signaling cascades has been extensively described, branches of the networks have not yet been completely identified. This is especially true in terms of the upstream sensors, where the picture is not well defined, especially in higher eukaryotes. In the immediate future, we therefore foresee that more sophisticated approaches using CRISPR/Cas9 and RNAi-based libraries will provide a means to perform systematic genome-wide genetic screens to reveal missing components of these pathways.

Additionally, there are other features of MAPK signaling that might have been overlooked. An example of such a feature is the peptide-mediated blockage of p38/JNK interaction with importins, which reduces their nuclear export. The presence of this peptide impaired MAPK nuclear localization and decreased cell proliferation and tumor growth to a larger extent than the presence of commercial p38 inhibitors. These data open up a new perspective on MAPK regulation and need to be further examined as they could provide a new therapeutic intervention strategy to regulate MAPK activity [32]. It is clear that nuclear localization stimulates the encounter of MAPKs with defined chromatin loci, where their targets are located to provide specificity for gene induction; however, how the kinases are directed to these regions is not clear. Similarly, the molecular mechanisms of transcriptional termination have only recently been uncovered, and there are few reports regarding the targets and the control of MAPKs in termination, as many noncoding RNAs have been shown to be regulated by MAPKs such as Hog1/p38 and ERK2 [87, 88].

Remarkably, upstream MAPK pathway components as transcriptional regulators are also unclear, although the recruitment of several MAP2Ks (MEK1/2, MKK6) to chromatin has been detected. Furthermore, an extreme case has been reported in which, upon insulin stimulation, the entire signaling pathway from the insulin receptor to the ERK signaling cascades is recruited to insulin inducible loci [89]. The functions and consequences of such recruitment require further investigation. Due to their master regulatory role, MAPKs have generated a lot of interest. It is clear that controlling MAPK activity could provide a means of controlling cell behavior. Additionally, understanding the consequences of heterogeneity for MAPK-regulated events will be crucial for understanding differential responses to extracellular stimuli and therapeutic treatments. In conclusion, it is of utmost importance that MAPK-mediated mechanisms of controlling gene expression are fully characterized in order to further identify druggable targets/processes that are relevant to human diseases.

Acknowledgements

MNR is a recipient of a Maria de Maeztu Postdoctoral Fellowship (Doctores Banco de Santander-María de Maeztu at Universitat Pompeu Fabra). GMC is a recipient of an FPI Predoctoral Fellowship (Spanish Ministry of Economyv Industry and Competitiveness (MINECO). The FP and EdN's laboratory is supported by grants from the Spanish Ministry of Economy and Competitiveness (BFU2015-64437-P to FP and BFU2017-85152-P and FEDER to EN), the Catalan Government (2017 SGR 799), the Fundación Botín, and the Banco Santander through its Santander Universities Global Division to FP and the Unidad de Excelencia Maria de Maeztu, MDM-2014-0370. FP is a recipient of an ICREA Acadèmia (Generalitat de Catalunya). We gratefully acknowledge funding from the Spanish Ministry

of Economy, Industry, and Competitiveness (MINECO) through the Centers of Excellence Severo Ochoa award and from the CERCA Program of the Catalan Government.

Conflict of interest

The authors declare no conflict of interest.

Author details

Mariona Nadal-Ribelles[1,2†], Carme Solé[1,2†], Gerard Martínez-Cebrián[1,2†], Francesc Posas[1,2*] and Eulàlia de Nadal[1,2*]

1 Departament de Ciències Experimentals i de la Salut, Cell Signaling Research Group, Universitat Pompeu Fabra (UPF), Barcelona, Spain

2 Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, Barcelona, Spain

*Address all correspondence to: eulalia.nadal@upf.edu and francesc.posas@upf.edu

† The authors contributed equally.

References

[1] Seger R, Wexler S. The MAPK signaling cascades. In: Bradshaw RA, Stahl PD, editors. Encyclopedia of Cell Biology. Academic Press. Elsevier Inc; 2015. pp. 122-127. DOI: 10.1016/B978-0-12-394447-4.30014-1

[2] Cargnello M, Roux PP. Activation and function of the MAPKs and their substrates, the MAPK-activated protein kinases. Microbiology and Molecular Biology Reviews. 2011;**75**:50-83. DOI: 10.1128/MMBR.00031-10

[3] Drew BA, Burow ME, Beckman BS. MEK5/ERK5 pathway: The first fifteen years. Biochimica et Biophysica Acta—Reviews on Cancer. 2012;**1825**:37-48. DOI: 10.1016/j.bbcan.2011.10.002

[4] Zeke A, Misheva M, Reményi A, Bogoyevitch MA. JNK signaling: Regulation and functions based on complex protein-protein partnerships. Microbiology and Molecular Biology Reviews. 2016;**80**:793-835. DOI: 10.1128/MMBR.00043-14

[5] Roskoski R. ERK1/2 MAP kinases: Structure, function, and regulation. Pharmacological Research. 2012;**66**:105-143. DOI: 10.1016/J.PHRS.2012.04.005

[6] Plotnikov A, Zehorai E, Procaccia S, Seger R. The MAPK cascades: Signaling components, nuclear roles and mechanisms of nuclear translocation. Biochimica et Biophysica Acta—Molecular Cell Research. 2011;**1813**:1619-1633. DOI: 10.1016/j.bbamcr.2010.12.012

[7] Bi W, Xiao L, Jia Y, Wu J, Xie Q, Ren J, et al. c-Jun N-terminal kinase enhances MST1-mediated pro-apoptotic signaling through phosphorylation at serine 82. Journal of Biological Chemistry. 2010;**285**:6259-6264. DOI: 10.1074/jbc.M109.038570

[8] Zeke A, Bastys T, Alexa A, Garai Á, Mészáros B, Kirsch K, et al. Systematic discovery of linear binding motifs targeting an ancient protein interaction surface on MAP kinases. Molecular Systems Biology. 2015;**11**:837. DOI: 10.15252/MSB.20156269

[9] Liu X, Zhang C-S, Lu C, Lin S-C, Wu J-W, Wang Z-X. A conserved motif in JNK/p38-specific MAPK phosphatases as a determinant for JNK1 recognition and inactivation. Nature Communications. 2016;**7**:10879. DOI: 10.1038/ncomms10879

[10] Chol KY, Satterberg B, Lyons DM, Elion EA. Ste5 tethers multiple protein kinases in the MAP kinase cascade required for mating in *S. cerevisiae*. Cell. 1994;**78**:499-512. DOI: 10.1016/0092-8674(94)90427-8

[11] Posas F, Saito H. Osmotic activation of the HOG MAPK pathway via Ste11p MAPKKK: Scaffold role of Pbs2p MAPKK. Science. 1997;**276**:1702-1708. DOI: 10.1126/science.276.5319.1702

[12] Kidger AM, Keyse SM. The regulation of oncogenic Ras/ERK signalling by dual-specificity mitogen activated protein kinase phosphatases (MKPs). Seminars in Cell & Developmental Biology. 2016;**50**:125-132. DOI: 10.1016/J.SEMCDB.2016.01.009

[13] Macia J, Regot S, Peeters T, Conde N, Solé R, Posas F. Dynamic signaling in the Hog1 MAPK pathway relies on high basal signal transduction. Science Signaling. 2009;**2**:ra13. DOI: 10.1126/scisignal.2000056

[14] Meister M, Tomasovic A, Banning A, Tikkanen R. Mitogen-activated protein (MAP) kinase scaffolding proteins: A recount. International Journal of

Molecular Sciences. 2013;**14**:4854-4884. DOI: 10.3390/ijms14034854

[15] Whelan JT, Hollis SE, Cha DS, Asch AS, Lee MH. Post-transcriptional regulation of the Ras-ERK/MAPK signaling pathway. Journal of Cellular Physiology. 2012;**227**:1235-1241. DOI: 10.1002/jcp.22899

[16] Chen RE, Thorner J. Function and regulation in MAPK signaling pathways: Lessons learned from the yeast *Saccharomyces cerevisiae*. Biochimica et Biophysica Acta—Molecular Cell Research. 2007;**1773**:1311-1340. DOI: 10.1016/j.bbamcr.2007.05.003

[17] Engelberg D, Perlman R, Levitzki A. Transmembrane signaling in *Saccharomyces cerevisiae* as a model for signaling in metazoans: State of the art after 25 years. Cellular Signalling. 2014;**26**:2865-2878. DOI: 10.1016/j.cellsig.2014.09.003

[18] Merlini L, Dudin O, Martin SG. Mate and fuse: How yeast cells do it. Open Biology. 2013;**3**:130008-130008. DOI: 10.1098/rsob.130008

[19] Saito H, Posas F. Response to hyperosmotic stress. Genetics. 2012;**192**:289-318. DOI: 10.1534/genetics.112.140863

[20] Cullen PJ, Sprague GF. The regulation of filamentous growth in yeast. Genetics. 2012;**190**:23-49. DOI: 10.1534/genetics.111.127456

[21] Sanz A, García R, Rodríguez-Peña J, Arroyo J. The CWI pathway: Regulation of the transcriptional adaptive response to cell wall stress in yeast. Journal of Fungi. 2017;**4**:1-12. DOI: 10.3390/jof4010001

[22] Piekarska I, Rytka J, Rempola B. Regulation of Sporulation in the Yeast Saccharomyces cerevisiae. Acta Biochimica Polonica. 2010;**57**(3):241-250

[23] Ryu H, Chung M, Dobrzyński M, Fey D, Blum Y, Lee SS, et al. Frequency modulation of ERK activation dynamics rewires cell fate. Molecular Systems Biology. 2015;**11**:838. DOI: 10.15252/MSB.20156458

[24] Ventura J-J, Hübner A, Zhang C, Flavell RA, Shokat KM, Davis RJ. Chemical genetic analysis of the time course of signal transduction by JNK. Molecular Cell. 2006;**21**:701-710. DOI: 10.1016/J.MOLCEL.2006.01.018

[25] Yoon S, Seger R. The extracellular signal-regulated kinase: Multiple substrates regulate diverse cellular functions. Growth Factors. 2006;**24**:21-44. DOI: 10.1080/02699050500284218

[26] Duch A, de Nadal E, Posas F. The p38 and Hog1 SAPKs control cell cycle progression in response to environmental stresses. FEBS Letters. 2012;**586**:2925-2931. DOI: 10.1016/j.febslet.2012.07.034

[27] Joaquin M, Gubern A, González-Nuñez D, Josué Ruiz E, Ferreiro I, de Nadal E, et al. The p57 CDKi integrates stress signals into cell-cycle progression to promote cell survival upon stress. The EMBO Journal. 2012;**31**:2952-2964. DOI: 10.1038/emboj.2012.122

[28] Gubern A, Joaquin M, Marquès M, Maseres P, Garcia-Garcia J, Amat R, et al. The N-terminal phosphorylation of RB by p38 bypasses its inactivation by CDKs and prevents proliferation in cancer cells. Molecular Cell. 2016;**64**:25-36. DOI: 10.1016/j.molcel.2016.08.015

[29] Trempolec N, Dave-Coll N, Nebreda AR. SnapShot: p38 MAPK substrates. Cell. 2013;**152**:924-924.e1. DOI: 10.1016/J.CELL.2013.01.047

[30] Casar B, Pinto A, Crespo P. Essential role of ERK dimers in the activation of cytoplasmic but not nuclear substrates by ERK-scaffold complexes. Molecular

Cell. 2008;**31**:708-721. DOI: 10.1016/j.molcel.2008.07.024

[31] Kosako H, Yamaguchi N, Aranami C, Ushiyama M, Kose S, Imamoto N, et al. Phosphoproteomics reveals new ERK MAP kinase targets and links ERK to nucleoporin-mediated nuclear transport. Nature Structural & Molecular Biology. 2009;**16**:1026-1035. DOI: 10.1038/nsmb.1656

[32] Maik-Rachline G, Zehorai E, Hanoch T, Blenis J, Seger R. The nuclear translocation of the kinases p38 and JNK promotes inflammation-induced cancer. Science Signaling. 2018;**11**:eaao3428. DOI: 10.1126/scisignal.aao3428

[33] Ferrigno P, Posas F, Koepp D, Saito H, Silver PA. Regulated nucleo/cytoplasmic exchange of HOG1 MAPK requires the importin β homologs NMD5 and XPO1. EMBO Journal. 1998;**17**:5606-5614. DOI: 10.1093/emboj/17.19.5606

[34] Perry RL, Parker MH, Rudnicki MA. Activated MEK1 binds the nuclear MyoD transcriptional complex to repress transactivation. Molecular Cell. 2001;**8**:291-301. DOI: 10.1016/S1097-2765(01)00302-1

[35] Ferreiro I, Barragan M, Gubern A, Ballestar E, Joaquin M, Posas F. The p38 SAPK is recruited to chromatin via its interaction with transcription factors. Journal of Biological Chemistry. 2010;**285**:31819-31828. DOI: 10.1074/jbc.M110.155846

[36] Tullai JW, Schaffer ME, Mullenbrock S, Sholder G, Kasif S, Cooper GM. Immediate-early and delayed primary response genes are distinct in function and genomic architecture. Journal of Biological Chemistry. 2007;**282**:23981-23995. DOI: 10.1074/jbc.M702044200

[37] Ferreiro I, Joaquin M, Islam A, Gomez-Lopez G, Barragan M, Lombardía L, et al. Whole genome analysis of p38 SAPK-mediated gene expression upon stress. BMC Genomics. 2010;**11**:144. DOI: 10.1186/1471-2164-11-144

[38] Yang SH, Sharrocks AD. Whitmarsh AJ. MAP kinase signalling cascades and transcriptional regulation. Gene. 2013;**513**:1-13. DOI: 10.1016/j.gene.2012.10.033

[39] Turjanski AG, Vaqué JP, Gutkind JS. MAP kinases and the control of nuclear events. Oncogene. 2007;**26**:3240-3253. DOI: 10.1038/sj.onc.1210415

[40] Arthur JSC. MSK activation and physiological roles. Frontiers in Bioscience : A Journal and Virtual Library. 2008;**13**:5866-5879. DOI: 10.2741/3122

[41] Roberts CJ, Nelson B, Marton MJ, Stoughton R, Meyer MR, Bennett HA, et al. Signaling and circuitry of multiple MAPK pathways revealed by a matrix of global gene expression profiles. Science. 2000;**287**:873-880. DOI: 10.1126/science.287.5454.873

[42] Capaldi AP, Kaplan T, Liu Y, Habib N, Regev A, Friedman N, et al. Structure and function of a transcriptional network activated by the MAPK Hog1. Nature Genetics. 2008;**40**:1300-1306. DOI: 10.1038/ng.235

[43] van Wageningen S, Kemmeren P, Lijnzaad P, Margaritis T, Benschop JJ, de Castro IJ, et al. Functional overlap and regulatory links shape genetic interactions between signaling pathways. Cell. 2010;**143**:991-1004. DOI: 10.1016/J.CELL.2010.11.021

[44] Pon JR, Marra MA. MEF2 transcription factors: Developmental regulators and emerging cancer genes. Oncotarget. 2016;**7**:2297-2312. DOI: 10.18632/oncotarget.6223

[45] de Nadal E, Ammerer G, Posas F. Controlling gene expression in response to stress. Nature Reviews Genetics. 2011;**12**:833-845. DOI: 10.1038/nrg3055

[46] de Nadal E, Posas F. Osmostress-induced gene expression—A model to understand how stress-activated protein kinases (SAPKs) regulate transcription. FEBS Journal. 2015;**282**:3275-3285. DOI: 10.1111/febs.13323

[47] Lawrence MC, Shao C, McGlynn K, Naziruddin B, Levy MF, Cobb MH. Multiple chromatin-bound protein kinases assemble factors that regulate insulin gene transcription. Proceedings of the National Academy of Sciences. 2009;**106**:22181-22186. DOI: 10.1073/pnas.0912596106

[48] Pokholok DK, Zeitlinger J, Hannett NM, Reynolds DB, Young RA. Activated signal transduction kinases frequently occupy target genes. Science. 2006;**313**:533-536. DOI: 10.1126/science.1127677

[49] Zapater M, Sohrmann M, Peter M, Posas F, de Nadal E. Selective requirement for SAGA in Hog1-mediated gene expression depending on the severity of the external osmostress conditions. Molecular and Cellular Biology. 2007;**27**:3900-3910. DOI: 10.1128/MCB.00089-07

[50] Solé C, Nadal-Ribelles M, Kraft C, Peter M, Posas F, de Nadal E. Control of Ubp3 ubiquitin protease activity by the Hog1 SAPK modulates transcription upon osmostress. The EMBO Journal. 2011;**30**:3274-3284. DOI: 10.1038/emboj.2011.227

[51] Yue J, Lai F, Beckedorff F, Zhang A, Pastori C, Shiekhattar R. Integrator orchestrates RAS/ERK1/2 signaling transcriptional programs. Genes & Development. 2017;**31**:1809-1820. DOI: 10.1101/gad.301697.117

[52] Kim K-Y, Levin DE. Mpk1 MAPK association with the Paf1 complex blocks Sen1-mediated premature transcription termination. Cell. 2011;**144**:745-756. DOI: 10.1016/j.cell.2011.01.034

[53] Nadal-Ribelles M, Conde N, Flores O, González-Vallinas J, Eyras E, Orozco M, et al. Hog1 bypasses stress-mediated down-regulation of transcription by RNA polymerase II redistribution and chromatin remodeling. Genome Biology. 2012;**13**:R106. DOI: 10.1186/gb-2012-13-11-r106

[54] Proft M, Mas G, de Nadal E, Vendrell A, Noriega N, Struhl K, et al. The stress-activated Hog1 kinase is a selective transcriptional elongation factor for genes responding to osmotic stress. Molecular Cell. 2006;**23**:241-250. DOI: 10.1016/j.molcel.2006.05.031

[55] Silva A, Cavero S, Begley V, Solé C, Böttcher R, Chávez S, et al. Regulation of transcription elongation in response to osmostress. PLOS Genetics. 2017;**13**:e1007090. DOI: 10.1371/journal.pgen.1007090

[56] Fujita T, Ryser S, Piuz I, Schlegel W. Up-regulation of P-TEFb by the MEK1-extracellular signal-regulated kinase signaling pathway contributes to stimulated transcription elongation of immediate early genes in neuroendocrine cells. Molecular and Cellular Biology. 2008;**28**:1630-1643. DOI: 10.1128/MCB.01767-07

[57] Yurko N, Liu X, Yamazaki T, Hoque M, Tian B, Manley JL. MPK1/SLT2 links multiple stress responses with gene expression in budding yeast by phosphorylating Tyr1 of the RNAP II CTD. Molecular Cell. 2017;**68**:913-925.e3. DOI: 10.1016/j.molcel.2017.11.020

[58] Vilborg A, Passarelli MC, Yario TA, Tycowski KT, Steitz JA. Widespread inducible transcription downstream

of human genes. Molecular Cell. 2015;**59**:449-461. DOI: 10.1016/j.molcel.2015.06.016

[59] Cohen-Armon M, Visochek L, Rozensal D, Kalal A, Geistrikh I, Klein R, et al. DNA-independent PARP-1 activation by phosphorylated ERK2 increases Elk1 activity: A link to histone acetylation. Molecular Cell. 2007;**25**:297-308. DOI: 10.1016/J.MOLCEL.2006.12.012

[60] de Nadal E, Zapater M, Alepuz PM, Sumoy L, Mas G, Posas F. The MAPK Hog1 recruits Rpd3 histone deacetylase to activate osmoresponsive genes. Nature. 2004;**427**:370-374. DOI: 10.1038/nature02258

[61] Nadal-Ribelles M, Mas G, Millán-Zambrano G, Solé C, Ammerer G, Chávez S, et al. H3K4 monomethylation dictates nucleosome dynamics and chromatin remodeling at stress-responsive genes. Nucleic Acids Research. 2015;**43**:4937-4949. DOI: 10.1093/nar/gkv220

[62] Mas G, de Nadal E, Dechant R, Rodríguez de la Concepción ML, Logie C, Jimeno-González S, et al. Recruitment of a chromatin remodelling complex by the Hog1 MAP kinase to stress genes. The EMBO Journal. 2009;**28**:326-336. DOI: 10.1038/emboj.2008.299

[63] Zippo A, Serafini R, Rocchigiani M, Pennacchini S, Krepelova A, Oliviero S. Histone crosstalk between H3S10ph and H4K16ac generates a histone code that mediates transcription elongation. Cell. 2009;**138**:1122-1136. DOI: 10.1016/j.cell.2009.07.031

[64] Rampalli S, Li L, Mak E, Ge K, Brand M, Tapscott SJ, et al. p38 MAPK signaling regulates recruitment of Ash2L-containing methyltransferase complexes to specific genes during differentiation. Nature Structural & Molecular Biology. 2007;**14**:1150-1156. DOI: 10.1038/nsmb1316

[65] Vendrell A, Martínez-Pastor M, González-Novo A, Pascual-Ahuir A, Sinclair DA, Proft M, et al. Sir2 histone deacetylase prevents programmed cell death caused by sustained activation of the Hog1 stress-activated protein kinase. EMBO Reports. 2011;**12**:1062-1068. DOI: 10.1038/embor.2011.154

[66] Fan J, Yang X, Wang W, Wood WH, Becker KG, Gorospe M. Global analysis of stress-regulated mRNA turnover by using cDNA arrays. Proceedings of the National Academy of Sciences. 2002;**99**:10611-10616. DOI: 10.1073/pnas.162212399

[67] Romero-Santacreu L, Moreno J, Pérez-Ortín JE, Alepuz P. Specific and global regulation of mRNA stability during osmotic stress in *Saccharomyces cerevisiae*. RNA Journal. 2009;**15**:1110-1120. DOI: 10.1261/rna.1435709

[68] Astakhova AA, Chistyakov DV, Sergeeva MG, Reiser G. Regulation of the ARE-binding proteins, TTP (tristetraprolin) and HuR (human antigen R), in inflammatory response in astrocytes. Neurochemistry International. 2018;**118**:82-90. DOI: 10.1016/j.neuint.2018.04.014

[69] Marchesi N, Thongon N, Pascale A, Provenzani A, Koskela A, Korhonen E, et al. Autophagy stimulus promotes early HuR protein activation and p62/SQSTM1 protein synthesis in ARPE-19 cells by triggering Erk1/2, p38MAPK, and JNK kinase pathways. Oxidative Medicine and Cellular Longevity. 2018;**2018**:1-15. DOI: 10.1155/2018/4956080

[70] Sakurai M, Shiromoto Y, Ota H, Song C, Kossenkov AV, Wickramasinghe J, et al. ADAR1 controls apoptosis of stressed cells by inhibiting Staufen1-mediated mRNA decay. Nature Structural and Molecular Biology.

2017;**24**:534-543. DOI: 10.1038/nsmb.3403

[71] Ayyar KK, Reddy KVR. MAPK and NF-κB signalling pathways regulate the expression of miRNA, let-7f in human endocervical epithelial cells. Journal of Cellular Biochemistry. 2018;**119**:4751-4759. DOI: 10.1002/jcb.26665

[72] Carmody SR, Tran EJ, Apponi LH, Corbett AH, Wente SR. The mitogen-activated protein kinase Slt2 regulates nuclear retention of non-heat shock mRNAs during heat shock-induced stress. Molecular and Cellular Biology. 2010;**30**:5168-5179. DOI: 10.1128/MCB.00735-10

[73] Regot S, de Nadal E, Rodrigúez-Navarro S, Gonzaĺez-Novo A, Peŕez-Fernandez J, Gadal O, et al. The Hog1 stress-activated protein kinase targets nucleoporins to control mrna export upon stress. Journal of Biological Chemistry. 2013;**288**:17384-17398. DOI: 10.1074/jbc.M112.444042

[74] Sabio G, Cerezo-Guisado MI, Del Reino P, Iñesta-Vaquera FA, Rousseau S, Arthur JSC, et al. Cuenda a. p38gamma regulates interaction of nuclear PSF and RNA with the tumour-suppressor hDlg in response to osmotic shock. Journal of Cell Science. 2010;**123**:2596-2604. DOI: 10.1242/jcs.066514

[75] Seidel P, Sun Q, Costa L, Lardinois D, Tamm M, Roth M. The MNK-1/eIF4E pathway as a new therapeutic pathway to target inflammation and remodelling in asthma. Cellular Signalling. 2016;**28**:1555-1562. DOI: 10.1016/j.cellsig.2016.07.004

[76] Chabot B, Shkreta L. Defective control of pre-messenger RNA splicing in human disease. Journal of Cell Biology. 2016;**212**:13-27. DOI: 10.1083/jcb.201510032

[77] Li W, Reeb AN, Lin B, Subramanian P, Fey EE, Knoverek CR, et al. Heat shock-induced phosphorylation of TAR DNA-binding protein 43 (TDP-43) by MAPK/ERK kinase regulates TDP-43 function. Journal of Biological Chemistry. 2017;**292**:5089-5100. DOI: 10.1074/jbc.M116.753913

[78] Lin J-C, Hsu M, Tarn W-Y. Cell stress modulates the function of splicing regulatory protein RBM4 in translation control. Proceedings of the National Academy of Sciences. 2007;**104**:2235-2240. DOI: 10.1073/pnas.0611015104

[79] Van Oordt WVDH, Diaz-Meco MT, Lozano J, Krainer AR, Moscat J, Cáceres JF. The MKK(3/6)-p38-signaling cascade alters the subcellular distribution of hnRNP A1 and modulates alternative splicing regulation. Journal of Cell Biology. 2000;**149**:307-316. DOI: 10.1083/jcb.149.2.307

[80] Shomron N. Stress alters the subcellular distribution of hSlu7 and thus modulates alternative splicing. Journal of Cell Science. 2005;**118**:1151-1159. DOI: 10.1242/jcs.01720

[81] Roux PP, Topisirovic I. Regulation of mRNA translation by signaling pathways. Cold Spring Harbor Perspectives in Biology. 2012;**4**:1-24. DOI: 10.1101/cshperspect.a012252

[82] Teige M, Scheikl E, Reiser V, Ruis H, Ammerer G. Rck2, a member of the calmodulin-protein kinase family, links protein synthesis to high osmolarity MAP kinase signaling in budding yeast. Proceedings of the National Academy of Sciences. 2001;**98**:5625-5630. DOI: 10.1073/pnas.091610798

[83] Waskiewicz AJ, Flynn A, Proud CG, Cooper JA. Mitogen-activated protein kinases activate the serine/threonine kinases Mnk1 and Mnk2. The EMBO Journal. 1997;**16**:1909-1920. DOI: 10.1093/emboj/16.8.1909

[84] Roux PP, Shahbazian D, Vu H, Holz MK, Cohen MS, Taunton J, et al. RAS/

ERK Signaling promotes site-specific ribosomal protein S6 phosphorylation via RSK and stimulates cap-dependent translation. Journal of Biological Chemistry. 2007;**282**:14056-14064. DOI: 10.1074/jbc.M700906200

[85] Shahbazian D, Roux PP, Mieulet V, Cohen MS, Raught B, Taunton J, et al. The mTOR/PI3K and MAPK pathways converge on eIF4B to control its phosphorylation and activity. The EMBO Journal. 2006;**25**:2781-2791. DOI: 10.1038/sj.emboj.7601166

[86] Warringer J, Hult M, Regot S, Posas F, Sunnerhagen P. The HOG pathway dictates the short-term translational response after hyperosmotic shock. Molecular Biology of the Cell. 2010;**21**:3080-3092. DOI: 10.1091/mbc.E10-01-0006

[87] Göke J, Chan Y-S, Yan J, Vingron M, Ng H-H. Genome-wide kinase-chromatin interactions reveal the regulatory network of ERK Signaling in human embryonic stem cells. Molecular Cell. 2013;**50**:844-855. DOI: 10.1016/J.MOLCEL.2013.04.030

[88] Nadal-Ribelles M, Solé C, Xu Z, Steinmetz LM, de Nadal E, Posas F. Control of Cdc28 CDK1 by a stress-induced lncRNA. Molecular Cell. 2014;**53**:549-561. DOI: 10.1016/j.molcel.2014.01.006

[89] Nelson JD, LeBoeuf RC, Bomsztyk K. Direct recruitment of insulin receptor and ERK signaling cascade to insulin-inducible gene loci. Diabetes. 2011;**60**:127-137. DOI: 10.2337/db09-1806

Chapter 3

Autophagy-Related Gene Expression Changes Are Found in Pancreatic Cancer and Neurodegenerative Diseases

Doaa M. Ali and Martin R. Berger

Abstract

Genetic alterations can cause cancer, including pancreatic cancer (PC) as well as certain neurodegenerative diseases. Our lab has recently identified genes that are modulated during pancreatic cancer liver metastasis, and some are known to have a role in neurobiology or neurodegenerative diseases. Autophagy or self-eating portrays the lysosomal-dependent degradation and recycling of protein aggregates and defective organisms in eukaryotic cells. Deregulation of autophagy as a cellular mechanism is common in neurodegenerative diseases as well as cancer and may represent a platform by which some genes can affect both disorders. This is exemplified for optineurin, which is an autophagy receptor that was found among genes with intensive modulation of expression in PC liver metastasis. Our results on this autophagy receptor draw the attention to the expression status of this and other autophagy genes in pancreatic cancer progression.

Keywords: pancreatic cancer, nervous system, neurodegenerative diseases, autophagy, optineurin

1. Introduction

Recent findings from microarray analyses of cancer cells have shown a growing list of genes with modulated expression, which are known to have importance in diseases other than cancer. This is in particular true for genes or gene families that have been identified to play a role in some neurodegenerative diseases or are a factor in cells of the nervous system, which may have a controlling role on the growth and development of cancer in general and of pancreatic cancer in particular. One cellular property, the alteration of which seems to be related to both types of diseases, is autophagy. In the lines below, we will discuss whether deregulation of autophagy could be a mechanism, which is in common between certain neurodegenerative diseases and pancreatic cancer, and thus may represent a link between two extremely different diseases.

2. Disrupted autophagy links cancer and neurodegenerative diseases

Autophagy is a Greek term that means self-eating and it portrays the lysosomal-dependent way of degrading and recycling cytosolic components in eukaryotic

cells. Autophagy is categorized into three different types, that is, microautophagy, macroautophagy, and chaperone-mediated autophagy (**Figure 1**), which are different in terms of their cargo and the mechanism of their occurrence [1, 2].

Micro-autophagy refers to the process where minute parts of the cytoplasm become sequestered and later on completely engulfed by lysosomal invaginations [3].

Chaperone-mediated autophagy is a selective pathway of autophagy, where proteins are targeted by the presence of a pentapeptide motif in the amino acid sequence of the substrate proteins, which is recognized by a chaperone protein. This complex is then delivered to the lysosomes for degradation in a receptor-dependent manner [2]. The motif should be accessible to the chaperone protein regardless of its position in the protein, but under normal conditions, it is concealed within the core when the protein is appropriately folded. Substrates for CMA are recognized by chaperones as heat shock cognate protein of 70 kD (hsp70) and cochaperones, which may be responsible for the unfolding of the protein before substrate-chaperone interaction can take place, but without direct interactions in some cases. Cochaperones include proteins PINK as hsp90, hsp40, and Bcl-2–associated athanogene 2 (Bag-1). Then, the substrates bind to the cytosolic tail of lysosome-associated membrane protein type 2A (LAMP-2A) and through its multimerization are translocated toward the lysosomal lumen, and this is seen as a limiting step for the pathway [2, 4, 5].

Macroautophagy, referred to below as autophagy, is the most widely investigated type of autophagy and entails the formation of what is called an autophagosome, which is a double membrane structure that is used to deliver cargos later through fusion to lysosomes or endosomes [6]. It is formed through several chronological steps starting from nucleation, elongation, to closure of a phagophore or isolation membrane that leads to the autophagosome formation. Proteins responsible to drive the autophagy process were basically discovered through analyzing the yeast genome and are named autophagy-related (ATG) proteins [6, 7].

Once stimulated, the autophagic process starts with the assembly of the initiation complex (ULK1 complex) and the nucleation complex (BECN1 complex) at the phagophore assembly site and this forms the basis for recruiting other ATG proteins and the elongation of the phagophore membrane [1]. ATG8/LC3 becomes bound to the inner and outer membrane of the phagophore following cleavage by the ATG4 protease and following conjugation to phosphatidylethanolamine (PE). Before

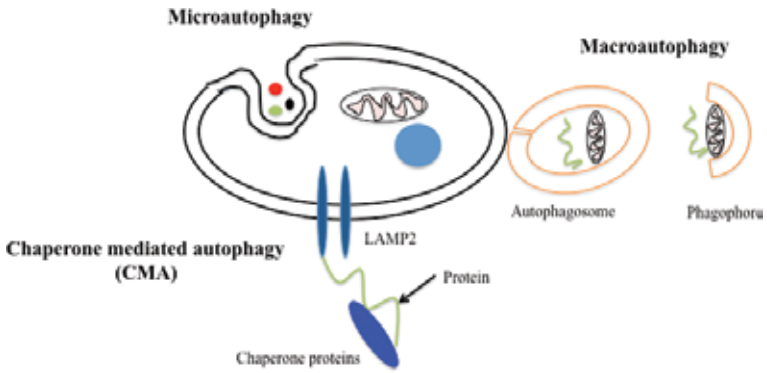

Figure 1.
Schematic representation of the three different types of autophagy showing the underlying difference in the mechanisms, microautophagy by invaginations of the lysosomes, macroautophagy through the formation of autophagosomes, and CMA with the targeted protein recognized by chaperones and delivered to the lysosomes through the LAMP2 protein.

closure, all ATGs dissociate and are recycled except for ATG8/LC3, which is located on the outer membrane. This protein is recycled after closure with the assistance of ATG4, while lysosomal enzymes in the autophagosome lumen degrade the ATG8/LC3 attached to the inner membrane.

Autophagosomes later unite with late endosomes and lysosomes to proceed for degradation [1]. LAMP proteins again control the fusion step and protect against degradation of the lysosomal membrane [1, 8].

Autophagy can be either selective or nonselective. During the selective process, specific cargos as aggregated proteins, damaged mitochondria, excess peroxisomes, and invading pathogens undergo degradation after being recognized via autophagy receptors. These autophagy receptors have the ability to recognize degradation signals on cargo proteins and also bind LC3/GABARAP proteins on the forming autophagosome. Among the identified autophagy receptors are p62/SQSTM1 (p62/Sequestosome 1), OPTN (optineurin), NBR1 (neighbor of BRCA1), and NDP52 (nuclear dot protein 52 kDa). All of them possess an ubiquitin-binding domain and LC3-interacting regions (LIRs) [9, 10].

Autophagy is a highly conserved process in mammals and is strictly regulated. One of the major regulators is mTOR through its complex 1 (mTORC1) and its activation exerts an inhibitory effect on autophagy induction. A similar effect is shown with PI3 kinases class I, whereas PI3 kinase class III activity is required for autophagosome formation. Starvation and amino acid depletion result in stimulation of the process. Other regulatory and autophagy inducers are based on an increase in cytosolic calcium, inhibition of inositol triphosphate, or starvation-induced autophagy [6, 11, 12]. Starvation also positively regulates autophagy by activation of AMPK that directly phosphorylates ULK1 or by inhibition of mTORC1 activity, or through inhibitory phosphorylation of nonautophagic BECN1 complexes [1]. Oxidative stress [13], DNA damage [14], and hypoxia [15] are all among the cellular mechanisms that can induce autophagy.

Autophagy has several cellular functions; it provides nutrients, eliminates damaged proteins and organelles, combats against invading microorganisms/pathogens, and in general keeps the homeostasis and balance in cells [1]. Because of its important role and significant cellular implications, defects or deregulation of this process was detected in different human diseases [1]. Infections and pathogens modulate autophagy according to their requirements to secure their survival in host cells. Certain steps of autophagy are hindered in neurodegenerative disorders and proteinopathies are a feature of these disorders. Autophagy is cytoprotective in cardiovascular tissue under physiological conditions, but it is induced in many cardiovascular diseases and is also deregulated in cancer, diabetes, and immune disorders [1].

2.1. Role of autophagy in cancer

In cancer, autophagy can either inhibit or promote cancer development and progression, and this varies according to the genetic lesions, tumor type, and stage. Combating mutagenic reactive oxygen species (ROS) accumulation, DNA damage, genomic instability, and oncogenic proteins are part of the protective functions of autophagy against tumor induction, as they induce autophagy when initiated [16, 17].

With regard to its onco-stimulatory role, downregulation of some autophagy-related genes, as Beclin-1 or ATG5, results in reduced growth of metastatic carcinoma cell lines, while that of ATG7 will promote apoptosis of colon cancer cells. Autophagy also permits tumors to resist stress and apoptotic signals and is connected in advanced cancer to poor prognosis and invasiveness. It increases ATP levels that support cell survival during hypoxia and starvation. Thus,

autophagy keeps healthy cells away from malignant transformation through maintaining the cellular homeostasis, but after tumor establishment, it may result in increased tumor progression and invasiveness [16, 17]. Autophagy supports tumor development through improving resistance of cells to endogenous apoptotic signals as well as resistance to chemotherapy and maintains the cancer stem cell compartment [17].

Pancreatic tumors show an elevated level of basal autophagy, more than any other epithelial cancer. The role of autophagy in PDAC needs to be clarified; as some investigations and the effect of anticancer drugs that generate ROS and induce autophagy give rise to the idea that autophagy is protective against cancer and stimulates apoptotic cancer cell death following treatment [18]. On the other hand, a number of reports suggested the role of autophagy in PDAC development. Autophagy was anticipated to be the cellular mechanism responsible for cancer development and progression from pancreatitis to overt carcinogenesis in the presence of K-ras mutations. Different observations supported a harmful role of autophagy, as being strongly induced in the inflammatory process. Inhibition of tumor development was observed after using the autophagy inhibitor chloroquine. Finally, LC3 is overexpressed in pancreatic cancer and downregulation of autophagy genes significantly reduces the growth and colony formation of PDAC cells in vitro [19, 20].

2.2. Role of autophagy in neurodegenerative diseases

Autophagic removal of proteins and damaged organelles is vital for the appropriate function of neurons, not only under pathological conditions, but also at baseline level under normal circumstances, as denoted by some studies [21].

Autophagosomes are very rarely detected in healthy neurons owing to the fact that the brain has a very low level of basal autophagy, or because the robust efficiency of the autophagy processes prevents the autophagosomes from accumulation [21, 22].

Dysfunction in autophagy affects the neuronal function and results in the occurrence of neurodegenerative diseases [23].

Induction of autophagy was anticipated to be an early stress response in axonal dystrophy and to contribute to axon remodeling [24]. Autophagy was shown to play a role in controlling microtubule dynamics and axon regeneration. Autophagy induction encouraged neurite growth, lessened the inhibitory effects of myelin, and reduced the formation of retraction bulbs after axonal injury in cultured cortical neurons [25].

Neurodegeneration is associated with deterioration of cognitive capabilities as well as motor functions with progressive damage to the nerve structure and functions. Alzheimer's, Parkinson's, amyotrophic lateral sclerosis, and Huntington's diseases represent the most famous neurodegenerative diseases [26]. They share in common the accumulation of toxic misfolded protein aggregates, as extracellular amyloid protein (Aβ) plaques and intracellular accumulation of tau protein in Alzheimer's disease, α-synuclein and ubiquitin in Parkinson's disease, and mutant Huntingtin protein in Huntington's disease [26]. Defective mitochondria, generating excessive reactive oxygen species (ROS), represent another feature of neurodegenerative diseases, as ROS harm the cellular constituents including DNA, lipids, and proteins [26]. The misfolded proteins are responsible for neuronal damage and death, and hence autophagy becomes a crucial target for the management of neurodegenerative diseases [27].

Recent studies revealed that alterations in autophagy-related genes trigger neurodegenerative diseases [28].

In Alzheimer's disease, transmembrane amyloid precursor protein (APP) is processed by β- and γ-secretases, which thus produce Aβ peptide [28, 29]. Autophagosomes show γ-secretase activity and represent a possibly active compartment for the production of Aβ and contribute to its deposition in affected neurons in Alzheimer's disease [30]. Dysfunction of Presenilin-1, associated with familial Alzheimer's disease, hinders proteolysis in autolysosomes and results in accumulation of Aβ peptides in autophagosomes [28]. Beclin-1 deficiency impedes the autophagic clearance of APP as an autophagy substrate, and it increases the pathology of Alzheimer's disease [31].

Huntington's disease is an inherited neurodegenerative disease and shows a range of declining motor, behavioral, and cognitive functions. Mutations in the huntingtin (HTT) gene are associated with cytosine-adenine-guanine (CAG) expansion, encoding a polyglutamine (polyQ) at the N-terminus of the gene [32]. This mutation affects the interaction of huntingtin with other proteins leading to all the neuropathological changes observed in the disease [33]. Huntingtin protein was shown to function as a scaffold protein for selective macroautophagy through interaction with autophagy pathway components; it cooperates with the autophagy receptor p62 and enables its connection with the essential autophagosome element LC3 and with ubiquitinated substrates. Huntingtin also binds to ULK1, the initiator kinase for autophagy, and this interaction liberates ULK1 from the negative regulation exerted by mTOR [34].

Parkinson's disease (PD) is a common progressive neurological disorder characterized by tremors, bradykinesia, rigidity, and loss of postural reflexes. The disease is caused by the loss of dopaminergic neurons in the *substantia nigra* [28, 35]. Different stages of Parkinson's disease showed deregulation of autophagy with the aggregation of α-synuclein in 'Lewy bodies.' Disrupted mitophagy is a major mechanism in the pathogenesis of the disease through mutations of PINK1 and PARKIN, which are essential for mitochondrial biogenesis and recycling [36]. Physiologically, PINK1 is a target of the ubiquitin-proteasome system after it has been processed by the mitochondrial protease presenilin-associated rhomboid-like (PARL). In depolarized mitochondria, this process is inhibited and PINK1 accumulates on the outer mitochondrial membrane, autophosphorylates, and recruits Parkin to damaged mitochondria. Parkin is an enzyme 3 (E3) ubiquitin ligase, and it was suggested that its ubiquitinated proteins recruit the autophagy receptor p62 to be integrated in an autophagosome for the autophagy degradation of damaged mitochondria (mitophagy) [28].

The two autophagy receptors previously linked to xenophagy, nuclear dot protein 52 kDa (NDP52) and optineurin (OPTN), were identified in a recent study to be the primary receptors for PINK1- and Parkin-mediated mitophagy. In this study, PINK1 recruits these two autophagy receptors, but not p62, to activate mitophagy directly, independent from Parkin. Upon their recruitment to mitochondria, NDP52 and OPTN recruit other autophagy components as ULK1, DFCP1, and WIPI1, thus revealing a function for these autophagy receptors upstream of LC3. While PINK1 begins mitophagy in the absence of Parkin, mitophagy is significantly boosted in the presence of Parkin [37].

Amyotrophic lateral sclerosis is a deadly progressive neurodegenerative disease characterized by the degeneration and deterioration of motor neurons. Beside the association of several genes in ALS, it has been reported that mutations in the Sigma-1 receptor (SigmaR1) are related to the autosomal recessive familial form of ALS [28, 38]. SigmaR1 controls calcium transport and its reduced expression can lead to the increased release of calcium from ER, the depolarization of mitochondrial membrane potential, and apoptosis. It reduces autophagic flux and autophagic degradation [28, 39]. Recent exome sequencing studies identified TANK-binding

kinase 1 (TBK1) as an important protein in both sporadic and familial ALS. TBK1 phosphorylates OPTN [40], and it was reported that a mutation of OPTN is also associated with ALS [41]. Furthermore, neuron-specific Atg5 and Atg7 knockout mice are associated with motor defects [42]. Beclin-1, the major player in the autophagic initiation process, has been shown to promote ALS by interacting with superoxide dismutase 1 (SOD1) and its downregulation was associated with aggregation of amyloid-β tau tangles, which are indicators of Alzheimer's disease (AD).

3. Link between autophagy and the nicotinic cholinergic system

The role of neurology genes in the progression and metastasis of cancer in general has been emphasized recently. In many reviews, Hildegard M. Schuller established the theory that modulation of the autonomic nervous system is responsible for driving cancer development, progression, and resistance to chemotherapy not only in non–small-cell lung cancer and pancreatic ductal adenocarcinoma but also in other cancers [43–48].

She anticipated the different modifiable risk factors of PDAC development, including smoking, chronic psychological stress, and habitual ingestion of alcohol, to act mostly via the sensitization (α7nAChR) and desensitization (α4nAChR) of nAChRs that drive the expression of proteins, which regulate the synthesis and release of catecholamines and GABA and result in the hyperactivity of Gs-mediated cAMP signaling [47]. α7nAChR activation is responsible for the production of stress neurotransmitters, which are known to promote cancer features, as well as of serotonin, dopamine, and glutamate, while the desensitization of α4nAChR is the main reason for reduced GABA levels, which is known to be a protective neurotransmitter [45]. This is believed to disturb the normal equilibrium and homeostasis of the signaling neurotransmitters in favor of cancer development, progression, and reduced responsiveness to treatment [44].

A link between the nicotinic cholinergic system and autophagy can be thought as follows: in one study, nicotine hindered the macrophage clearance of mycobacterium tuberculosis via inhibition of autophagosome formation in infected T helper cells and macrophages [49]. In another study, nicotine enhanced the proteasome activity and the total protein ubiquitination as well as autophagy as was proven by the occurrence of autophagic vacuoles and increased MAP LC3-II at protein level. These mechanisms help in the downregulation of connexin 43 and limit the communication in endothelial cells based on the involvement of nAChR α4β2 and α3β2, but not α7 [50]. Jeong et al. have shown that melatonin's neuroprotective effects against prion-mediated neural damage are owing to the activation of autophagy as an outcome of α7nAchR regulation [51].

4. Modulated expression of some neurologic genes in PDAC with special reference to autophagy

Our lab has identified a group of genes, which are modulated during pancreatic cancer liver metastasis [52]. The microarray data are derived from ASML rat PDAC cells that were reisolated from rat liver after they had been implanted intraportally and colonized the rat liver for various periods of time. These genes were arranged according to their fold change versus control cells (more than threefold modulation of expression) at early, intermediate, advanced, and final stages of metastasis, and the resulting genes were investigated by the Ingenuity Pathway Analysis program [53]. Among these genes, a subgroup was identified to have a dual role in

neurobiology and cancer metastasis. Thus, from ca. 30,000 genes, as found in a microarray experiment, 14 genes were selected as most promising. They are shown in **Table 1**. Data describing these genes are retrieved from the NCBI database [54].

From this list, some genes attracted our attention, including the autophagy receptor OPTN as well as MAOA/MAOB and CAV1.

What attracted our interest in this group of genes was that they share certain roles in cancer and neurology, and some were reported to have some direct or indirect link to autophagy.

The overexpression of the well-known mitochondrial enzyme MAOA, linked with many psychiatric illnesses, was shown to be associated with prostate cancer. MAOA caused neuroendocrine differentiation of prostate cancer cells through

	Gene symbol	Summary
Upregulated genes		
Receptor (G protein–coupled) activity modifying protein 1	Ramp1	Expression is increased in the striatum after repeated L-DOPA administration in a rat model of Parkinson's disease.
Activating transcription factor 3	Atf3	Expression is rapidly upregulated in response to nerve injury.
Aristaless-like homeobox 3	Alx3	Preferential methylation of this gene's promoter is associated with advanced-stage neuroblastoma.
Sortilin-related VPS10 domain-containing receptor 2	Sorcs2	These genes are strongly expressed in the central nervous system.
MAP/microtubule affinity-regulating kinase 4	Mark4	Expression is a potential marker for cancer, and the protein may also play a role in Alzheimer's disease.
Chloride intracellular channel 6	Clic6	Elevated expression is correlated with poor prognosis in human gliomas.
Downregulated genes		
Caveolin 1, caveolae protein	Cav1	The protein helps in coupling integrins to the Ras-ERK pathway and in promoting cell cycle progression.
Neuritin 1	Nrn1	The encoded protein promotes neurite outgrowth suggesting its role in promoting neuritogenesis.
Monoamine oxidase A/B	MaoA/MaoB	They have an important role in the metabolism of neuroactive and vasoactive amines in the central nervous system and peripheral tissues.
5-hydroxytryptamine (serotonin) receptor 6, G protein–coupled	Htr6	This receptor is thought to regulate cholinergic neuronal transmission in the brain.
Ly6/neurotoxin 1	Lynx1	It has the capacity to enhance nicotinic acetylcholine receptor function in the presence of acetylcholine.
Calcium channel, voltage-dependent, beta 4 subunit	Cacnb4	Certain mutations in this gene have been associated with idiopathic generalized epilepsy (IGE) and juvenile myoclonic epilepsy (JME). Multiple transcript variants for different isoforms have been found.
Optineurin	Optn	Optineurin plays a role in normal-tension glaucoma and adult-onset primary open angle glaucoma.

Table 1.
Genes with modulated expression.

activation of autophagy, ROS production, and downregulation of repressor element-1 silencing transcription factor (REST) [55].

However, at mRNA level, MAOA was not measurably expressed in our panel of human PDAC cell lines. Another aspect, which prevented further experiments, was that the irreversible MAOA-inhibiting drug clorgyline halted the proliferation of BXPC3 cells only at very high concentrations, which rendered further studies of this gene unpromising.

Deficiency of the caveolae protein CAV1 stimulates the basal and inducible autophagy as a cell survival mechanism under starvation. This is a recently described function of CAV1 and lipid rafts in breast cancer development via modulation of lysosomal function and autophagy [56]. Another study reported that CAV1 regulates autophagy positively under oxidative stress and cerebral ischemic injury. CAV1 deficiency limited localization of Beclin-1 (BECN1) to the mitochondria and eliminated LC3 foci formation in response to hydrogen peroxide in the brain of CAV1 knockout mice [57]. CAV1 is related to tumorigenesis and metastasis. CAV1 overexpression in PDAC is associated with poor clinical outcome, as well as chemo- and radioresistance [58]. Moreover, signs of premature neuronal aging and degeneration are evident in CAV1 knockout mice together with increased Aβ, P-tau, and astrogliosis [59].

The gene optineurin (OPTN), which is an autophagy receptor, was also among the genes listed in **Table 1**. Since OPTN has an important role in many neurodegenerative diseases as well as a function as an autophagy receptor, thus linking autophagy in cancer and neurodegenerative diseases, a comprehensive overview of this gene is given as well as an outline of the expression of other autophagy genes in PDAC.

5. Optineurin

OPTN is linked to many neurodegenerative diseases including normal tension glaucoma, primary open-angle glaucoma [60], as well as amyotrophic lateral sclerosis [41].

The OPTN gene is located on chromosome 10p13 and translates into a protein of 67-kDa. The mRNA consists of 16 exons; the first 3 are noncoding, while the coding exons give rise to a protein composed of 577 amino acids [61, 62]. OPTN was found to be head to head oriented with the gene coiled-coil domain containing 3 (CCDC3) with a distance of about 98-kb between their 5'UTRs [63]. OPTN has a half-life of around 8 h and its degradation engages the ubiquitin proteasomal system as its level increases following exposure to a proteasome inhibitor, but not to autophagic or lysosomal inhibitors [64].

OPTN encompasses several domains: an NF-κB-essential molecule-like domain, leucine zipper motif, coiled-coil motifs, an ubiquitin-binding domain (UBD), an LC3-interacting region, and a carboxyl (C)-terminal C_2H_2 type of zinc finger [65, 66].

As a cytoplasmic protein [60], OPTN colocalizes together with myosin VI and Rab8 around the Golgi complex and in vesicles at the plasma membrane [67]. It shows a high level of expression in certain tissues such as retina, brain, heart, skeletal muscle, placenta, testis, and kidney [68].

It interacts with itself to form homo-oligomers [69] and also with other molecules such as Ras-related protein 8 (Rab8) [70], huntingtin [71], myosin VI [72], transferrin receptor [68], LC3/GABARAP [62], polo-like kinase 1 [73, 74], TBK1 [75], as well as metabotropic glutamate receptor, transcription factor IIIA, serine/threonine kinase receptor-interacting protein 1, CYLD lysine 63 deubiquitinase (cylindromatosis, CYLD), and HECT (homologous to the E6-AP carboxyl terminus) domain and ankyrin

repeat containing E3 ubiquitin protein ligase 1 (HACE1) [65]. The structure of OPTN and some interactions are shown in **Figure 2**.

OPTN plays different roles under physiological conditions including membrane trafficking, maintenance of the Golgi apparatus, exocytosis, protein secretion, cell division control, regulation of NF-κB, and host defense against pathogens [61, 62, 65, 67].

Its antiviral preventive response was based on its ability to regulate the interferon response in a cell cycle–dependent manner owing to its nuclear translocation together with deubiquitinating enzyme CYLD during the G2/M phase of the cell cycle, which abolishes the inhibitory effect it exerts on TBK1. As a result, the TBK1 activity is increased with enhanced interferon production [76].

The overexpression of OPTN was demonstrated to be protective against H_2O_2-mediated cell death, a function, which is compromised by a mutated form of OPTN (E50K), resulting in cells that are less fit to survive under stress conditions [70].

Beside all the above stated roles, OPTN is considered as a disease-linked gene. Mutations of OPTN or its altered expression are associated with multiple diseases including normal tension glaucoma and primary open-angle glaucoma [60] as well as plenty of neurodegenerative diseases including amyotrophic lateral sclerosis [41], ubiquitin-positive intraneuronal inclusions in ALS with dementia, basophilic inclusions in the basophilic type of ALS, neurofibrillary tangles and dystrophic neurites in Alzheimer's disease, Lewy bodies and Lewy neurites in Parkinson's disease, ballooned neurons in Creutzfeldt-Jakob disease, glial cytoplasmic inclusions in multiple system atrophy, and Pick bodies in Pick's disease with unknown significance [77]. Reduced OPTN expression in humans might increase the risk of developing Crohn's disease [78] and dispose to the occurrence of Paget's disease by enhancing osteoclast differentiation, as OPTN is a recently identified regulator of bone resorption [79].

The OPTN protein was identified as a selective autophagy receptor such as the multidomain scaffold/adaptor protein p62/sequestosome-1 (p62/SQSTM-1) and nuclear domain 10 protein 52 (NDP52) in terms of binding to polyubiquitinated cargoes and brings them to autophagosomes via its LC3-interacting region [65, 80–82]. In this context, it can help cells to get rid of pathogens as salmonella [80], defective mitochondria [82], and misfolded protein aggregates [83] or can have a role in tumor suppression [81].

The role of OPTN in autophagy can also be independent from ubiquitination [65], where it can distinguish several protein aggregates through its C-terminal coiled-coil domain [84].

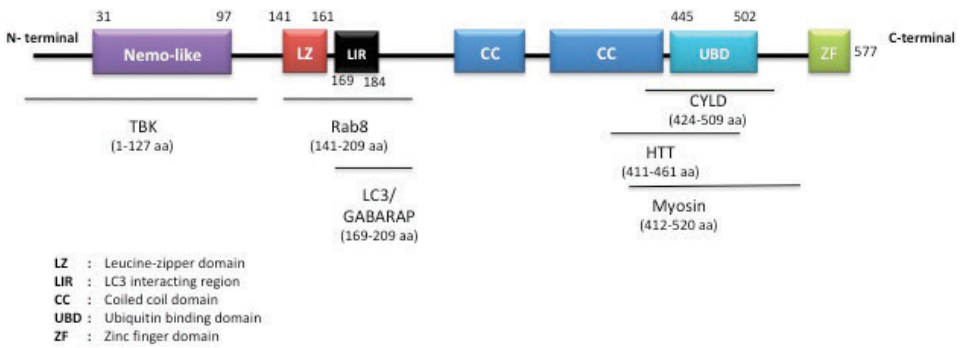

Figure 2.
Graphic illustration of OPTN structural domains and the localization of these domains relative to its amino acid sequence as well as some of its interacting proteins.

The role of OPTN in autophagy extends beyond being an autophagy receptor, as OPTN also plays a role as an autophagy inducer. Overexpression of OPTN in its wild type or mutated E50K forms, or its upregulation by cytokine treatment, was linked to elevated LC3-II levels in retinal ganglion cells, while the level of the proteasome activity marker PSMB5 (proteasome regulatory β 5) was reduced denoting induction of autophagy [64]. In vivo results confirmed the induction of autophagy and the reduction of ubiquitin proteasome pathway upon the injection of wild type and E50K OPTN transfected vectors in rat eyes [85].

Mutation or altered expression of OPTN may result in many diseases as glaucoma, ALS, other neurodegenerative diseases [65], or even cancer [81] owing to the implication of mitochondrial dysfunction and protein aggregation [65]. Mutations in both autophagy receptors p62 and OPTN in Paget's disease of the bone may attribute to an autophagy-related mechanism of developing this disease [65].

An overview of the RNA-seq data generated by the Cancer Genome Atlas (TCGA) revealed high expression of the OPTN gene across several cancer types (reported as median number fragments per kilo-base of exon per million reads (FPKM)) and was based on mRNA expression in cancer tissues (see **Figure 3**). In this regard and with data retrieved from the Protein Atlas website [86, 87], pancreatic cancer represents the tumor with the second highest OPTN expression and is topped only by renal cancer.

Prompted by this result, we analyzed the expression of autophagy genes and autophagy receptors in PDAC in a TCGA cohort with a sample size of 179. Most of the autophagy genes were above average genomic expression, represented by >7.5 log2 rsem (RNA-seq by expectation maximization) except for ATG10, which was below average. OPTN was second highest expressed of all autophagy genes, preceded only by SQSTM1 (**Figure 4**).

Based on these findings, we reasoned that OPTN might be involved in important mechanisms associated with cancer as well as with neurodegenerative diseases that require further analysis.

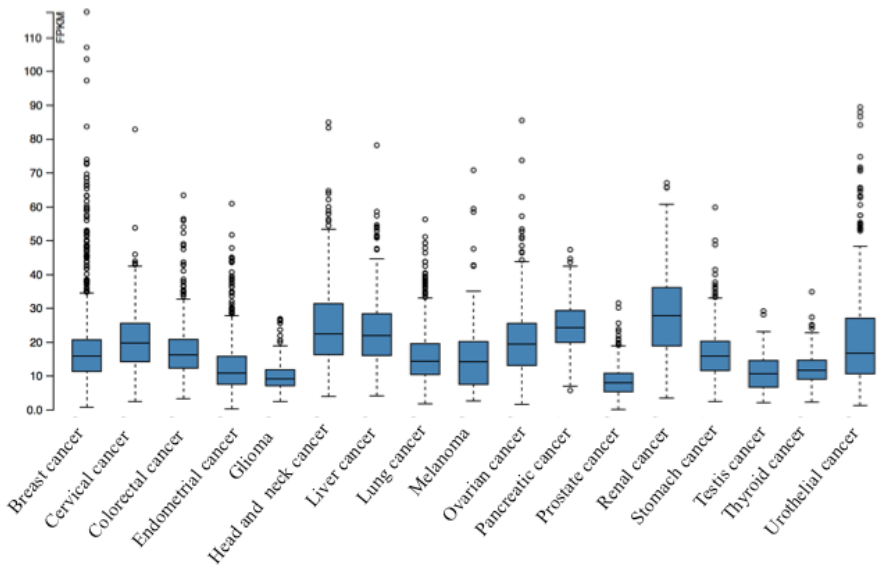

Figure 3.
Overview of the RNA-seq data in TCGA of 17 cancer cohorts showing the highest OPTN expression in renal cell carcinoma followed by PDAC (expressed in terms of median FPKM) [86–88].

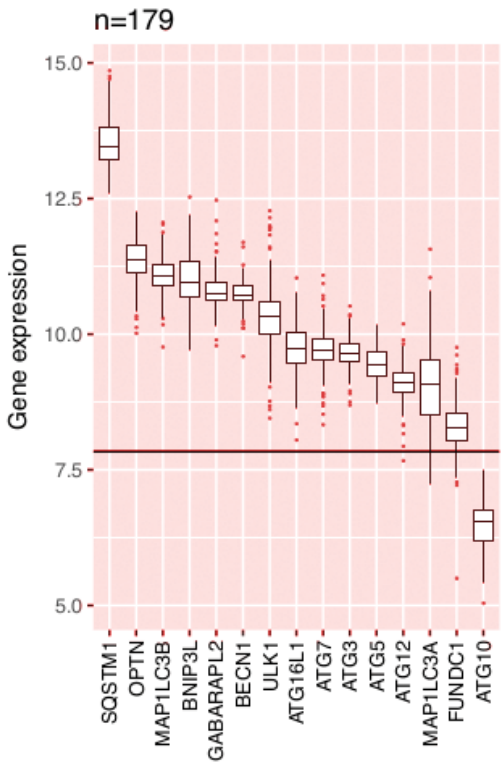

Figure 4.
Expression of autophagy-related genes and receptors in a PDAC cohort with a sample size of 179 patients showing high expression of all autophagy-related genes (> 7.5 log2 rsem) except for ATG10.

6. Conclusions

In summary, autophagy represents a link between the nervous system and cancer in general as well as pancreatic cancer in particular. The modulated expression of autophagy-related genes in cancer and neurodegenerative diseases highlights the importance of this mechanism and suggests further studies on their regulation and effective targeting. Autophagy is a new background for certain genes in cancer as well as in neurology, and understanding this process may well serve as a platform for understanding the pathogenesis of these diseases in different organs. In addition, there seems to be a special role for OPTN in different neurodegenerative diseases as well as cancer and this protein could well be a target in cancer treatment. The link between nicotinic acetylcholine receptors and autophagy still requires further studies to be properly delineated.

Acknowledgements

The authors express their gratitude to Dr. Ashwini K. Sharma for the generation of **Figure 4** and his continuous support with questions related to bioinformatics. Doaa M. Ali was supported by a PhD Grant from the Deutscher Akademischer Austauschdienst (DAAD).

Conflict of interest

The authors have no conflicting interests.

Acronyms and abbreviations

OPTN	Optineurin
PDAC	pancreatic ductal adenocarcinoma
CMA	chaperone-mediated autophagy
ATG	autophagy-related protein
BECN1	Beclin-1
ULK1	Unc-51 like autophagy activating kinase
MAP1LC3A	microtubule-associated protein 1 light chain 3 alpha
MAP1LC3B	microtubule-associated protein 1 light chain 3 beta
GABARAPL2	GABA type A receptor-associated protein like 2
SQSTM1	sequestosome 1
BNIP3L	BCL2-interacting protein 3 like
BNIP3	BCL2-interacting protein 3
FUNDC1	FUN14 domain containing 1

Author details

Doaa M. Ali[1,2] and Martin R. Berger[1*]

1 Toxicology and Chemotherapy Unit, German Cancer Research Centre (DKFZ), Heidelberg, Germany

2 Department of Pharmacology and Experimental Therapeutics, Medical Research Institute, Alexandria University, Egypt

*Address all correspondence to: m.berger@dkfz.de

References

[1] Lippai M, Szatmari Z. Autophagy-from molecular mechanisms to clinical relevance. Cell Biology and Toxicology. 2017;**33**:145-168. DOI: 10.1007/s10565-016-9374-5

[2] Bejarano E, Cuervo AM. Chaperone-mediated autophagy. Proceedings of the American Thoracic Society. 2010;7:29-39. DOI: 10.1513/pats.200909-102JS

[3] Mijaljica D, Prescott M, Devenish RJ. Microautophagy in mammalian cells: Revisiting a 40-year-old conundrum. Autophagy. 2014;(7):673-682. DOI: 10.4161/auto.7.7.14733

[4] Agarraberes FA, Dice JF. A molecular chaperone complex at the lysosomal membrane is required for protein translocation. Journal of Cell Science. 2001;**114**:2491-2499

[5] Bandyopadhyay U, Kaushik S, Varticovski L, Cuervo AM. The chaperone-mediated autophagy receptor organizes in dynamic protein complexes at the lysosomal membrane. Molecular and Cellular Biology. 2008;**28**:5747-5763. DOI: 10.1128/MCB.02070-07

[6] Eskelinen E-L. New insights into the mechanisms of macroautophagy in mammalian cells. International Review of Cell and Molecular Biology. 2008:207-247

[7] Klionsky DJ, Cregg JM, Dunn WA Jr, Emr SD, Sakai Y, Sandoval IV, et al. A unified nomenclature for yeast autophagy-related genes. Developmental Cell. 2003;**45**:539-545

[8] Huynh KK, Eskelinen EL, Scott CC, Malevanets A, Saftig P, Grinstein S. LAMP proteins are required for fusion of lysosomes with phagosomes. The EMBO Journal. 2007;**26**:313-324. DOI: 10.1038/sj.emboj.7601511

[9] Deng Z, Purtell K, Lachance V, Wold MS, Chen S, Yue Z. Autophagy receptors and neurodegenerative diseases. Trends in Cell Biology. 2017;**27**:491-504. DOI: 10.1016/j.tcb.2017.01.001

[10] Stolz A, Ernst A, Dikic I. Cargo recognition and trafficking in selective autophagy. Nature Cell Biology. 2014;**16**:495-501. DOI: 10.1038/ncb2979

[11] Petiot A, Ogier-Denis E, Blommaart EF, Meijer AJ, Codogno P. Distinct classes of phosphatidylinositol 3′-kinases are involved in signaling pathways that control macroautophagy in HT-29 cells. The Journal of Biological Chemistry. 2000;**275**:992-998

[12] Jewell JL, Russell RC, Guan KL. Amino acid signalling upstream of mTOR. Nature Reviews. Molecular Cell Biology. 2013;**14**:133-139. DOI: 10.1038/nrm3522

[13] Filomeni G, De Zio D, Cecconi F. Oxidative stress and autophagy: The clash between damage and metabolic needs. Cell Death and Differentiation. 2015;**22**:377-388. DOI: 10.1038/cdd.2014.150

[14] Gomes LR, Menck CFM, Leandro GS. Autophagy roles in the modulation of DNA repair pathways. International Journal of Molecular Sciences. 2017;**18**. DOI: 10.3390/ijms18112351

[15] Fang Y, Tan J, Zhang Q. Signaling pathways and mechanisms of hypoxia-induced autophagy in the animal cells. Cell Biology International. 2015;**39**:891-898. DOI: 10.1002/cbin.10463

[16] New M, Van Acker T, Long JS, Sakamaki JI, Ryan KM, Tooze SA. Molecular pathways controlling autophagy in pancreatic cancer. Frontiers in Oncology. 2017;7:28. DOI: 10.3389/fonc.2017.00028

[17] Galluzzi L, Pietrocola F, Bravo-San Pedro JM, Amaravadi RK, Baehrecke EH, Cecconi F, et al. Autophagy in malignant transformation and cancer progression. The EMBO Journal. 2015;**34**:856-880. DOI: 10.15252/embj.201490784

[18] Grasso D, Garcia MN, Iovanna JL. Autophagy in pancreatic cancer. International Journal of Cell Biology. 2012;**2012**:760498. DOI: 10.1155/2012/760498

[19] Kang R, Tang D. Autophagy in pancreatic cancer pathogenesis and treatment. American Journal of Cancer Research. 2012;**42**:383-396

[20] Iovanna JL. Autophagy contributes to the initiation of pancreatic cancer. Medical Science (Paris). 2017;**33**:335-339. DOI: 10.1051/medsci/20173303022

[21] Nikoletopoulou V, Papandreou ME, Tavernarakis N. Autophagy in the physiology and pathology of the central nervous system. Cell Death and Differentiation. 2015;**22**:398-407. DOI: 10.1038/cdd.2014.204

[22] Boland B, Kumar A, Lee S, Platt FM, Wegiel J, Yu WH, et al. Autophagy induction and autophagosome clearance in neurons: Relationship to autophagic pathology in Alzheimer's disease. The Journal of Neuroscience. 2008;**28**:6926-6937. DOI: 10.1523/JNEUROSCI.0800-08.2008

[23] Hara T, Nakamura K, Matsui M, Yamamoto A, Nakahara Y, Suzuki-Migishima R, et al. Suppression of basal autophagy in neural cells causes neurodegenerative disease in mice. Nature. 2006;**441**:885-889. DOI: 10.1038/nature04724

[24] Wang QJ, Ding Y, Kohtz DS, Mizushima N, Cristea IM, Rout MP, et al. Induction of autophagy in axonal dystrophy and degeneration. The Journal of Neuroscience. 2006;**26**:8057-8068. DOI: 10.1523/JNEUROSCI.2261-06.2006

[25] He M, Ding Y, Chu C, Tang J, Xiao Q, Luo ZG. Autophagy induction stabilizes microtubules and promotes axon regeneration after spinal cord injury. Proceedings of the National Academy of Sciences of the United States of America. 2016;**113**:11324-11329. DOI: 10.1073/pnas.1611282113

[26] Kim M, Ho A, Lee JH. Autophagy and human neurodegenerative diseases-A Fly's perspective. International Journal of Molecular Sciences. 2017;**18**. DOI: 10.3390/ijms18071596

[27] Fujikake N, Shin M, Shimizu S. Association between autophagy and neurodegenerative diseases. Frontiers in Neuroscience. 2018;**12**:255. DOI: 10.3389/fnins.2018.00255

[28] Kiriyama Y, Nochi H. The function of autophagy in neurodegenerative diseases. International Journal of Molecular Sciences. 2015;**16**:26797-26812. DOI: 10.3390/ijms161125990

[29] Zhou ZD, Chan CH, Ma QH, Xu XH, Xiao ZC, Tan EK. The roles of amyloid precursor protein (APP) in neurogenesis: Implications to pathogenesis and therapy of Alzheimer disease. Cell Adhesion & Migration. 2011;**45**:280-292

[30] Yu WH, Kumar A, Peterhoff C, Shapiro Kulnane L, Uchiyama Y, Lamb BT, et al. Autophagic vacuoles are enriched in amyloid precursor protein-secretase activities: Implications for beta-amyloid peptide over-production and localization in Alzheimer's disease. The International Journal of Biochemistry & Cell Biology. 2004;**36**:2531-2540. DOI: 10.1016/j.biocel.2004.05.010

[31] Jaeger PA, Pickford F, Sun CH, Lucin KM, Masliah E, Wyss-Coray T. Regulation of amyloid precursor protein processing by the Beclin 1 complex. PLoS One. 2010;**5**:e11102. DOI: 10.1371/journal.pone.0011102

[32] Dayalu P, Albin RL. Huntington disease: Pathogenesis and treatment. Neurologic Clinics. 2015;**33**:101-114. DOI: 10.1016/j.ncl.2014.09.003

[33] Li SH, Li XJ. Huntingtin-protein interactions and the pathogenesis of Huntington's disease. Trends in Genetics. 2004;**20**:146-154. DOI: 10.1016/j.tig.2004.01.008

[34] Rui YN, Xu Z, Patel B, Chen Z, Chen D, Tito A, et al. Huntingtin functions as a scaffold for selective macroautophagy. Nature Cell Biology. 2015;**17**:262-275. DOI: 10.1038/ncb3101

[35] Jankovic J. Parkinson's disease: Clinical features and diagnosis. Journal of Neurology, Neurosurgery, and Psychiatry. 2008;**479**:368-376. DOI: 10.1136/jnnp.2007.131045

[36] Gasser T. Molecular pathogenesis of Parkinson disease: Insights from genetic studies. Expert Reviews in Molecular Medicine. 2009;**11**:e22. DOI: 10.1017/S1462399409001148

[37] Lazarou M, Sliter DA, Kane LA, Sarraf SA, Wang C, Burman JL, et al. The ubiquitin kinase PINK1 recruits autophagy receptors to induce mitophagy. Nature. 2015;**524**:309-314. DOI: 10.1038/nature14893

[38] Chen S, Sayana P, Zhang X, Le W. Genetics of amyotrophic lateral sclerosis: An update. Molecular Neurodegeneration. 2013;**8**:28. DOI: 10.1186/1750-1326-8-28

[39] Fukunaga K, Shinoda Y, Tagashira H. The role of SIGMAR1 gene mutation and mitochondrial dysfunction in amyotrophic lateral sclerosis. Journal of Pharmacological Sciences. 2015;**127**:36-41. DOI: 10.1016/j.jphs.2014.12.012

[40] Cirulli ET, Lasseigne BN, Petrovski S, Sapp PC, Dion PA, Leblond CS, et al. Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. Science. 2015;**347**:1436-1441. DOI: 10.1126/science.aaa3650

[41] Maruyama H, Morino H, Ito H, Izumi Y, Kato H, Watanabe Y, et al. Mutations of optineurin in amyotrophic lateral sclerosis. Nature. 2010;**465**:223-226. DOI: 10.1038/nature08971

[42] Metaxakis A, Ploumi C, Tavernarakis N. Autophagy in age-associated neurodegeneration. Cell. 2018;**7**. DOI: 10.3390/cells7050037

[43] Schuller HM. Neurotransmitter receptor-mediated signaling pathways as modulators of carcinogenesis. Progress in Experimental Tumor Research. 2007;**39**:45-63. DOI: 10.1159/000100045

[44] Schuller HM. Neurotransmission and cancer: Implications for prevention and therapy. Anti-Cancer Drugs. 2008;**19**:655-671. DOI: 10.1097/CAD.0b013e3283025b58

[45] Schuller HM. Regulatory role of the alpha7nAChR in cancer. Current Drug Targets. 2012;**13**:680-687

[46] Schuller HM. Effects of tobacco constituents and psychological stress on the beta-adrenergic regulation of non-small cell lung cancer and pancreatic cancer: Implications for intervention. Cancer Biomarkers. 2013;**13**:133-144. DOI: 10.3233/CBM-130323

[47] Schuller HM. The neuro-psychological axis of pancreatic cancer as a novel target for intervention. Pancreatic Disorders & Therapy. 2013;**03**. DOI: 10.4172/2165-7092.1000124

[48] Schuller HM, Zhang L, Weddle DL, Castonguay A, Walker K, Miller MS. The cyclooxygenase inhibitor ibuprofen and the FLAP inhibitor MK886 inhibit pancreatic carcinogenesis induced in hamsters by transplacental exposure to ethanol and

the tobacco carcinogen NNK. Journal of Cancer Research and Clinical Oncology. 2002;**128**:525-532. DOI: 10.1007/s00432-002-0365-y

[49] Bai X, Stitzel JA, Bai A, Zambrano CA, Phillips M, Marrack P, et al. Nicotine impairs macrophage control of Mycobacterium tuberculosis. American Journal of Respiratory Cell and Molecular Biology. 2017;**57**:324-333. DOI: 10.1165/rcmb.2016-0270OC

[50] Duerrschmidt N, Hagen A, Gaertner C, Wermke A, Nowicki M, Spanel-Borowski K, et al. Nicotine effects on human endothelial intercellular communication via alpha4beta2 and alpha3beta2 nicotinic acetylcholine receptor subtypes. Naunyn-Schmiedeberg's Archives of Pharmacology. 2012;**385**:621-632. DOI: 10.1007/s00210-012-0738-y

[51] Jeong JK, Park SY. Melatonin regulates the autophagic flux via activation of alpha-7 nicotinic acetylcholine receptors. Journal of Pineal Research. 2015;**59**:24-37. DOI: 10.1111/jpi.12235

[52] Al-Taee KK, Ansari S, Hielscher T, Berger MR, Adwan H. Metastasis-related processes show various degrees of activation in different stages of pancreatic cancer rat liver metastasis. Oncology Research and Treatment. 2014;**37**:464-470. DOI: 10.1159/000365496

[53] IPA (Qiagen Inc.). 2014. Available from: https://apps.ingenuity.com

[54] Geer LY, Marchler-Bauer A, Geer RC, Han L, He J, He S, et al. The NCBI BioSystems database. Nucleic Acids Research. 2010;**38**:D492-D496. DOI: 10.1093/nar/gkp858

[55] Lin YC, Chang YT, Campbell M, Lin TP, Pan CC, Lee HC, et al. MAOA-a novel decision maker of apoptosis and autophagy in hormone refractory neuroendocrine prostate cancer cells. Scientific Reports. 2017;**7**:46338. DOI: 10.1038/srep46338

[56] Shi Y, Tan SH, Ng S, Zhou J, Yang ND, Koo GB, et al. Critical role of CAV1/caveolin-1 in cell stress responses in human breast cancer cells via modulation of lysosomal function and autophagy. Autophagy. 2015;**11**:769-784. DOI: 10.1080/15548627.2015.1034411

[57] Nah J, Yoo SM, Jung S, Jeong EI, Park M, Kaang BK, et al. Phosphorylated CAV1 activates autophagy through an interaction with BECN1 under oxidative stress. Cell Death & Disease. 2017;**8**:e2822. DOI: 10.1038/cddis.2017.71

[58] Lin HH, Lin HK, Lin IH, Chiou YW, Chen HW, Liu CY, et al. Mechanical phenotype of cancer cells: Cell softening and loss of stiffness sensing. Oncotarget. 2015;**6**:20946-20958. DOI: 10.18632/oncotarget.4173

[59] Head BP, Peart JN, Panneerselvam M, Yokoyama T, Pearn ML, Niesman IR, et al. Loss of caveolin-1 accelerates neurodegeneration and aging. PLoS One. 2010;**5**:e15697. DOI: 10.1371/journal.pone.0015697

[60] Chalasani ML, Swarup G, Balasubramanian D. Optineurin and its mutants: Molecules associated with some forms of glaucoma. Ophthalmic Research. 2009;**442**:176-184. DOI: 10.1159/000232400

[61] Kachaner D, Genin P, Laplantine E, Weil R. Toward an integrative view of Optineurin functions. Cell Cycle. 2012;**11**:2808-2818. DOI: 10.4161/cc.20946

[62] Ying H, Yue BY. Cellular and molecular biology of optineurin. International Review of Cell and Molecular Biology. 2012;**294**:223-258. DOI: 10.1016/B978-0-12-394305-7.00005-7

[63] Schilter KF, Reis LM, Sorokina EA, Semina EV. Identification of an Alu-repeat-mediated deletion of OPTN upstream region in a patient with a complex ocular phenotype. Molecular Genetics & Genomic Medicine. 2015;**3**:490-499. DOI: 10.1002/mgg3.159

[64] Shen X, Ying H, Qiu Y, Park JS, Shyam R, Chi ZL, et al. Processing of optineurin in neuronal cells. The Journal of Biological Chemistry. 2011;**286**:3618-3629. DOI: 10.1074/jbc.M110.175810

[65] Ying H, Yue BY. Optineurin: The autophagy connection. Experimental Eye Research. 2016;**144**:73-80. DOI: 10.1016/j.exer.2015.06.029

[66] Turturro S, Shen X, Shyam R, Yue BY, Ying H. Effects of mutations and deletions in the human optineurin gene. Springerplus. 2014;**3**:99. DOI: 10.1186/2193-1801-3-99

[67] Sahlender DA, Roberts RC, Arden SD, Spudich G, Taylor MJ, Luzio JP, et al. Optineurin links myosin VI to the Golgi complex and is involved in Golgi organization and exocytosis. The Journal of Cell Biology. 2005;**169**:285-295. DOI: 10.1083/jcb.200501162

[68] Nagabhushana A, Chalasani ML, Jain N, Radha V, Rangaraj N, Balasubramanian D, et al. Regulation of endocytic trafficking of transferrin receptor by optineurin and its impairment by a glaucoma-associated mutant. BMC Cell Biology. 2010;**11**:4. DOI: 10.1186/1471-2121-11-4

[69] Gao J, Ohtsubo M, Hotta Y, Minoshima S. Oligomerization of optineurin and its oxidative stress- or E50K mutation-driven covalent cross-linking: Possible relationship with glaucoma pathology. PLoS One. 2014;**9**:e101206. DOI: 10.1371/journal.pone.0101206

[70] De Marco N, Buono M, Troise F, Diez-Roux G. Optineurin increases cell survival and translocates to the nucleus in a Rab8-dependent manner upon an apoptotic stimulus. The Journal of Biological Chemistry. 2006;**281**:16147-16156. DOI: 10.1074/jbc.M601467200

[71] del Toro D, Alberch J, Lazaro-Dieguez F, Martin-Ibanez R, Xifro X, Egea G, et al. Mutant huntingtin impairs post-Golgi trafficking to lysosomes by delocalizing optineurin/Rab8 complex from the Golgi apparatus. Molecular Biology of the Cell. 2009;**20**:1478-1492. DOI: 10.1091/mbc.E08-07-0726

[72] Bond LM, Peden AA, Kendrick-Jones J, Sellers JR, Buss F. Myosin VI and its binding partner optineurin are involved in secretory vesicle fusion at the plasma membrane. Molecular Biology of the Cell. 2011;**22**:54-65. DOI: 10.1091/mbc.E10-06-0553

[73] Kachaner D, Filipe J, Laplantine E, Bauch A, Bennett KL, Superti-Furga G, et al. Plk1-dependent phosphorylation of optineurin provides a negative feedback mechanism for mitotic progression. Molecular Cell. 2012;**445**:553-566. DOI: 10.1016/j.molcel.2011.12.030

[74] Kachaner D, Laplantine E, Genin P, Weil R. Optineurin: A new vision of cell division control. Cell Cycle. 2012;**11**:1481-1482. DOI: 10.4161/cc.20116

[75] Gleason CE, Ordureau A, Gourlay R, Arthur JS, Cohen P. Polyubiquitin binding to optineurin is required for optimal activation of TANK-binding kinase 1 and production of interferon beta. The Journal of Biological Chemistry. 2011;**286**:35663-35674. DOI: 10.1074/jbc.M111.267567

[76] Genin P, Cuvelier F, Lambin S, Corte-Real Filipe J, Autrusseau E, Laurent C, et al. Optineurin regulates the interferon response in a cell cycle-dependent manner. PLoS Pathogens. 2015;**411**:e1004877. DOI: 10.1371/journal.ppat.1004877

[77] Osawa T, Mizuno Y, Fujita Y, Takatama M, Nakazato Y, Okamoto K. Optineurin in neurodegenerative diseases. Neuropathology. 2011;**31**:569-574. DOI: 10.1111/j.1440-1789.2011.01199.x

[78] Chew TS, O'Shea NR, Sewell GW, Oehlers SH, Mulvey CM, Crosier PS, et al. Optineurin deficiency in mice contributes to impaired cytokine secretion and neutrophil recruitment in bacteria-driven colitis. Disease Models & Mechanisms. 2015;**8**:817-829. DOI: 10.1242/dmm.020362

[79] Obaid R, Wani SE, Azfer A, Hurd T, Jones R, Cohen P, et al. Optineurin negatively regulates osteoclast differentiation by modulating NF-kappaB and interferon signaling: Implications for Paget's disease. Cell Reports. 2015;**13**:1096-1102. DOI: 10.1016/j.celrep.2015.09.071

[80] Wild P, Farhan H, McEwan DG, Wagner S, Rogov VV, Brady NR, et al. Phosphorylation of the autophagy receptor optineurin restricts Salmonella growth. Science. 2011;**333**:228-233. DOI: 10.1126/science.1205405

[81] Liu Z, Chen P, Gao H, Gu Y, Yang J, Peng H, et al. Ubiquitylation of autophagy receptor Optineurin by HACE1 activates selective autophagy for tumor suppression. Cancer Cell. 2014;**26**:106-120. DOI: 10.1016/j.ccr.2014.05.015

[82] Wong YC, Holzbaur EL. Optineurin is an autophagy receptor for damaged mitochondria in parkin-mediated mitophagy that is disrupted by an ALS-linked mutation. Proceedings of the National Academy of Sciences of the United States of America. 2014;**111**:E4439-E4448. DOI: 10.1073/pnas.1405752111

[83] Shen WC, Li HY, Chen GC, Chern Y, Tu PH. Mutations in the ubiquitin-binding domain of OPTN/optineurin interfere with autophagy-mediated degradation of misfolded proteins by a dominant-negative mechanism. Autophagy. 2015;**411**:685-700. DOI: 10.4161/auto.36098

[84] Korac J, Schaeffer V, Kovacevic I, Clement AM, Jungblut B, Behl C, et al. Ubiquitin-independent function of optineurin in autophagic clearance of protein aggregates. Journal of Cell Science. 2013;**126**:580-592. DOI: 10.1242/jcs.114926

[85] Ying H, Turturro S, Nguyen T, Shen X, Zelkha R, Johnson EC, et al. Induction of autophagy in rats upon overexpression of wild-type and mutant optineurin gene. BMC Cell Biology. 2015;**16**:14. DOI: 10.1186/s12860-015-0060-x

[86] Thul PJ, Akesson L, Wiking M, Mahdessian D, Geladaki A, Ait Blal H, et al. A subcellular map of the human proteome. Science. 2017;**356**. DOI: 10.1126/science.aal3321

[87] Uhlen M, Zhang C, Lee S, Sjostedt E, Fagerberg L, Bidkhori G, et al. A pathology atlas of the human cancer transcriptome. Science. 2017;**357**. DOI: 10.1126/science.aan2507

[88] Courtesy of Human Protein Atlas. *OPTN*. OPTN RNA expression overview 2016. Available from: https://ww.proteinatlas.org/ENSG00000123240-OPTN/pathology.

Chapter 4

Oral Pathology: Gene Expression in Odontogenic Cysts

Naida Hadziabdic and Amina Kurtovic-Kozaric

Abstract

Odontogenic cysts are a group of common pathological lesions of the jaw. Typically, they can be found randomly on X-rays as round benign lesions. However, some of them can behave aggressively with a tendency toward malignancy. Among odontogenic cysts with benign pathology, up to 60% of all jaw cysts are radicular cysts, which originate from root canal infection. Pathogenesis involves the interaction between osteoblasts, osteocytes, and osteoclasts as well as the expression of RANK-RANKL/OPG signaling system. Furthermore, collagenases (e.g., MMPs) are expressed in epithelial lining of the cyst. Among odontogenic cysts with potentially aggressive behavior, odontogenic keratocysts (OKCs) have a high rate of recurrence and very debatable treatment options; they can be associated with Gorlin syndrome. Keratocysts have developmental origin and show variability in their gene expression profiles. Their etiology is closely related to genetic factors, especially mutations in different members of Shh signaling pathway, including *PTCH* gene.

Keywords: odontogenic cysts, radicular cyst, keratocyst, RANK-RANKL/OPG, PTCH gene

1. Introduction

Odontogenic cysts are pathological cavities located in the jaw bones, filled with fluid surrounded by epithelial lining and fibrous connective tissue. Two most common odontogenic cysts with epithelial lining are inflammatory and developmental cysts (**Figure 1**). Among inflammatory cysts, the most characteristic one is the radicular cyst, which can be found in 60% of all odontogenic cysts. Radicular cyst originates from root canal infection and is of benign nature [1]. Cysts that originate from tissues involved in tooth development are called developmental cysts. Among them, the most interesting one is odontogenic keratocyst because of its unique and unexplored characteristics [2].

This chapter will deal with gene expression profiles of radicular cysts as the most common member of odontogenic cysts and keratocysts as the most debatable member of epithelial developmental cysts in order to uncover possible mechanism of pathogenesis of these two types of jaw cysts.

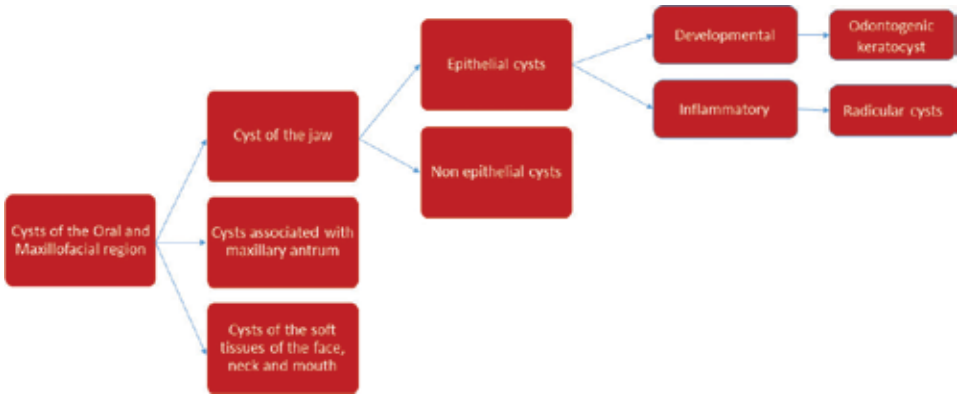

Figure 1.
Modified classification of cysts found in oral and maxillofacial regions [1].

2. Radicular cysts

2.1 Radicular and residual cysts

Radicular cysts are inflammatory cyst of jaw bones developing as a result of proliferation of the epithelium on the site of inflammation. They are the types of inflammatory lesions, which develop from epithelial residues of Hertwig lining in the process of periapical periodontitis after pulp necrosis. Cysts, which develop like this, are predominantly located around root apex and/or on the lateral side of the root, and in that case, they are called lateral radicular cysts. It often happens that radicular cyst remains in jaw bones after removal of the offending tooth, and in that case, it is called as residual cyst [1].

Radicular and residual cysts are by far the most common cystic lesions in the jaws, comprising 30% of all odontogenic cysts over a 30-year period according to results published by Jones et al. [3]. According to the age distribution of radicular cysts, most occur in patients who are in the third decade of life [3, 4]. There are many cases of patients who are between 40 and 60 years old, and after that, the number of cases decreases. Only few cases are seen in children even though dental caries commonly occurs in children.

Numerous studies have shown that radicular cysts occur more often in men than in women. The lower frequency in women may be because they take more care about mouth hygiene, especially the maxillary anterior incisors, where occurrence of cysts is very common. However, trauma to maxillary anterior teeth usually occurs to men [3].

Generally, radicular cysts show no symptoms and are discovered only after taking periapical radiographs of teeth with non-vital pulps. However, radicular cyst is the most common cause of swelling of the jaws, which patients often complain for. In the beginning, the enlargement is bony hard but as the cyst grows, the covering bone becomes very thin, and the swelling exhibits "egg shell crackling" or "springiness." The lesion becomes fluctuant only when the cyst has completely eroded the bone. In the mandible there is labial or buccal enlargement, rarely lingual, whereas in the maxilla, it is usually buccal or palatal [1].

Among other symptoms, pain and infection are very common in radicular cysts. Unless there is an infection, patients with radicular cyst usually feel no pain. But, it seems like there is no correlation between symptoms and infection. In the study done by Vier and Figueiredo, 21 out of 24 cysts were described as cavities which are filled with pus [5]. Authors did not correlate this finding and clinical symptoms, but it is unlikely that there is a relationship because most cysts are symptomless. In some cases patients complain of pain even though there is no evidence of infection and no

evidence of acute inflammation is present histologically after the cyst is removed [5]. On the other hand, some patients have histologically inflamed and clinically infected cysts, but they feel no pain [6]. A number of authors believe that there are cyst-prone individuals who are susceptible to developing radicular cysts because it often happens that more than one cyst is seen in one patient [7]. This can be supported and explained by the fact that occurrence of cysts is rare in relation to the large numbers of carious teeth which have dead pulps. Possibly, an immune mechanism inhibits formation of the cyst in most individuals, and patients who are prone to cyst development have a defect in suppression mechanism and immunological surveillance [8]. There is a possibility that some individuals are genetically susceptible to radicular cyst development.

Residual radicular cysts are cysts which remain after the non-vital offending tooth is removed. Studies have shown that they represent approximately 10% of all odontogenic cysts [9, 10]. In the study, Nair et al. considered that the type of cyst was important because of its persistence after treatment [11]. His findings confirmed the work of Simon [12] who stated that there were two types of radicular cyst: true radicular cyst, which contains a closed cavity lined by the epithelium, and the periapical pocket cyst in which the epithelium is attached to the margins of the apical foramen, so that the cyst lumen stays open to the root canal which is affected. Pocket cyst heals after treatment and tooth extraction, but true cyst, which is completely enclosed, is self-sustaining and persists even if there is no cause present [12].

2.2 Pathogenesis of radicular cysts

The pathogenesis of radicular cysts can be separated into three phases: the phase of initiation, the phase of cyst formation, and the phase of enlargement. Many studies have been done to investigate the mechanisms which are involved into these three phases.

2.2.1 The phase of initiation

Scientists agree that the epithelial cell rests of Malassez found in the periodontal ligament in periapical granulomas which are connected to necrotic and inflamed pulps are, in fact, the source of cyst linings [13]. The epithelial cell rests begin to multiply by inflammation which results from bacteria and debris discharged from the dead pulp. Bacterial endotoxins released from the necrotic pulp may be the key factor which initiates the inflammation and immune response and cause epithelial proliferation [1]. Meghji et al. investigated cyst fluids and grew cyst explants from radicular cysts and other cyst types and revealed that there are higher levels of endotoxins in radicular than in the other cysts [13].

Immunological studies are crucial in further understanding of granulomas and cysts, and they indicate that humoral and cell-mediated processes are involved in the pathogenesis of these lesions. Immunological studies have also demonstrated that inflammatory cytokines have an important role in the proliferation of epithelial cell rests. Stern et al. demonstrated that T lymphocyte infiltrates are involved in the development of periapical granulomas [1]. After a large number of studies, it became clear that endotoxins and inflammatory cytokines are highly involved in stimulation of the epithelial proliferation and they function as chemotactic and pro-inflammatory molecules as well [14, 15].

2.2.2 Phase of cyst formation

Pathogenesis of a radicular cyst is the next phase of cyst development, and it involves the process of cavity being lined by odontogenic epithelium. Two theories have been proposed, and both of them are reasonable and may function

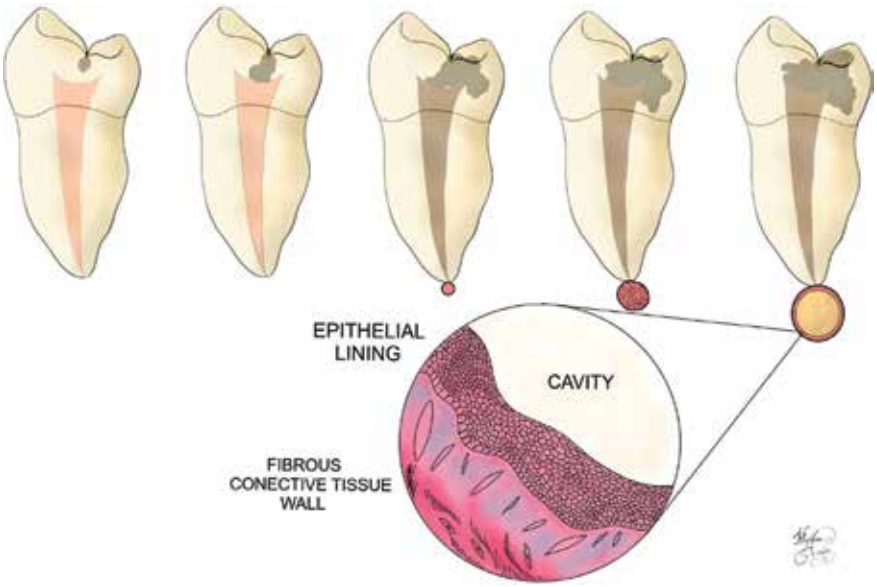

Figure 2.
Schematic representation of the development of radicular cyst. Starting from the left, superficial caries forms first, followed by medium and deep caries. Untreated deep caries leads to total pulp inflammation. Subsequently, necrosis of the pulp appears, which becomes infected. The resulting infection of the root canal initiates the epithelial remnants of Hertwig sheath (Malassez epithelial rests) to proliferate. Once the cells proliferate, the epithelial nest is formed. When the epithelial nest reaches the size of 1 cm, the center becomes necrotic leading to the formation of future cystic cavity, which becomes lined with the epithelium. For unknown reasons, this epithelium starts secreting fluid, which is called cystic fluid. These steps lead to the formation of radicular cyst, a round cavity filled with fluid and lined with the epithelium and fibrous connective tissue. This description of the radicular cyst development is the prevailing theory.

independently. The primarily accepted theory proposes that epithelial cells multiply and enclose the surface of connective tissue of an abscess cavity or cavity, which resulted from the breakdown of connective tissue by activity of proteolytic enzymes [16]. The secondary one, which is supported more, states that radicular cyst forms inside of the multiplying epithelial mass in periapical granuloma by cell death in the center (**Figure 2**) [16].

2.2.3 Growth and enlargement of the radicular cyst

The cyst enlargement is the final stage in radicular cyst pathogenesis. Toller's studies showed that osmosis contributes to the increase in the size of cysts [17]. Ward et al. proved this by simulating the growth of odontogenic cyst by mathematical modeling [18]. This modeling not only confirmed the results of Kubota et al. but also demonstrated that as the cyst became larger, cell proliferation played bigger part than osmotic pressure [19]. Harris and Toller suggested that cyst enlargement depends on epithelial proliferation which continues if inflammatory stimulus is present [20]. **Figure 3** shows large radicular cyst in the upper jaw.

2.3 Gene expression in radicular cysts

As described previously, among odontogenic cysts with benign pathology, up to 60% of all jaw cysts are radicular cysts, which originate from root canal infection caused by various microorganisms. The consequence of the radicular cyst development is the concomitant resorption of the surrounding bone tissues and periradicular periodontal ligament (PDL; [21]). Studies that have analyzed gene expression in radicular

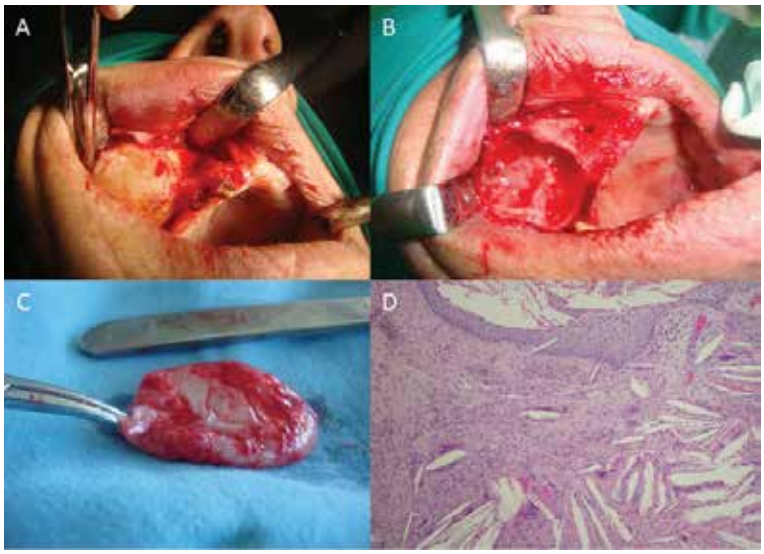

Figure 3.
Large radicular cyst located in the upper jaw on the right side. The appearance of the radicular cyst after mucoperiosteal flap was raised (A). Bone cavity after the cyst enucleation (B). The macroscopic view of the enucleated cyst (C). The histological morphology of radicular cyst with typical cholesterol crystals in the form of clefts, stained with H&E, 40× magnification (D).

cysts have mostly focused on genes that are involved in processes such as bone metabolism, inflammation, and tumorigenesis. Regarding bone metabolism, a gene known as receptor activator of nuclear factor-B ligand (RANKL) has been extensively studied because of its role in bone resorption around the tooth apex. This gene is part of a pathway that activates osteoclasts and is inhibited by a protein called osteoprotegerin (OPG). The role of RANK-RANKL-OPG signaling pathway in radicular cyst pathogenesis is further described in Section 2.3.2. Regarding genes involved in inflammatory processes, studies have analyzed expression of genes that code for chemokines and chemokine receptors that are involved in T helper type 1 (Th1) and Th2 responses that are characterized by the generation of interleukin-2 (IL-2), IL-12, and interferon-c (IFN-c) and by IL-4, IL-5, IL-6, IL-10, and IL-13, respectively [22, 23]. Regarding genes involved in tumorigenesis, *TP53* has been well analyzed in radicular cysts, where it shows low expression. Besides, *TP53*, *PCNA*, *FHIT*, and *Ki67* genes were analyzed and also showed insignificant changes compared to controls [24–26].

However, the most extensively studied genes in the pathogenesis of radicular cysts belong to the family of matrix metalloproteinases (MMPs). Their role in the cyst formation and development is described in Section 2.3.1.

2.3.1 Matrix metalloproteinases (MMPs)

The family of genes that are most commonly associated with the development of these lesions are matrix metalloproteinases (MMPs), which are metal-dependent endopeptidases that represent the major class of enzymes responsible for extracellular matrix degradation. It has been shown that MMPs have a crucial role in collagen degradation during periodontal tissue destruction [27–29]. Schematic representation of different classes of MMPs and an example of molecular structure is shown in **Figure 4**. Most commonly differentially expressed MMPs in oral diseases are presented in **Table 1**. For example, MMP-1, MMP-2, MMP-3, MMP-8, MMP-9, and MMP-13 play a role in the pathogenesis of periapical inflammatory lesions [27–31]. MMP-1 is a crucial enzyme in the initiation of osteoclastic bone

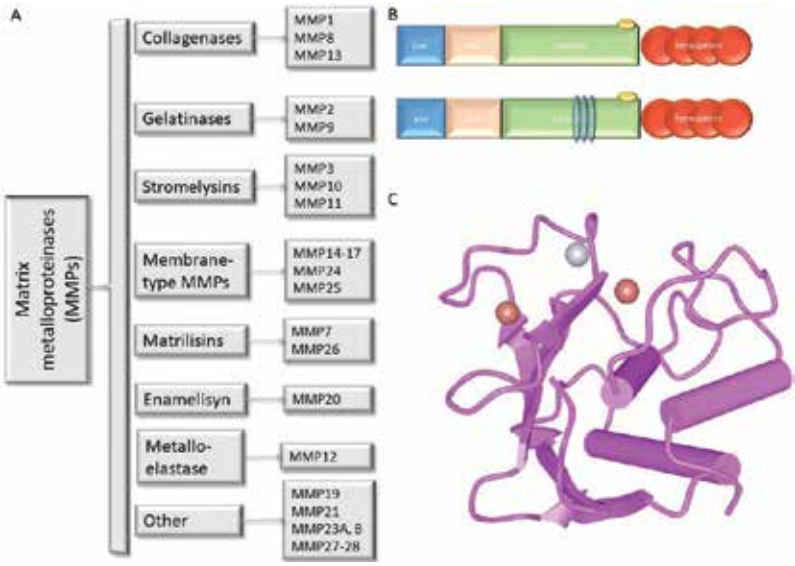

Figure 4.
Schematic classification, domain, and crystal structure of MMPs. Functional classification of MMPs is shown in A. MMPs are classified according to their function into collagenases, gelatinases, stromelysins, membrane-type MMPs, matrilysin, enamelin, metalloelastases, and others. Domain structure of collagenases (MMP-1, MMP-8, MMP-13) and gelatinases (MMP-2, MMP-9) is shown in B. The domain structure consists of signal peptide sequence (in blue color, labeled "pro"); prodomain that inhibits the catalytic domain, making it inaccessible to substrates (in pink color, labeled "pre"); catalytic domain that contains zinc atom (in green color, labeled "catalytic"); and hinge domain which links the catalytic and hemopexin domain (in red color, labeled "hemopexin"). Gelatinases also contain fibronectin-like domain repeats which aid in substrate binding (in blue color). Crystal structure of MMP-1 is shown in C.

resorption [32], because it degrades the collagen of the unmineralized layer found on the surface of the bone. After collagen degradation, the collagen fragments are produced, which function in osteoclast activation [33]. Studies in rats have shown that MMP-1 expression is elevated in the active phase of periapical lesion development [34]. MMP-1 also causes expansion of radicular cysts in humans [35].

MMP-2 and MMP-9 function in the degradation of the extracellular matrix (ECM), particularly during the active stages of lesion development [30], and their expression is increased in those lesions compared to control healthy tissue (gingiva, PDL, or oral mucosa).

MMPs play an important role in normal cellular processes such as tissue growth, bone resorption, and remodeling (wound healing and angiogenesis) [36]. MMPs are secreted in their proenzyme state and require extracellular activation. They are regulated by endogenously secreted inhibitors, called TIMPs (the tissue inhibitor of metalloproteinases). MMPs cleave native, nondenatured collagens with long uninterrupted triple helices and can function as collagenases (**Table 1**) [37]. Thus, the molecular basis of MMP function involves proteolytic cleavage of different substrates and subsequent activation of transforming growth factor-beta (TGF-beta), insulin-like growth factors (IGF), vascular endothelial growth factors (VEGF), and RANKL pathway.

Normal tissues generally show low expression of MMPs. However, during pathological states that require destruction of extracellular matrix, expression of MMPs can be drastically increased [29]. In normal and healthy tissues, components of extracellular matrix are in constant balance between degradation and protein synthesis. It has been shown that elevated MMP levels correlate with nonhealing [38, 39].

In pathological state such as apical periodontitis, MMP expression and secretion are increased, suggesting the direct role of MMPs in tissue remodeling and

Name of enzyme	Class	Substrate	Expression in normal tissue and oral disease
MMP-1 (collagenase-1, interstitial collagenase)	Collagenase	Collagen I, II, III, VII, VIII, and X and gelatin	Bone resorption in periapical lesions and human periapical lesions degrades nonmineralized extracellular matrix and stimulates osteoclastogenesis by collagen degradation on bone surface, expression in cystic wall and cystic fluid, pulpitis, squamocellular carcinoma, and normal tissues (stomach and gallbladder)
MMP-8 (collagenase-2, neutrophil collagenase)	Collagenase	Collagen I, II, III, V, VII, VIII, and X and gelatin	Chronic pulp inflammation, human periapical lesions, pulpitis, caries, and normal tissues (bone marrow and spleen)
MMP-13 (collagenase-3)	Collagenase	Collagen I, II, III, IV, V, VII, IX, and X	Chronic pulp inflammation, human periapical lesions, pulpitis, and normal tissue (vagina, lung)
MMP-9 (gelatinase B)	Gelatinase	Collagen III, IV, V, VII, X, and XI	Periodontitis, caries, oral squamocellular carcinoma metastasis, invasiveness and shorter survival, and human periapical lesions
MMP-2 (gelatinase A)	Gelatinase	Collagen I, II, III, IV, V, VII, X, and XI	Periodontitis, caries, oral squamocellular carcinoma metastasis, invasiveness and shorter survival, collagen type IV degradation, ameloblastoma, human periapical lesions, and normal tissue (gallbladder, uterine, endometrium, cervix)
MMP-3		Collagen III, IV, V, VII, IX, X, and XI	Oral squamocellular carcinoma, human periapical lesions, and normal tissue (endometrium, salivary gland)

Table 1.
List of MMPs most commonly associated with and expressed in oral disease and healthy tissues.

destruction during lesion development [40–42]. Regulation of their expression is primarily controlled at the transcriptional level even though regulatory mechanisms have still not been fully elucidated [27, 43, 44]. Some of the reasons include the influence of promoter polymorphisms, epigenetic mechanisms, and posttranscriptional processes [43, 45].

Promoter polymorphisms have been detected in MMP promoter regions, suggesting that the changes in MMP expression can predispose individuals to develop periapical inflammatory lesions. Similarly, these MMP promoter polymorphisms can lead to progression of disease [42, 46]. For example, Menezes-Silva et al. investigated genetic predisposition to periapical disease by testing 16 SNP polymorphisms in *MMP2*, *MMP3*, *MMP9*, *MMP13*, *MMP14*, and *TIMP2* genes [42]. They found that polymorphisms in *MMP2* and *MMP3* genes are associated with the development of periapical lesions, suggesting that these markers could assist in prevention and healing process [42].

Besides MMP promoter polymorphisms, another regulatory mechanism for MMP expression has been found at the epigenetic level, where methylation of MMP promoters can lead to gene inactivation and subsequent decrease in transcription [47]. Campos et al. have studied the methylation state of *MMP-2* and *MMP-9* in periapical inflammatory lesions [48]. Their results show that *MMP-2* gene was partially methylated in periapical granuloma, radicular cysts, and normal oral mucosa,

and subsequent association between methylation status and gene expression was not possible [48]. Regarding *MMP-9*, the study found that this gene was more unmethylated in periapical granulomas and radicular cysts than in healthy mucosa, which implies that *MMP-9* mRNA expression is increased and may be epigenetically controlled [48]. Effects of DNA methylation on MMP genes can contribute to individual susceptibility to the development of periapical granuloma and radicular cysts as periapical inflammatory lesions. Moreover, it may also play a role in the patient's response to therapy [48].

Besides MMP studies, other genes such as *FOXP3* have shown interesting results in periapical granulomas and radicular cysts. It has been shown that the *FOXP3* gene promoter methylation was inversely correlated with *FOXP3* transcript levels, suggesting that *FOXP3* may be crucial in determining periapical lesion development [49].

In conclusion, here we presented studies that suggest that genetic predisposition to frequent development of periapical inflammatory lesions could be caused by the presence of polymorphisms in MMP gene promoters or by epigenetic mechanisms such as differential methylation status of MMP genes.

2.3.2 RANKL expression

RANKL-RANK-OPG system discovery has changed our understanding of bone biology (**Figure 5**). This is a signaling system that is crucial for skeletal homeostasis because it maintains the balance between bone resorption by osteoclasts and bone formation by osteoblasts. RANKL is a ligand for its receptor RANK which can be found on osteoclast progenitor cells. After binding, RANK-RANKL system activates NF-κB pathway and upregulation of NFATc1 protein, which is a master regulator of cytokines important for osteoclastogenesis. The disruption of RANKL-RANK system leads to the inhibition of bone resorption. RANKL belongs to the TNF superfamily of proteins. The name RANKL stands for receptor activator of nuclear factor-κB ligand, also known as osteoprotegerin ligand (OPGL), osteoclast differentiation factor (ODF), TNF-related activation-induced cytokine (TRANCE), and TNF ligand superfamily member 11 (TNFSF11) [50, 51]. It is a homotrimeric protein that is membrane bound on osteoblasts and activated T cells. It can also be

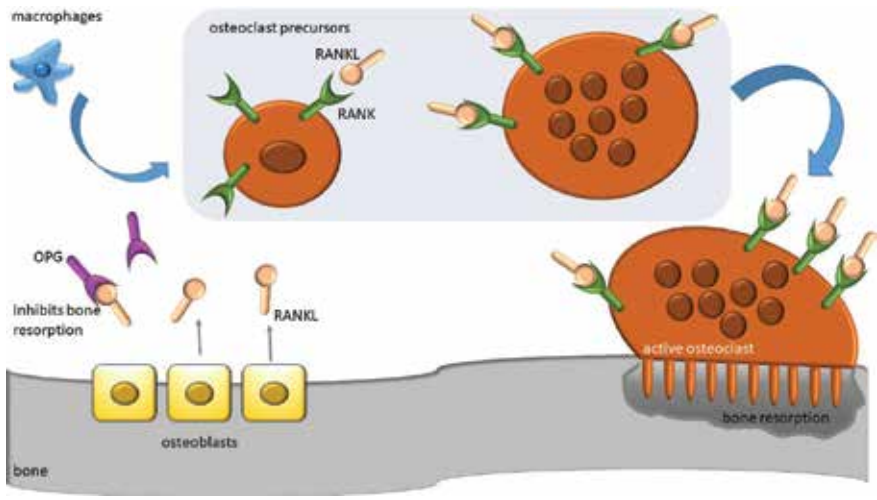

Figure 5.
RANKL-RANK-OPG system as regulators of bone resorption. Bone metabolism is a dynamic process that balances bone formation and bone resorption. RANKL (ligand) is secreted by osteoblasts and binds to the RANK receptor on osteoclast precursors and activated osteoclasts, which in turn result in bone resorption. The inhibition of this process is mediated by OPG.

secreted by T cells. RANKL is proteolytically cleaved by MMP-3 or MMP-7. Besides T cells and osteoclasts, RANKL expression can be seen in lymph nodes, thymus, mammary glands, spleen, and bone marrow. In tumor cells, RANKL is associated with migration and bone metastasis. Thus, RANKL is a key regulator of bone metabolism, specifically a regulator of osteoclastogenesis and osteoclastoactivation in a normal and pathological states [50, 51]. RANK is a receptor for RANKL and also a member of TNF superfamily. It is a homotrimeric transmembrane protein. Its expression is generally less than RANKL, and high expression is seen in mammary glands and cancer cells. OPG is an inhibitor of RANK-RANKL system. It stands for osteoprotegerin, also known as osteoclastogenesis inhibitory factor (OCIF) or tumor necrosis factor receptor superfamily member 11B, is a cytokine receptor of the TNF receptor superfamily encoded by the *TNFRSF11B* gene. It is a 380 amino acid glycoprotein that is found in soluble form as either monomer or dimer. The OPG dimer is crucial for RANK-RANKL inhibition because OPG dimer increases the affinity of OPG for RANKL.

RANKL affects the development of the periapical inflammatory lesions by activating osteoclasts, thus inducing pathological bone resorption [52]. RANKL protein expression was first shown in radicular cysts [53], which colocalized with osteoclasts. Subsequent studies analyzed transcript levels in inflammatory granulomas, which were increased compared to healthy PDL tissue [52]. RANKL expression was found in infiltrating leucocytes, specifically monocytes and dendritic cells, which were shown to be the main cells that secrete this protein. Another study compared RANKL and OPG levels in apical granulomas and radicular cysts [54], finding that both OPG and RANKL expressions were higher in granulomas than in cysts, but their ratio was comparable in these two types of periapical inflammatory lesions. Fukada and colleagues found that RANKL transcript levels were significantly higher in granuloma than in radicular cysts [55]. At the protein level, no difference was observed in RANKL and OPG levels in a study conducted by Fan and colleagues [56]. In a study on endodontically involved disease, RANKL expression was higher in lesions with more intense inflammation, but the ratio RANKL/OPG in relation to inflammation was not increased [57].

3. Odontogenic keratocysts

Odontogenic keratocysts (OKCs) represent a rare form of odontogenic cysts which originate from dental lamina remnants or eventually from the basal layer of upper and lower jaw oral epithelium before the odontogenesis ended. Since it was first described in 1876, this form of cysts grabs scientific attention mostly because of its developmental variabilities, histological appearance, and genetic basis [2, 58]. In the past few years, the World Health Organization (WHO) made an attempt to create more appropriate classification of these cysts. Recently, they were considered as keratocystic odontogenic tumors for their aggressive behavior, high mitotic rate, and association with genetic and chromosomal abnormalities. The newest WHO classification reclassifies them again as odontogenic keratocysts because *PTCH1* gene mutations were detected, similarly to other developmental cysts such as dentigerous cyst [59–61]. Despite many classifications, pathologists and surgeons face difficulties in the establishment of proper diagnosis. This is because keratocysts cannot be clinically and radiographically distinguished from other odontogenic cysts. Moreover, it is still debatable what the optimal therapeutic approach is in the treatment of keratocysts in order to prevent recidive, which is a characteristic of this disease [2].

Odontogenic keratocyst is histologically characterized with stratified squamous epithelium, which is five to eight layers thick with palisaded hyperchromatic basal cell layer and "corrugated" parakeratotic epithelial cells on luminal surface [2, 62, 63].

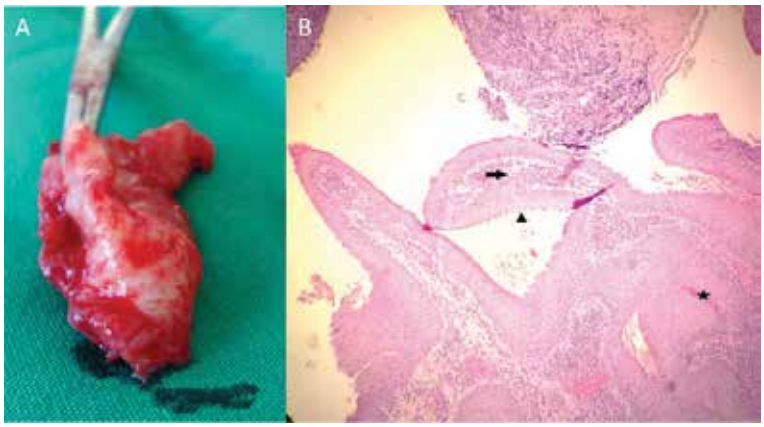

Figure 6.
Macroscopic view (A) and histological morphology (B) of odontogenic keratocyst. Inflamed fibrous wall lined with nonspecific stratified squamous epithelium (arrow) with focally corrugated surface (arrow head). The presence of satellite cysts and solid islands in the cystic wall is marked with a star.

Parakeratinization of the surface layer is one of the histological features that predicts recurrence together with the higher level of cell proliferative activity in the epithelium, binding in the basal layer of the epithelium, supraepithelial split of the epithelial lining, presence of daughter cells or also called satellite cells (**Figure 6**) [64].

Considering clinical and histological features together with unclear etiology and therapeutic strategy odontogenic keratocyst still remains debatable. Controversies in nature of these lesions are reflection of limited knowledge about their origin. Recent molecular and genetic discoveries try to elucidate these pathological entities and thus approach closer to better therapy modalities.

3.1 Molecular basis of odontogenic keratocysts

Odontogenic keratocysts are locally aggressive and have a tendency to recur over time. Keratocysts can happen sporadically or in conjunction with Gorlin-Goltz syndrome (Gorlin syndrome, OMIM 109400) [62, 64]. Gorlin syndrome is an autosomally inherited disease with a germline mutation in *PTCH* gene, located on chromosome 9q22. The incidence is 1:60,000 newborns. Patients with this syndrome present with a spectrum of developmental abnormalities which affect skin, nervous system, eyes, endocrine system, and bones. Skin abnormalities include basal cell carcinoma, benign dermoid cysts and tumors, palmar and plantar keratosis, and dermal calcinosis. Dental and bone abnormalities include multiple keratocysts, mild mandibular prognathism, kyphoscoliosis, and other vertebral defects like bifurcation of ribs, spina bifida, and others. Eye abnormalities include hypertelorism, wide nasal bridge, congenital blindness and strabismus. Neurological abnormalities include mental retardation, calcification of dura mater, and others. For the diagnosis of Gorlin syndrome, two major or one major and two minor criteria are needed [62, 64, 65].

The major criteria are as follows: multiple basal cell carcinoma or one tumor diagnosed before 20 years of age, histologically proven odontogenic keratocyst, three or more palmar or plantar pits, bilamellar calcification of the falx cerebri, fused or markedly splayed ribs, and first-degree relatives with Gorlin syndrome. The minor criteria include macrocephaly; congenital cleft lip or palate, frontal bossing, coarse face, or hypertelorism; other skeletal abnormalities; radiological deformities like bridging of the sella turcica, vertebral abnormalities, and ovarian fibroma; and medulloblastoma.

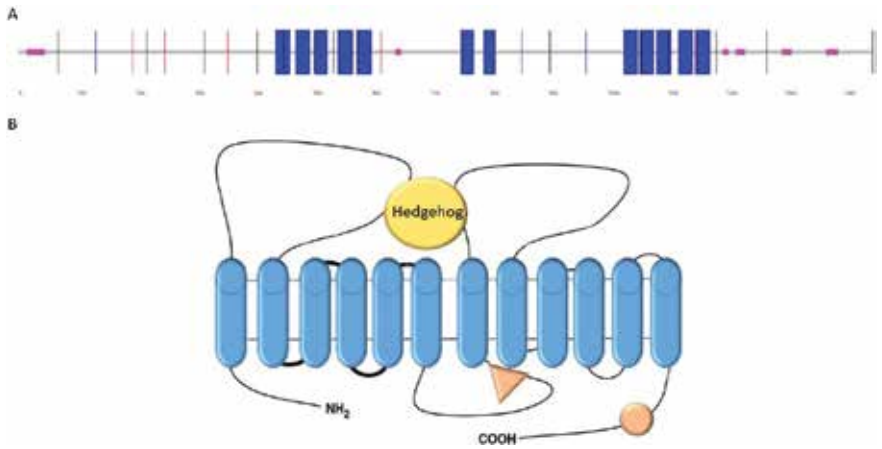

Figure 7.
Schematic representation of PTCH molecular structure. (A) Domain structure of PTCH protein with 12 transmembrane domains (blue) and low complexity regions (pink). It contains 23 coding exons spanning 74 kb and encoding a 1447 amino acid receptor glycoprotein. (B) PTCH is a membrane-bound receptor for hedgehog protein (yellow). PTCH1 gene produces patched protein, which is a homolog of Drosophila segment polarity gene. This receptor is a part of the hedgehog signaling cascade, which include downstream proteins such as smoothened (Smo) and disheveled (Dsh).

The appearance of keratocysts in Gorlin syndrome patients can be seen in up to 92% of all patients. Even though it has been shown that the syndromic form of keratocysts contains much higher numbers of proliferating epithelial cells and satellite cysts within the cystic wall, pathological examination cannot differentiate between syndromic and sporadic forms of this disease. Similarly, it has been shown that the syndromic form has higher rate of recurrence than the sporadic form. However, the syndromic form has shown more aggressive behavior than the sporadic form. Thus, radiographic appearance of multiple keratocyst formations should raise suspicion of a possible Gorlin syndrome [63].

PTCH1 gene produces Patched protein, which is a homolog of *Drosophila* segment polarity gene. It contains 23 coding exons spanning 74 kb and encoding a 1447 amino acid receptor glycoprotein. This receptor is a part of the Hedgehog signaling cascade, which includes downstream proteins such as Smoothened (Smo) and Disheveled (Dsh). Molecular structure including schematic domain representation is shown in **Figure 7**.

PTCH is a tumor suppressor and as such requires both gene copies to be mutated in tumors. Thus, patients with Gorlin syndrome already have one inherited mutation and can acquire the second mutation easier than patients with sporadic form. It has been shown that mutations in this gene often occur as LOH (loss of heterozygosity), meaning that both copies are mutated through different genetic mechanisms such as deletions, mutations, gene silencing, and others. After PTCH mutation, keratocysts can acquire additional genetic alterations which accelerate tumor development. When LOH was analyzed in sporadic form, several additional tumor suppressor genes were affected such as *TP53*. The association between these mutations and appearance of satellite microcysts in the cystic wall is particularly interesting [64].

3.2 Gene expression in odontogenic keratocysts

The hallmark of odontogenic keratocysts is mutation in *PTCH* gene, which is a receptor in Sonic hedgehog signaling (Shh) pathway. This pathway is important for proper differentiation of embryonic cells, and mutations in this pathway lead

to abnormal cellular proliferation and differentiation, which is strongly associated with the development of certain tumors [66, 67]. This section will describe genes that are involved in pathogenesis of odontogenic keratocysts through Sonic hedgehog pathway (Shh). The members of this pathway include secreted SHH ligand, which binds to its receptor *PTCH*, a 12-pass transmembrane protein, as described previously (**Figure 7**). In the absence of SHH ligand, PTCH inhibits a transmembrane protein called Smoothened (SMO).

In odontogenic keratocysts, *PTCH* mutations include LOH, deletions, point mutations, and others. The consequence of *PTCH* mutations are the constitutive activation of Shh pathway. When *PTCH* is mutated and inactivated, it is expected that SMO would be disinhibited and thus activated. This is exactly what is found in odontogenic keratocysts, where Smo overexpression has been shown by transcriptional and immunohistochemical studies [67–69]. Besides overexpression of SMO and PTCH, other downstream genes such as *GLI1*, *CCND1*, and *BCL-2* have been shown to be overexpressed in odontogenic keratocysts, indicating that these SHH pathway genes contribute to the development of these lesions.

Another downstream target of Shh signaling is a transcription factor SOX-2 [70] that is expressed in progenitor cells in epithelial tissues. SOX-2 expression is associated with elements of tooth development, especially in the region of the third molars in the lower jaw, which is the place where OKCs are usually located [71, 72].

Besides studies on Shh pathway, a list of genes that are upregulated in odontogenic keratocysts is shown in **Table 2**. Bioinformatic analysis has shown that other genes, such as *TP53* and *PCNA*, appear as the leaders and initiators of gene expression that is important for the development of odontogenic keratocysts [61, 73, 74]. Their analysis has shown that genes related to cell cycle and apoptosis are often dysregulated in these cysts, implying the recurrence of these cysts. Studies have shown that *TP53*, *PCNA*, p63, and Ki-67 expression is higher in keratocysts than in other types of odontogenic cysts. *TP53* is a tumor suppressor gene with several different functions in the cell including apoptosis, cell cycle arrest, and DNA repair.

The second gene that is found to be associated with odontogenic keratocysts is proliferating cell nuclear antigen (*PCNA*), which encodes a protein located in the nucleus and associated with DNA polymerase delta. It acts as a homotrimer and is

Gene name	Method of detection	References
PCNA	IHC	[74]
CCND1	IHC and real-time PCR	[67]
IL-6	IHC	[75]
VEGFA	IHC	[76]
BCL2	IHC, real-time PCR, Western blot	[77]
FHIT	IHC	[25]
GLI1	IHC	[78]
TP63	IHC	[76]
KRT6B	IHC	[79]
TP53	Bioinformatic analysis	[73]
SMO	IHC and transcriptional analysis	[67–69]

IHC, immunohistochemistry; PCR, polymerase chain reaction.

Table 2.
Genes that show high expression in odontogenic keratocysts, detected either at the RNA or protein level (modified from [63]).

implicated in the leading strand synthesis during DNA replication. DNA damage induces ubiquitination of this protein through RAD6-dependent DNA repair pathway. Expression of both *TP53* and *PCNA* in different odontogenic lesions was higher in the suprabasal layer of keratocysts than in radicular cyst. Overexpression of *PCNA* in the suprabasal layer implies its neoplastic nature and a tendency toward recurrence.

4. Conclusion

This chapter summarizes gene expression profiles of radicular cysts as the most common member of odontogenic cysts and keratocysts, a specific entity of epithelial developmental cysts, in order to uncover possible mechanism of pathogenesis that would help in the timely diagnosis and discovery of novel therapeutic options for these two types of jaw cysts. Pathogenesis of radicular cysts is associated with differential expression of genes involved in bone metabolism (RANK-RANKL-OPG pathway) and inflammation (chemokines and their receptors). However, the most extensively studied genes in the pathogenesis of radicular cysts belong to the family of matrix metalloproteinases (MMPs), which show increased expression.

Specific entities of odontogenic cysts are odontogenic keratocysts, which are prone to recidive. This trait of keratocysts to recur makes them similar to tumors, which can be also seen in their gene expression profiles. The hallmark of odontogenic keratocysts is mutation in *PTCH* gene, which is a receptor in Sonic hedgehog signaling (Shh) pathway. Mutations in PTCH gene lead to the constitutive activation of this pathway. Besides overexpression of PTCH, other downstream genes such as *SMO*, *GLI1*, *CCND1*, and *BCL-2* have been shown to be overexpressed in odontogenic keratocysts, indicating that these SHH pathway genes contribute to the development of these lesions.

Acknowledgements

Authors would like to thank Armin Klančević, DMD, for his contribution by drawing **Figure 2**.

Conflict of interest

Authors declare no conflict of interest.

Author details

Naida Hadziabdic[1*] and Amina Kurtovic-Kozaric[2]

1 Department of Oral Surgery, Faculty of Dental Medicine, University of Sarajevo, Sarajevo, Bosnia and Herzegovina

2 Department of Pathology, Clinical Center of the University of Sarajevo, Sarajevo, Bosnia and Herzegovina

*Address all correspondence to: nsulejma@yahoo.com

IntechOpen

References

[1] Shear M, Speight P. Cysts of the Oral and Maxillofacial Regions. John Wiley & Sons; 15 Apr 2008

[2] Nayak M, Singh A, Singhvi A, Sharma R. Odontogenic keratocyst: What is in the name? Journal of Natural Science, Biology and Medicine. 2013;**4**:282. DOI: 10.4103/0976-9668.116968

[3] Jones AV, Craig GT, Franklin CD. Range and demographics of odontogenic cysts diagnosed in a UK population over a 30-year period. Journal of Oral Pathology & Medicine. 2006;**35**:500-507. DOI: 10.1111/j.1600-0714.2006.00455.x

[4] Donath K. Odontogenic and nonodontogenic jaw cysts. Deutsche zahnärztliche Zeitschrift. 1985;**40**: 502-509 passim

[5] Vier FV, Figueiredo JAP. Prevalence of different periapical lesions associated with human teeth and their correlation with the presence and extension of apical external root resorption. International Endodontic Journal. 2002;**35**:710-719

[6] Shear M. Clinical statistics of dental cysts. The Journal of the Dental Association of South Africa. 1961;**16**:360-364

[7] Stoelinga PJ, Bronkhorst FB. The incidence, multiple presentation and recurrence of aggressive cysts of the jaws. Journal of Cranio-Maxillo-Facial Surgery. 1988;**16**:184-195. DOI: 10.1016/S1010-5182(88)80044-1

[8] Toller P, Lancet EH-T. Immunoglobulins and immunoglobulin-containing cells in cysts of the jaws. Lancet. 1969;**294**:178-181

[9] High AS, Hirschmann PN. Age changes in residual radicular cysts. Journal of Oral Pathology. 1986;**15**:524-528

[10] High AS, Hirschmann PN. Symptomatic residual radicular cysts. Journal of Oral Pathology. 1988;**17**:70-72

[11] Ramachandran Nair PN, Pajarola G, Schroeder HE. Types and incidence of human periapical lesions obtained with extracted teeth. Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontics. 1996;**81**:93-102

[12] Simon JHS. Incidence of periapical cysts in relation to the root canal. Journal of Endodontia. 1980;**6**:845-848. DOI: 10.1016/S0099-2399(80)80039-2

[13] Meghji S, Qureshi W, Henderson B, Harris M. The role of endotoxin and cytokines in the pathogenesis of odontogenic cysts. Archives of Oral Biology. 1996;**41**:523-531

[14] Stern MH, Dreizen S, Mackler BF, Levy BM. Isolation and characterization of inflammatory cells from the human periapical granuloma. Journal of Dental Research. 1982;**61**:1408-1412. DOI: 10.1177/00220345820610120601

[15] Browne RM. The pathogenesis of odontogenic cysts: A review. Journal of Oral Pathology. 1975;**4**:31-46

[16] Nair PNR. Pathogenesis of apical periodontitis and the causes of endodontic failures. Critical Reviews in Oral Biology and Medicine. 2004;**15**:348-381

[17] Summers L. The incidence of epithelium in periapical granulomas and the mechanism of cavitation in apical dental cysts in man. Archives of Oral Biology. 1974;**19**:1177-1180

[18] Ward JP, Magar V, Franks SJ, Landini G. A mathematical model of the

dynamics of odontogenic cyst growth. Analytical and Quantitative Cytology and Histology. 2004;**26**:39-46

[19] Kubota Y, Yamashiro T, Oka S, Ninomiya T, Ogata S, Shirasuna K. Relation between size of odontogenic jaw cysts and the pressure of fluid within. The British Journal of Oral & Maxillofacial Surgery. 2004;**42**:391-395. DOI: 10.1016/j.bjoms.2004.02.032

[20] Harris M, Toller P. The pathogenesis of dental cysts. British Medical Bulletin. 1975;**31**:159-163

[21] Kawashima N, Stashenko P. Expression of bone-resorptive and regulatory cytokines in murine periapical inflammation. Archives of Oral Biology. 1999;**44**:55-66. DOI: 10.1016/S0003-9969(98)00094-6

[22] Jankovic D, Liu Z, Gause WC. Th1- and Th2-cell commitment during infectious disease: Asymmetry in divergent pathways. Trends in Immunology. 2001;**22**:450-457

[23] Silva TAA, Garlet GPP, Fukada SYY, Silva JSS, Cunha FQQ. Chemokines in oral inflammatory diseases: Apical periodontitis and periodontal disease. Journal of Dental Research. 2007;**86**:306-319. DOI: 10.1177/154405910708600403

[24] De Oliveira Gomes C, Drummond SN, Jham BC, Abdo EN, Mesquita RA. A survey of 460 supernumerary teeth in Brazilian children and adolescents. International Journal of Paediatric Dentistry. 2008;**18**:98-106. DOI: 10.1111/j.1365-263X.2007.00862.x

[25] Malčić A, Jukić S, Anić I, Pavelić B, Kapitanović S, Krušlin B, et al. Alterations of FHIT and P53 genes in keratocystic odontogenic tumor, dentigerous and radicular cyst. Journal of Oral Pathology & Medicine. 2008;**37**:294-301. DOI: 10.1111/j.1600-0714.2007.00622.x

[26] Gadbail A, Chaudhary M, Patil S, Gawande M. Actual proliferating index and p53 protein expression as prognostic marker in odontogenic cysts. Oral Diseases. 2009;**15**:490-498. DOI: 10.1111/j.1601-0825.2009.01590.x

[27] Chakraborti S, Mandal M, Das S, Mandal A, Chakraborti T. Regulation of matrix metalloproteinases: An overview. Molecular Cell. 2003;**253**:269-285

[28] Paula-Silva FWG, da Silva LAB, Kapila YL. Matrix metalloproteinase expression in teeth with apical periodontitis is differentially modulated by the modality of root canal treatment. Journal of Endodontia. 2010;**36**:231-237. DOI: 10.1016/j.joen.2009.10.030

[29] Hadziabdic N, Kurtovic-Kozaric A, Pojskic N, Sulejmanagic N, Todorovic L. Gene-expression analysis of matrix metalloproteinases 1 and 2 and their tissue inhibitors in chronic periapical inflammatory lesions. Journal of Oral Pathology & Medicine. 2016;**45**:224-30. DOI: 10.1111/jop.12347

[30] Corotti MV, Zambuzzi WF, Paiva KBS, Menezes R, Pinto LC, Lara VS, et al. Immunolocalization of matrix metalloproteinases-2 and -9 during apical periodontitis development. Archives of Oral Biology. 2009;**54**:764-771. DOI: 10.1016/j.archoralbio.2009.04.013

[31] Matsui H, Yamasaki M, Nakata K, Amano K, Nakamura H. Expression of MMP-8 and MMP-13 in the development of periradicular lesions. International Endodontic Journal. 2011;**44**:739-745. DOI: 10.1111/j.1365-2591.2011.01880.x

[32] Sasaki K, Takagi M, Konttinen YT, Sasaki A, Tamaki Y, Ogino T, et al. Upregulation of matrix metalloproteinase (MMP)-1 and its activator MMP-3 of human osteoblast by uniaxial cyclic stimulation. The Journal of Biomedical Materials

Research Part B: Applied Biomaterials. 2007;**80B**:491-498. DOI: 10.1002/jbm.b.30622

[33] Holliday LS, Welgus HG, Fliszar CJ, Veith GM, Jeffrey JJ, Gluck SL. Initiation of osteoclast bone resorption by interstitial collagenase. The Journal of Biological Chemistry. 1997;**272**: 22053-22058. DOI: 10.1074/jbc.272.35.22053

[34] Lin SK, Chiang CP, Hong CY, Lin CP, Lan WH, Hsieh CC, et al. Immunolocalization of interstitial collagenase (MMP-1) and tissue inhibitor of metalloproteinases-1 (TIMP-1) in radicular cysts. Journal of Oral Pathology & Medicine. 1997;**26**:458-463. DOI: 10.1111/j.1600-0714.1997.tb00016.x

[35] Lin SK, Kok SH, Kuo MYP, Wang TJ, Wang JT, Yeh FTC, et al. Sequential expressions of MMP-1, TIMP-1, IL-6, and COX-2 genes in induced periapical lesions in rats. European Journal of Oral Sciences. 2002;**110**:246-253. DOI: 10.1034/j.1600-0447.2002.11227.x

[36] Delaissé JM, Engsig MT, Everts V, Del Carmen Ovejero M, Ferreras M, Lund L, et al. Proteinases in bone resorption: Obvious and less obvious roles. Clinica Chimica Acta. 2000;**291**:223-234. DOI: 10.1016/S0009-8981(99)00230-2

[37] Guise TA. Breaking down bone: New insight into site-specific mechanisms of breast cancer osteolysis mediated by metalloproteinases. Genes & Development. 2009;**23**:2117-2123. DOI: 10.1101/gad.1854909

[38] Beidler SK, Douillet CD, Berndt DF, Keagy BA, Rich PB, Marston WA. Multiplexed analysis of matrix metalloproteinases in leg ulcer tissue of patients with chronic venous insufficiency before and after compression therapy. Wound Repair and Regeneration. 2008;**16**:642-648. DOI: 10.1111/j.1524-475X.2008.00415.x

[39] Liu Y, Min D, Bolton T, Nube V, Twigg SM, Yue DK, et al. Increased matrix metalloproteinase-9 predicts poor wound healing in diabetic. Diabetes Care. 2009;**32**:117-119. DOI: 10.2337/dc08-0763

[40] Carneiro E, Menezes R, Garlet GP, Garcia RB, Monteiro Bramante C, Figueira R, et al. Expression analysis of matrix metalloproteinase-9 in epithelialized and nonepithelialized apical periodontitis lesions. Oral Surgery, Oral Med Oral Pathol Oral Radiol Endodontology. 2009;**107**:127-132. DOI: 10.1016/j.tripleo.2008.07.030

[41] Letra A, Ghaneh G, Zhao M, Ray H, Francisconi CF, Garlet GP, et al. MMP-7 and TIMP-1, new targets in predicting poor wound healing in apical periodontitis. Journal of Endodontia. 2013;**39**:1141-1146. DOI: 10.1016/j.joen.2013.06.015

[42] Menezes-Silva R, Khaliq S, Deeley K, Letra A, Vieira AR. Genetic susceptibility to periapical disease: Conditional contribution of MMP2 and MMP3 genes to the development of periapical lesions and healing response. Journal of Endodontia. 2012;**38**: 604-607. DOI: 10.1016/j.joen.2012.02.009

[43] Yan C, Boyd DD. Regulation of matrix metalloproteinase gene expression. Journal of Cellular Physiology. 2007;**211**:19-26

[44] Ye S. Polymorphism in matrix metalloproteinase gene promoters: Implication in regulation of gene expression and susceptibility of various diseases. Matrix Biology. 2000;**19**:623-629. DOI: 10.1016/S0945-053X(00)00102-5

[45] Chou J, Shahi P, Werb Z. MicroRNA-mediated regulation of the tumor microenvironment. Cell Cycle. 2013;**12**:3262-3271. DOI: 10.4161/cc.26087

[46] Dill A, Letra A, Chaves De Souza L, Yadlapati M, Biguetti CC, Garlet GP, et al. Analysis of multiple cytokine polymorphisms in individuals with untreated deep carious lesions reveals il1b (RS1143643) as a susceptibility factor for periapical lesion development. Journal of Endodontia. 2015;**41**:197-200. DOI: 10.1016/j.joen.2014.10.016

[47] Esteller M. CpG island hypermethylation and tumor suppressor genes: A booming present, a brighter future. Oncogene. 2002;**21**:5427-5440. DOI: 10.1038/sj.onc.1205600

[48] Campos K, Gomes CC, Farias LC, Silva RM, Letra A, Gomez RS. DNA methylation of MMP9 is associated with high levels of MMP-9 messenger RNA in periapical inflammatory lesions. Journal of Endodontia. 2016;**42**:127-130. DOI: 10.1016/j.joen.2015.10.002

[49] Campos K, Franscisconi CF, Okehie V, De Souza LC, Trombone APF, Letra A, et al. FOXP3 DNA methylation levels as a potential biomarker in the development of periapical lesions. Journal of Endodontia. 2015;**41**:212-218. DOI: 10.1016/j.joen.2014.10.003

[50] Lacey DL, Timms E, Tan HL, Kelley MJ, Dunstan CR, Burgess T, et al. Osteoprotegerin ligand is a cytokine that regulates osteoclast differentiation and activation. Cell. 1998;**93**:165-176. DOI: 10.1016/S0092-8674(00)81569-X

[51] Thomson AW, Lotze MT, editors. The Cytokine Handbook, Two-Volume Set. 4 th ed. Elsevier; 2003. 1572 p

[52] Vernal R, Dezerega A, Dutzan N, Chaparro A, León R, Chandía S, et al. RANKL in human periapical granuloma: Possible involvement in periapical bone destruction. Oral Diseases. 2006;**12**:283-289. DOI: 10.1111/j.1601-0825.2005.01191.x

[53] Tay JYY, Bay BH, Yeo JF, Harris M, Meghji S, Dheen ST. Identification of RANKL in osteolytic lesions of the facial skeleton. Journal of Dental Research. 2004;**83**:349-353. DOI: 10.1177/154405910408300415

[54] Menezes R, Bramante CM, da Silva Paiva KB, Letra A, Carneiro E, Fernando Zambuzzi W, et al. Receptor activator NFκB-ligand and osteoprotegerin protein expression in human periapical cysts and granulomas. Oral Surgery, Oral Med Oral Pathol Oral Radiol Endodontology. 2006;**102**:404-409. DOI: 10.1016/j.tripleo.2005.10.054

[55] Fukada SY, Silva TA, Garlet GP, Rosa AL, Da Silva JS, Cunha FQ. Factors involved in the T helper type 1 and type 2 cell commitment and osteoclast regulation in inflammatory apical diseases. Oral Microbiology and Immunology. 2009;**24**:25-31. DOI: 10.1111/j.1399-302X.2008.00469.x

[56] Fan R, Zhang C-F, Gao Y. Expression of receptor activator of nuclear factor kappaB ligand and osteoprotegerin in chronic apical periodontitis. Beijing Da Xue Xue Bao. Yi Xue Ban. 2008;**40**:39-42

[57] Fan R, Sun B, Zhang C, Lü Y, Xuan W, Wang Q, et al. Receptor activator of nuclear factor kappa B ligand and osteoprotegerin expression in chronic apical periodontitis: Possible association with inflammatory cells. Chinese Medical Journal. 2011;**124**:2162-2166

[58] Chi AC, Neville BW. Odontogenic cysts and tumors. Surgical Pathology Clinics. 2011;**4**:1027-1091. DOI: 10.1016/j.path.2011.07.003

[59] Neville BW, Damm DD, Chi AC, Allen CM. Oral and Maxillofacial Pathology. 4th ed. Elsevier Health Sciences; 2015. 928p

[60] Neville BW. Update on current trends in oral and maxillofacial pathology. Head and Neck Pathology. 2007;**1**:75-80. DOI: 10.1007/s12105-007-0007-4

[61] Morgan TA, Burton CC, Qian F. A retrospective review of treatment of the odontogenic keratocyst. Journal of Oral and Maxillofacial Surgery. 2005;**63**: 635-639. DOI: 10.1016/j.joms.2004.07.026

[62] Grasmuck EA, Nelson BL. Keratocystic odontogenic tumor. Head and Neck Pathology. 2010;**4**:94-96. DOI: 10.1007/s12105-009-0146-x

[63] Koçak-Berberoğlu H, Çakarer S, Brkić A, Gürkan-Koseoglu B, Altuğ-Aydil B, Keskin C. Three-dimensional cone-beam computed tomography for diagnosis of keratocystic odontogenic tumours; evaluation of four cases. Medicina Oral, Patología Oral y Cirugía Bucal. 2012;**17**:1000-1005. DOI: 10.4317/medoral.17629

[64] Madras J, Lapointe H. Keratocystic odontogenic tumour: Reclassification of the odontogenic keratocyst from cyst to tumour. Tex Dent Journal. 2008;**125**:446-54

[65] Jawa DS, Sircar K, Somani R, Grover N, Jaidka S, Singh S. Gorlin-Goltz syndrome. Journal of oral and maxillofacial pathology: JOMFP. 2009;**13**:89-92. DOI: 10.4103/0973-029X.57677

[66] Diniz MG, Gomes CC, de Sousa SF, Xavier GM, Gomez RS. Oncogenic signalling pathways in benign odontogenic cysts and tumours. Oral Oncology. 2017;**72**:165-173. DOI: 10.1016/j.oraloncology.2017.07.021

[67] Gurgel CAS, Buim MEC, Carvalho KC, Sales CBS, Reis MG, de Souza RO, et al. Transcriptional profiles of SHH pathway genes in keratocystic odontogenic tumor and ameloblastoma. Journal of Oral Pathology & Medicine. 2014;**43**:619-626. DOI: 10.1111/jop.12180

[68] Rui Z, Li-Ying P, Jia-Fei Q, Ying-Ying H, Feng C, Tie-Jun L. Smoothened gene alterations in keratocystic odontogenic tumors. Head & Face Medicine. 2014;**10**:1-7. DOI: 10.1186/1746-160X-10-36

[69] Diniz MG, Borges ÉR, Guimarães ALS, Moreira PR, Brito JAR, Gomez MV, et al. PTCH1 isoforms in odontogenic keratocysts. Oral Oncology. 2009;**45**:291-295. DOI: 10.1016/j.oraloncology.2008.05.020

[70] Santini R, Pietrobono S, Pandolfi S, Montagnani V, D'Amico M, Penachioni JY, et al. SOX2 regulates self-renewal and tumorigenicity of human melanoma-initiating cells. Oncogene. 2014;**33**:4697-4708. DOI: 10.1038/onc.2014.71

[71] Juuri E, Jussila M, Seidel K, Holmes S, Wu P, Richman J, et al. Sox2 marks epithelial competence to generate teeth in mammals and reptiles. Development. 2013;**140**:1424-1432. DOI: 10.1242/dev.089599

[72] Yalvac ME, Ramazanoglu M, Rizvanov AA, Sahin F, Bayrak OF, Salli U, et al. Isolation and characterization of stem cells derived from human third molar tooth germs of young adults: Implications in neo-vascularization, osteo-, adipo-and neurogenesis. The Pharmacogenomics Journal. 2010;**10**:105-113. DOI: 10.1038/tpj.2009.40

[73] Santos EM, Santos HO, dos Santos Dias I, Santos SH, de Paula AM, Feltenberger JD, Guimarães AL, Farias LC. Bioinformatics analysis reveals

genes involved in the pathogenesis of ameloblastoma and keratocystic odontogenic tumor. International journal of molecular and cellular medicine. 2016;**5**:199-219. DOI: 10.22088/acadpub.BUMS.5.4.199

[74] Seyedmajidi M, Nafarzadeh S, Siadati S, Shafaee S, Bijani A, Keshmiri N. p53 and PCNA expression in keratocystic odontogenic tumors compared with selected odontogenic cysts. International Journal of Molecular and Cellular Medicine. 2013;**2**:185-193

[75] Sengüven B, Oygür T. Investigation of interleukin-1 alpha and interleukin-6 expression and interleukin-1 alpha gene polymorphism in keratocystic odontogenic tumors and ameloblastomas. Med Oral Patol Oral Cir Bucal. 2011;**16**:e467-72. DOI: 10.4317/medoral.16.e467

[76] Chen W-L, Ouyang K-X, Li H-G, Huang Z-Q, Li J-S, Wang J-G. Expression of inducible nitric oxide synthase and vascular endothelial growth factor in ameloblastoma. The Journal of Craniofacial Surgery. 2009;**20**:171-175. DOI: 10.1097/SCS.0b013e31818435cd

[77] Diniz MG, Gomes CC, De Castro WH, Guimarães ALS, De Paula AMB, Amm H, et al. MiR-15a/16-1 influences BCL2 expression in keratocystic odontogenic tumors. Cellular Oncology. 2012;**35**:285-291. DOI: 10.1007/s13402-012-0087-3

[78] Zhang T, Chen M, Lü Y, Xing Q, Chen W. A novel mutation of the PTCH1 gene activates the Shh/Gli signaling pathway in a Chinese family with nevoid basal cell carcinoma syndrome. Biochemical and Biophysical Research Communications. 2011;**409**:166-170. DOI: 10.1016/j.bbrc.2011.04.047

[79] Heikinheimo K, Jee KJ, Morgan PR, Nagy B, Knuutila S, Leivo I. Genetic changes in sporadic keratocystic odontogenic tumour (odontogenic keratocyst). European Cells & Materials. 2007;**14**:64

Section 3

Advances in Transcriptional Control

Chapter 5

Data Mining Approaches for Understanding of Regulation of Expression of the Urea Cycle Genes

Ljubica Caldovic

Abstract

Urea cycle converts ammonia, a waste product of protein catabolism and a neurotoxin, into non-toxic urea. Urea cycle disorders are a group of rare genetic diseases that have protein-restricted diet as a common treatment modality. Expression of urea cycle genes is regulated in concert by the dietary protein intake, but the mechanism of this regulation is not well understood. Data mining of databases such as ENCODE and Cistrome can be used to gain new information about regulatory elements, transcription factors, and epigenetic mechanisms that regulate expression of urea cycle genes. This can lead to better understanding of the common mechanism, which regulates urea cycle genes, and can generate testable hypotheses about regulation of gene expression and new treatments for urea cycle disorders.

Keywords: urea cycle, transcriptional regulation, transcription factor, data mining, ChIP-Seq, epigenetic, histone modification, chromatin remodeler

1. Introduction

Transcriptional regulation of gene expression is essential for development, tissue differentiation, and organisms' responses to changes in their environment. Maintenance of homeostasis would be impossible without regulation of expression of genes that code for enzymes in the carbohydrate, fat, and protein metabolism. Omnivorous mammals, such as humans, mice, and rats, can adapt their metabolism to varying proportions of amino acids, fats, and carbohydrates as sources of energy [1, 2]. Strict carnivores, such as cats, lack such adaptive mechanisms [3]. Diet rich in carbohydrates and fats triggers utilization of these nutrients as sources of energy and storage of excess sugars and fat in the form of glycogen and adipose tissue [1, 4]. On the other hand, a diet rich in proteins imposes changes in nitrogen balance because excess proteins and amino acids cannot be stored [1]. High intake of proteins, combined with low intake of carbohydrates and fats, leads to utilization of amino acids as energy sources and results in increased catabolism of amino acids [1, 4] and increased need to dispose of waste nitrogen, generated in this process. Transcriptional regulation of expression of enzymes in the pathways for degradation of nutrients as well as biosynthesis of molecules that can be stored is necessary for adaptations to these dietary changes.

Urea cycle is a liver metabolic pathway that converts ammonia into urea. Ammonia is a waste product of catabolism of dietary and cellular proteins, and a potent neurotoxin [5]. Defects in any of the enzymes of the urea cycle lead to partial or complete block of urea production and accumulation of ammonia, which damages the brain and can be lethal [5]. Therefore, the physiological role of urea cycle is to protect the brain from the toxic effects of ammonia.

Five enzymes of the urea cycle, carbamoyl phosphate synthetase 1 (CPS1), ornithine transcarbamylase (OTC), argininosuccinate synthase (ASS), argininosuccinate lyase (ASL) and arginase 1, catalyze conversion of ammonia into urea. In addition, N-acetylglutamate synthase (NAGS), ornithine/citrulline transporter (ORNT), and aspartate/glutamate transporter (also known as ARALAR2 or Citrin) are required for the normal function of the urea cycle. The pathway is partitioned between mitochondria and cytoplasm of the liver cells. NAGS, CPS1, and OTC are located in the mitochondrial matrix [6] while ASS, ASL, and arginase 1 are in the cytosol [6]. **Figure 1** shows enzymes, transporters, and intermediates of the urea cycle together with their cellular localization. The NAGS (EC 2.3.1.1) uses glutamate and acetyl coenzyme A to produce NAG, an obligatory allosteric activator of CPS1 [6–9]. Upon binding of NAGS, CPS1 (EC 6.3.4.16) consumes two molecules of ATP to catalyze formation of carbamoyl phosphate from ammonia and bicarbonate [6, 10]. Carbamoyl phosphate and ornithine, which is transported into mitochondria by ORNT, are converted to citrulline by OTC (EC 2.1.3.3), followed by the transport of citrulline from mitochondria into cytoplasm by ORNT [6]. ASS (EC 6.3.4.5) catalyzes formation of argininosuccinate from citrulline and aspartate, which is supplied by the Citrin [6]. The argininosuccinate is then converted

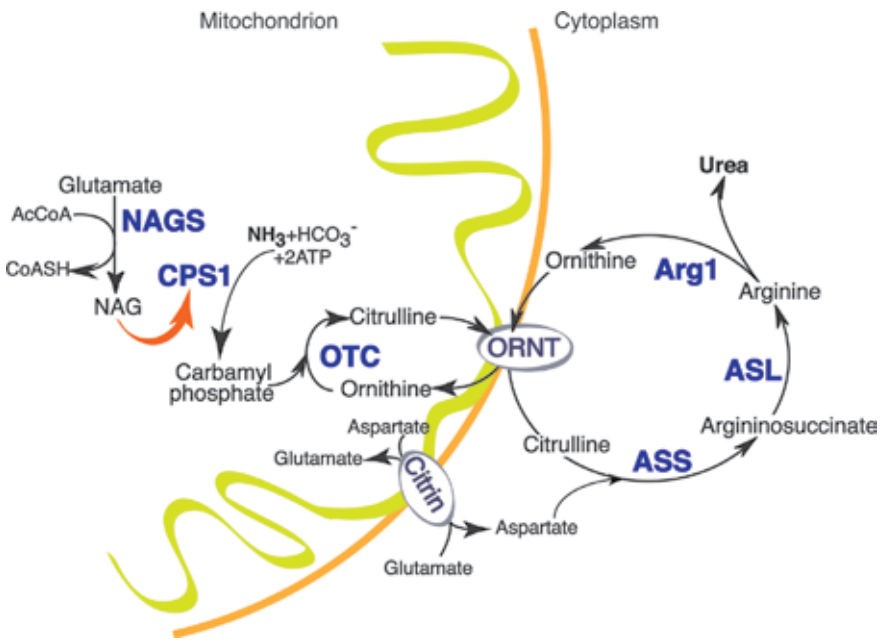

Figure 1.
Enzymes, transporters, and intermediates of the urea cycle. Urea cycle enzymes are shown in blue typeface: NAGS—N-acetylglutamate synthase, CPS1—carbamoyl phosphate synthetase 1, OTC—ornithine transcarbamylase, ASS—argininosuccinate synthase, ASL—argininosuccinate lyase, and Arg1—arginase 1. Urea cycle transporters ORNT (ornithine/citrulline transporter) and Citrin (glutamate/aspartate transporter) are shown as ovals. The orange arrow indicates activation of CPS1 by N-acetylglutamate (NAG). AcCoA—acetyl coenzyme A. CoASH—coenzyme A.

to arginine by the ASL (EC 4.3.2.1). Arginase 1 (EC 3.5.3.1) completes the cycle by hydrolyzing arginine into urea and ornithine, which is transported into mitochondria to be a substrate for OTC (**Figure 1**). Urea cycle genes and enzymes are not uniformly expressed in the liver; their expression follows a gradient from high in hepatocytes surrounding portal vein to no expression in hepatocytes surrounding central vein [11]. NAGS, CPSI, and OTC are also expressed in the small intestine, where they synthesize citrulline which is then transported to the kidneys for the biosynthesis of arginine by the ASS and ASL [6]. Argininosuccinate synthase and lyase also function in the NO signaling and are present in the tissues that express nitrous oxide synthase [5].

Long-term changes in dietary protein intake lead to adaptive changes in expression of urea cycle enzymes. Their expression increases in rats and monkeys fed high-protein diet and decreases upon feeding of low-protein diet [4, 12]. These adaptive changes seem to be mediated, at least in part, by the hormones glucagon, glucocorticoid, and insulin [13–25]. Glucagon and glucocorticoid hormones trigger changes in mRNA and protein levels of all five urea cycle enzymes, but the mechanisms responsible for these changes seem to differ for each enzyme [21]. It is not known whether changes in the dietary protein intake trigger similar changes in the expression of the NAGS gene because it has not been identified at the time. Also unknown are signaling cascades that mediate effects of hormones to regulate expression of urea cycle genes and whether specific amino acids and/or other metabolites act as sensors of the dietary protein intake.

Inspection of the regulatory regions of genes for urea cycle enzymes (**Figure 2**) does not reveal a common regulatory element that would bind one or more transcriptional factors to coordinately regulate transcription of all urea cycle genes [13, 26–36]. The studies of expression of urea cycle genes in knockout mice also show a lack of common regulatory mechanism. Ureagenesis is defective in mice lacking hepatocyte nuclear factor 4 α (HNF4α) due to absence of OTC mRNA and protein [37] as well as in mice lacking CCAAT/enhancer binding protein α (C/EBPα) due to lack of CPSI mRNA [38]. However, ureagenesis appears normal in mice lacking C/EBPβ although this transcriptional factor appears to regulate expression of the arginase 1 gene [39]. It is also unknown if short-term increases in nitrogen load following a meal trigger any change in expression of urea cycle enzymes.

This chapter focuses on the regulation of NAGS, CPS1, and OTC expression because their only known functions are protection of the brain from ammonia toxicity through participation in the urea cycle and intestinal biosynthesis of citrulline. The three genes share common expression pattern in the liver, intestine, and during development. Because of the role of urea cycle in protecting the brain from ammonia toxicity, expression of the three genes have been studied in much greater detail in the liver than intestinal cells. Detailed understanding of the transcriptional regulation of the urea cycle genes is important for our understanding of bodies' response to changes in the environment such as dietary changes as well as events that trigger increased catabolism of cellular proteins such as starvation, infections, and invasive medical procedures [6, 40–43]. Because regulation of expression of the mammalian CPS1 and OTC genes has been studied for more than three decades while expression of human NAGS, which was identified and cloned in 2002, took place less than a decade ago the approaches taken in these studies differed greatly. Knowledge of transcriptional regulation of mammalian CPS1 and OTC was gained through cloning of genomic DNA, construction of reporter gene plasmids with various fragments from the CPS1 and OTC regulatory regions and their expression in cultured cells and transgenic mice, whereas regulatory elements of the NAGS gene have been identified using comparative genomics approaches.

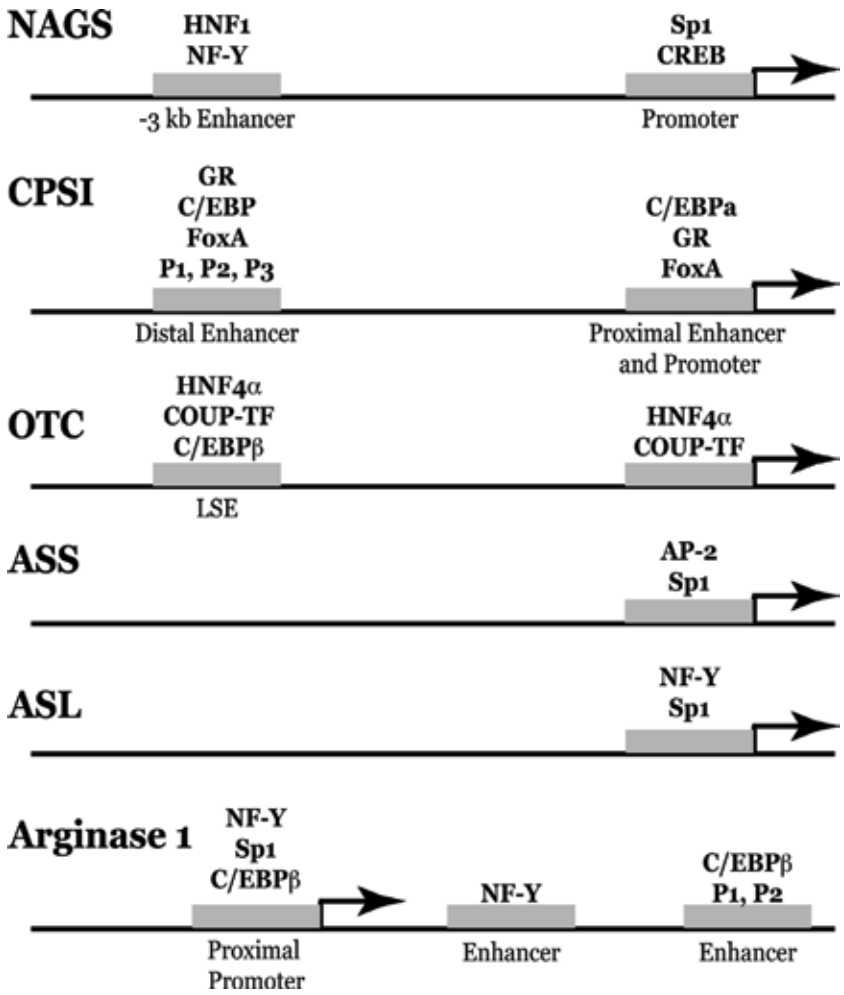

Figure 2.
Transcriptional regulatory elements of the urea cycle genes. Hooked arrows indicate transcriptional initiation sites. Gray boxes—cisacting regulatory elements. Proteins that bind to cis-acting regulatory elements are shown above each box. C/EBP—CCAAT/enhancer binding proteins; HNF4α—hepatic nuclear factor a; COUP-TF—chicken ovalbumin upstream promoter transcriptional factor; FoxA—forkhead box A transcriptional factor/ hepatic nuclear factor 3; GR—glucocorticoid receptor; P3—unidentified protein of approx. 75 kDa; P1, P2—two unidentified proteins; NF-Y—CCAAT—binding factor; AP-2—activator protein 2; Sp1—Sp1 transcriptional factor; CREB—cAMP response element binding transcription factor.

2. Transcriptional regulation of urea cycle genes

2.1 Transcriptional regulation of mammalian NAGS gene

Although the existence of mammalian NAGS gene and its product have been known since the 1950s [44], the gene remained elusive until 2002, when it was identified and cloned in mice and humans [45–48]. The human *NAGS* gene is located on chromosome 17 within band 17q21.31 and spans approximately 8.5 kb. This includes seven exons that encode a 1605 bp open reading frame, six introns, a promoter, and an enhancer located about 3 kb upstream of the transcription start sites [45, 47, 49–51]. The mouse *Nags* gene is located in the syntenic region on chromosome 11. Pairwise BLAST [52] was used for comparison of the regions upstream of the NAGS genes from seven mammals including human; this analysis revealed two conserved elements, one located immediately upstream of the first exon of

the *NAGS* gene and a putative regulatory element located about 3 kb upstream of the *NAGS* translation initiation site [50]. The pattern of DNA sequence conservation within the conserved region immediately upstream of the first exon of the *NAGS* gene suggested that it might consist of a promoter and a proximal regulatory element, which is similar to the CPS1 regulatory region that will be described in the next section. Cis-element over-representation (CLOVER) software [53] was then used to identify binding sites for specificity protein 1 (Sp1), cAMP response element binding/activating transcription factor (CREB/ATF), and CCAAT-enhancer binding protein (C/EBP) transcription factors in the putative *NAGS* promoter, while activator protein-2 (AP2), hepatic nuclear factor 1 (HNF1), nuclear factor-Y (NF-Y), and mothers against decapentaplegic homolog 3 (SMAD3) binding sites were found in the upstream regulatory element, which was named −3 kb enhancer. These findings were then experimentally verified.

Reporter gene assays were used to confirm that conserved regions located adjacent to and −3 kb upstream of the first *NAGS* exon indeed function as a promoter and enhancer in the HepG2 hepatoma cells [50]. Consistent with the absence of the TATA-box in the NAGS promoter, transcription of the *NAGS* mRNA in the liver and intestine initiates at multiple sites located between 50 and 150 bp upstream of the NAGS translation initiation codon [50]. Binding of the Sp1 and CREB transcription factors to the *NAGS* promoter, and binding of the HNF1 and NF-Y transcription factors to the −3 kb enhancer were confirmed with chromatin immunoprecipitation (ChIP) and DNA pull-down assays [50]. Binding of HNF1 to the −3 kb NAGS enhancer is responsible for the liver-specific expression of the *NAGS* gene [50] and the role of HNF1 transcription factor in expression of the *NAGS* gene was confirmed when a sequence variant that caused decreased HNF1 binding to its site was found in a patient with NAGS deficiency [49].

2.2 Transcriptional regulation of mammalian CPS1 gene

Regulatory region of the rat *Cps1* gene has been cloned in 1985 [54]. Almost all of our knowledge of transcriptional regulation of the CPS1 gene is based on experiments with the rat *Cps1* gene in the rat and human hepatoma cell lines and transgenic mice. The aim of these studies was to elucidate the mechanism of regulation of *Cps1* expression by glucagon and glucocorticoids as well as identify regulatory elements that restrict *Cps1* expression to periportal hepatocytes [55–57].

A promoter, located immediately upstream of the first *Cps1* exon, a proximal enhancer that is immediately adjacent to the promoter, a distal enhancer located about 6 kb upstream of the *Cps1* translation initiation codon, and another regulatory element located about 10 kb upstream of the *Cps1* translation initiation site are responsible for transcriptional regulation of the rat *Cps1* gene [27, 29, 56, 58]. Transcription of the rat *Cps1* mRNA is initiated 138–140 bp upstream of the translation initiation site by the promoter that has TATA and CAAT motifs [59] and binds C/EBP transcription factor [29]. The distal *Cps1* enhancer consists of a cAMP response unit (CRU) and a glucocorticoid response unit (GRU); each response unit binds multiple transcription factors that activate *Cps1* expression in response to glucagon and glucocorticoids, and are responsible for *Cps1* expression in periportal hepatocytes [26, 55, 56, 60]. The CRU binds CREB, HNF3, C/EBP transcription factors, and a yet to be identified protein P1 [60, 61], while GRU binds glucocorticoid receptor, hepatocyte nuclear factor 3/forkhead box A (HNF3/FOXA), C/EBP, a 75 kDa protein P3, and a yet to be identified protein P2 [26, 61, 62]. The distal *Cps1* enhancer activates *Cps1* transcription via proximal enhancer that binds C/EBP and glucocorticoid receptor and is located immediately upstream of the *Cps1* promoter [63]. These studies of the rat *Cps1* gene regulation rest on the premise that the rat

gene is a good model for human *CPS1*. However, the two species have different metabolic rates due to their different sizes and regulation of the CPS1 gene and urea cycle may differ in the two organisms. More recently, a region of the human *CPS1* gene that corresponds to the rat *Cps1* promoter and proximal enhancer has been shown to bind HNF3 and direct reporter gene expression in hepatoma cells [64]. Human *CPS1* gene is located on chromosome 2, band 2q34 where it spans approx. 125 kb and has 38 exons that encode a 1500 amino acids long protein.

2.3 Transcriptional regulation of mammalian OTC gene

The human *OTC* gene is 70 kb long and has 10 exons which contain a 1062 bp long coding sequence [65]. Transcription of the human *OTC* gene initiates at multiple transcription start sites (TSS) [66], while transcription of the mouse and rat *Otc* genes initiate at a single transcription start site located 136 and 98 bp upstream of the translation initiation codon [67, 68]. Within the rat *Otc* promoter, four regions, A–D, bind transcription factors that regulate expression of the *Otc* gene [31]. Region A is a negative regulator of *Otc* transcription [31], and transcription factors that bind to this region have not been identified. Regions B and C bind transcriptional activator HNF4, and transcriptional repressor chicken ovalbumin upstream promoter—transcription factor (COUP-TF) [30, 31]. Region D is located downstream of the transcription start site and its role in expression of the *Otc* gene remains to be elucidated [31].

The rat *Otc* promoter is sufficient for expression of transgenes in the liver and intestine of transgenic mice [69, 70]. An enhancer located approximately 11 kb upstream of the first exon of the rat *Otc* gene is responsible for a high level of expression of the *Otc* gene in the liver [31]. This −11 kb enhancer has four transcription factor-binding sites, designated I–IV [31]; sites I and II bind C/EBPβ, while transcription factor HNF4 binds to sites I and IV in the rat *Otc* enhancer to activate expression of the *Otc* gene [32, 37, 38]. Since comparative genomics studies revealed that the distance between regions that correspond to the −11 kb rat *Otc* enhancer and OTC promoter vary in mammalian genomes, this region was renamed as the liver-specific enhancer (LSE) [71].

2.4 Transcriptional regulation of the NAGS, CPS1, and OTC genes in the genomics era

Advances in sequencing technology-enabled sequencing of dozens of mammalian genomes and comparisons of their sequences revealed conserved regions in non-coding regions that could function as regulatory elements [72, 73]. This strategy was used to identify NAGS promoter and enhancer [50]. Next-generation sequencing also enabled examination of the function of non-coding regions in the human and mouse genomes including their chromatin structure, and binding of transcription and chromatin remodeling proteins to generate an Encyclopedia of Non-coding DNA Elements (ENCODE). These studies were first carried out in the limited number of cultured cell lines, but are now expanding to include tissues and cultured primary cells and their results have been stored in the ENCODE database [74, 75]. In addition to these large-scale projects, many individual labs have been performing ChIP-Seq experiments and the publically available results of their experiments are being gathered in the Cistrome database [76]. The advantage of the Cistrome database is ability to compare chromatin states and track changes in binding of transcription factors in response to signaling molecules, treatments, and environmental stimuli. Data mining of the ENCODE and Cistrome databases present an opportunity to identify novel regulatory elements in the *NAGS*, *CPS1*,

and *OTC* genes and transcription factors that bind to the regulatory elements. Both databases were queried for chromatin modifications and binding of transcription factors to the *NAGS*, *CPS1*, and *OTC* genes and their flanking regions in the liver tissue using following coordinates of the GRCh38/hg38 human genome assembly: chr17:43,994,682-44,012,832 for the *NAGS* gene, chr2:210,499,833-210,691,279 for the *CPS1* gene, and chrX:38,334,777-38,459,529 for the *OTC* gene. The following filters were applied to experimental matrix of the ENCODE database (www.encodeproject.org): organism—*Homo sapiens*, biosample type—tissue, organ—liver, project—ENCODE, genome assembly—GRCh38, assay category—ChIP-Seq, assay category—DNA binding, and target of assay—transcription factor, histone, broad and narrow histone mark, and chromatin remodeler. Results of the query were visualized using UCSC Genome Browser. Results of the ChIP-Seq experiments for the genomic region of interest can be downloaded as either wiggle or bed files using Tools and Table Browser menus of the UCSC Genome Browser. The ChIP-Seq data for each DNA binding protein and histone modification of interest can be acquired by selecting ENCODE Hub from the group menu, ENCODE ChIP-Seq from the track menu and experiment ID from the table menu of the Table Browser page. Cistrome Data Browser was used to query Cistrome database (www.cistrome.org); *Homo sapiens* was selected as species and hepatocyte as biological source. Experimental results that passed quality controls were visualized in the UCSC Genome Browser and results of experiments were obtained is the same way as for the ENCODE database.

The 5′-ends of each region were chosen based on the presence of RAD21, a component of cohesion, and CTCF binding that can indicate boundaries of chromatin domains, whereas the 3′-ends of the *NAGS* and *OTC* genomic region have been chosen to be either within or close to their downstream neighboring genes; the 3′-end of the *CPS1* genomic region was chosen to include a conserved region downstream of the last exon of the *CPS1* gene (**Figures 3A**, **4A**, and **5A**). RAD21 and CTCF bind to additional sites within *NAGS*, *CPS1*, and *OTC* genes and the role of cohesion and CTCF in expression of the three genes is yet to be determined (**Figures 3A**, **4A**, and **5A**).

The H3K4me3 histone 3 modifications that mark active promoters and H3K27ac modifications that mark active enhancers are present at upstream regions of all three genes (**Figures 3B**, **4B**, and **5B**). The ENCODE database also has the DNaseI sensitivity data from the human fetal liver tissue that show open chromatin state for the *NAGS* gene and closed chromatin for the *CPS1* and *OTC* genes (**Figures 3B**, **4B**, and **5B**). This difference could be due to presence of the ubiquitously expressed *TMEM101* gene downstream of the *NAGS* gene (**Figure 3**).

Query of the ChIP-Seq data in the ENCODE database confirmed binding of the Sp1, CREB/ATF3, HNF4A, HNF3/FOXA1, HNF3/FOXA2, and COUP-TF to the promoters and enhancers of the *NAGS*, *CPS1*, and *OTC* genes and identified binding of these and several other transcription factors to previously identified as well as novel regulatory elements (**Figures 3C**, **4C**, and **5C**). Transcription factors RXRA, COUP-TF and HNF4A, YY1 and JUND/AP1 also appear to bind *NAGS* promoter while RXRA, HNF4A, YY1, and REST bind the −3 kb *NAGS* enhancer (**Figure 3C**). The ChIP-Seq data also show binding of transcription factor to regions of the NAGS gene that could be novel regulatory elements. For example, Sp1, RXRA, and HNF4A bind to a region in the first intron of the *NAGS* gene. ChIP-Seq data also show that transcription factors bind to a region located between *NAGS* promoter and −3 kb enhancer as well as to the two regions upstream of the −3 kb enhancer; these sites could be novel regulatory elements of the *NAGS* gene (**Figure 3C**). The map of the *NAGS* genomic region shows that transcription of the *PYY* gene initiates at the *NAGS* promoter but in the opposite direction (**Figure 3D**). This is because a

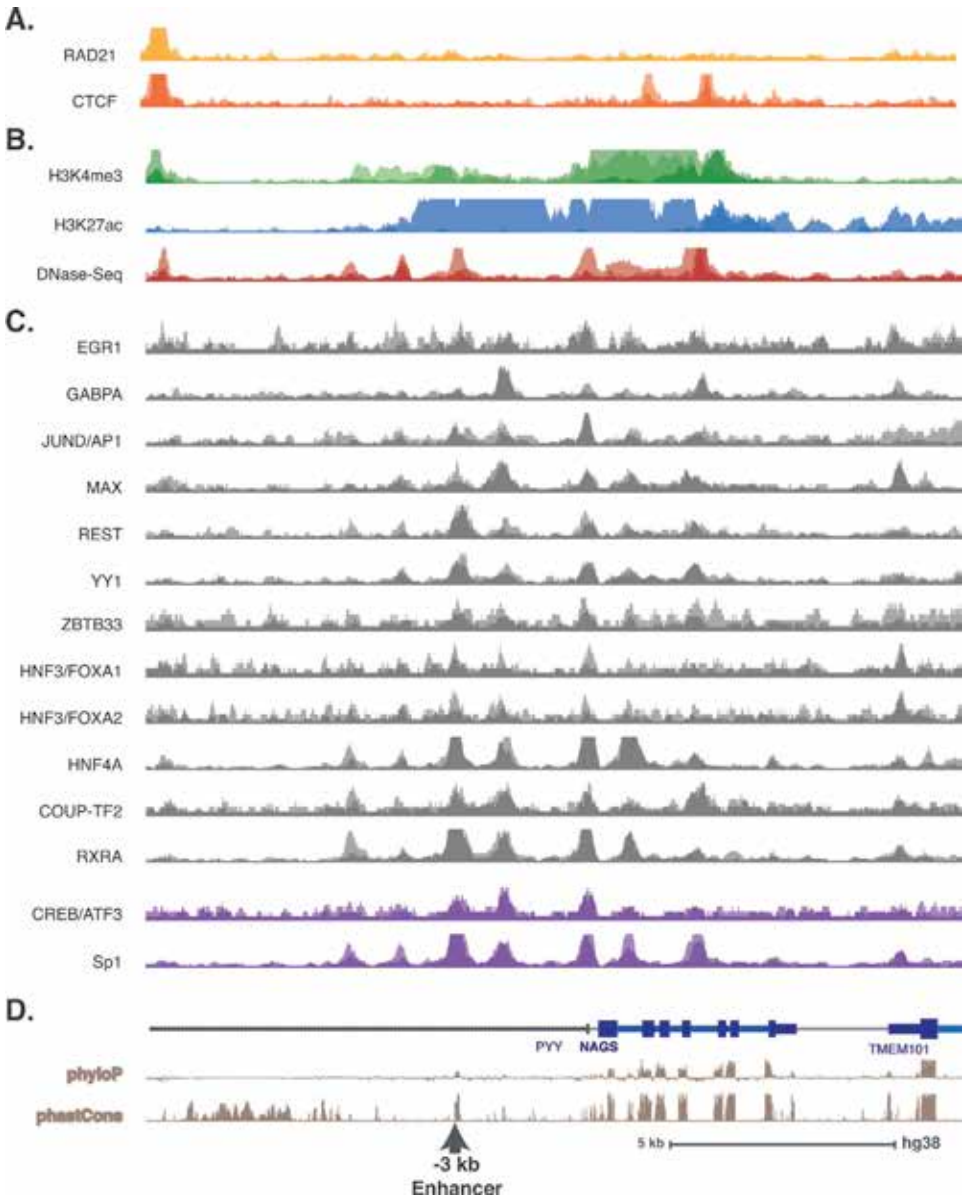

Figure 3.
Overview of the NAGS genomic region. (A) Binding sites for chromatin remodeling proteins. (B) Epigenetic marks. (C) Transcription factor binding sites. (D) Map and sequence conservation of the NAGS genomic region.

PYY cDNA (GenBank ID BC041057.1) has been isolated from a brain astrocytoma sample and sequenced [77]. This *PYY* transcript may have resulted from aberrant expression of the *PYY* gene in the brain astrocytoma cells since *PYY* is not expressed in the brain according to the Human Protein Atlas [78] and the GTEx track of the UCSC Genome Browser [79].

Transcription factors HNF3/FOXA1, HNF3/FOXA2, and CREB bind to the human *CPS1* promoter and the region upstream of the human *CPS1* that corresponds to the rat *Cps1* distal enhancer as well as additional sites located within the first intron of the *CPS1* gene and upstream of the distal enhancer (**Figure 4C**). Moreover, HNF4A, RXRA, SP1, YY1, JUND/AP2, and REST also bind to the *CPS1*

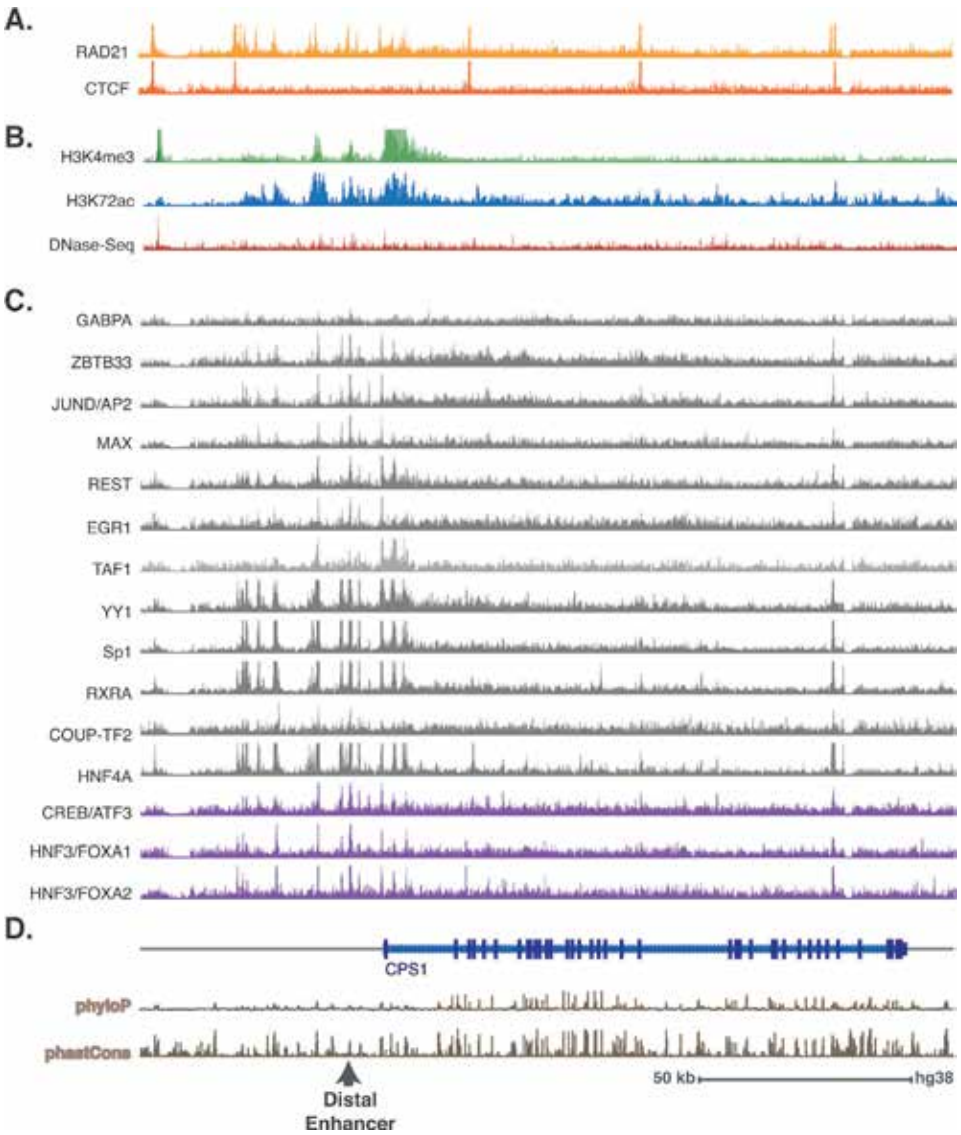

Figure 4.
Overview of the CPS1 genomic region. (A) Binding sites for chromatin remodeling proteins. (B) Epigenetic marks. (C) Transcription factor binding sites. (D) Map and sequence conservation of the CPS1 genomic region.

upstream region and first intron (**Figure 4C**). It is possible that yet to be identified proteins P1, P2, and P3 that bind to the rat *Cps1* distal enhancer are among these transcription factors. Likewise, the HNF4A and COUP-TF transcription factors that are known to bind to the *OTC* promoter and LSE also bind to sites upstream of the LSE and within first *OTC* intron, and additional transcription factors bind to these regions (**Figure 5C**). The novel regulatory elements that have been identified by the binding of transcription factors coincide with regions that are conserved in vertebrates as indicated by the phyloP and phastCons tracks of the UCSC Genome Browser (**Figures 3D**, **4D**, and **5D**).

These data mining efforts identified a common set of transcription factors that bind to the regulatory regions of the *NAGS*, *CPS1*, and *OTC* genes in the liver and may be responsible for the coordinated changes in their expression in response to dietary protein intake and hormonal signaling. The knowledge of transcription

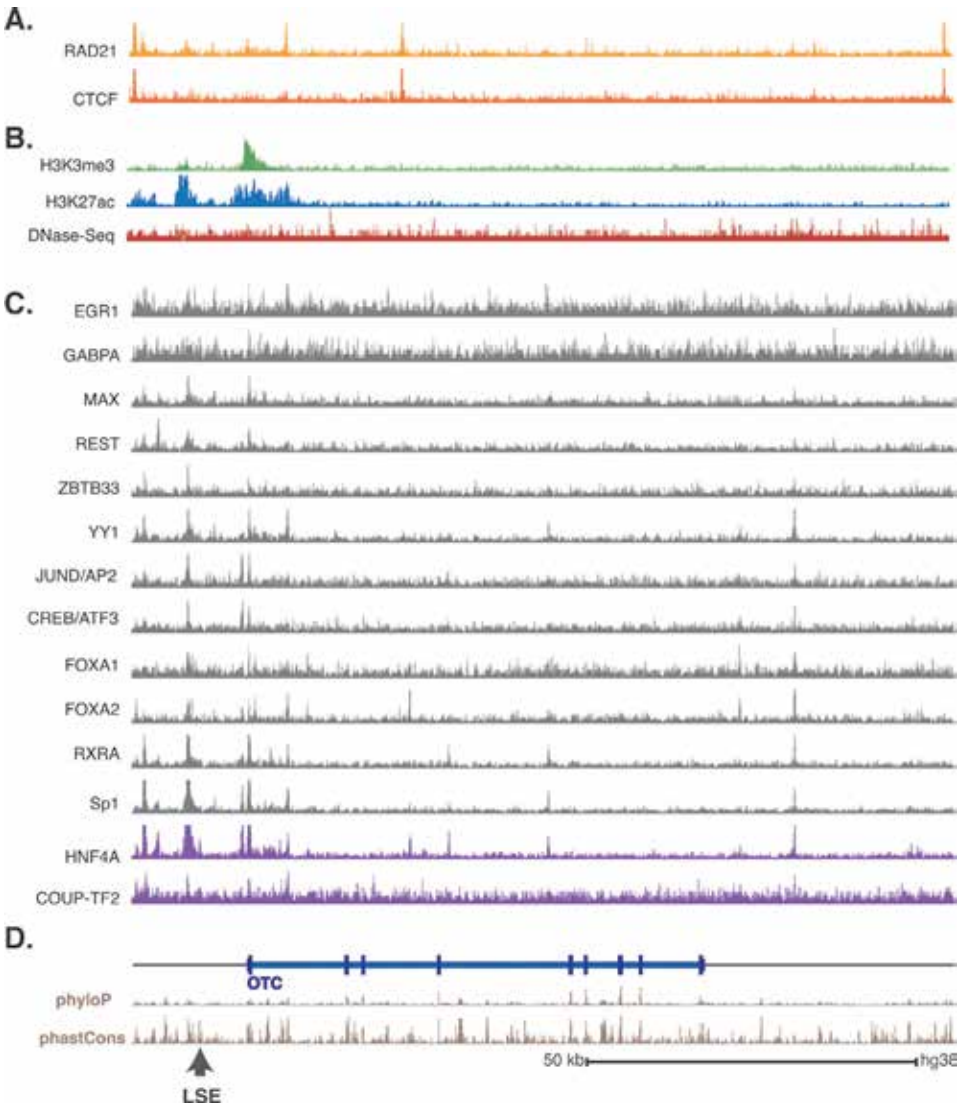

Figure 5.
Overview of the OTC genomic region. (A) Binding sites for chromatin remodeling proteins. (B) Epigenetic marks. (C) Transcription factor binding sites. (D) Map and sequence conservation of the OTC genomic region.

factors that regulate expression of urea cycle genes could provide a clue about amino acid(s) and metabolites that act as sensors of the dietary protein intake. Data mining of the ENCODE and Cistrome databases presented here revealed that transcription factor RXRA binds to regulatory elements of *NAGS*, *CPS1*, and *OTC* genes. The RXRA transcription factor regulates gene expression by forming heterodimers with several transcription factors including peroxisome proliferator-activated receptor gamma (PPARγ), which regulated glucose metabolism [80]. If future studies show that PPARγ-RXRA heterodimer regulates expression of the *NAGS*, *CPS1*, and *OTC* genes, that would suggest that glucose, rather than amino acid(s), might be a sensor of the balance of protein and carbohydrate intake that regulates expression of the urea cycle genes. Although similar data sets are not yet available for the human small intestine cells, the ENCODE project is ongoing and future data mining efforts will provide a more complete information about transcriptional regulation of the *NAGS*, *CPS1*, and *OTC* genes. Similarly, the Cistrome

database is growing and queries of its data may reveal molecular mechanisms of *NAGS*, *CPS1*, and *OTC* regulation through differential binding of transcription factors to their regulatory elements. The utility of the data mining approach goes beyond understanding of transcriptional regulation of genes. This approach can be used to explain deleterious effects of sequence variants on expression of genes that are associated with human diseases and identify drug targets for treatment of diseases that can benefit from increased expression of hypomorphic alleles. In the case of *NAGS*, *CPS1*, or *OTC* deficiencies, which have high plasma ammonia, or hyperammonemia, as a common symptom, partial defects in any of the three genes result in the decreased activity or abundance of the corresponding enzyme and decreased capacity for ureagenesis. Protein-restricted diet, which minimizes ammonia production, is standard therapy for patients with partial defects of the *NAGS*, *CPS1*, *OTC*, and other urea cycle genes that were not discussed in this chapter. However, protein-restricted diet also leads to decreased expression of urea cycle genes, including the defective one, leading to further decrease patient's capacity for ureagenesis and increased risk of hyperammonemia. A drug therapy that is based on transcriptional regulation of the *NAGS*, *CPS1*, and *OTC* genes might be able to increase their expression even when patients are on the protein-restricted diet and decrease patients' risk of hyperammonemia-induced brain damage.

Acknowledgements

This work was supported by public health service grants K01DK076846 and R01DK064913 from the National Institute of Diabetes Digestive and Kidney Diseases, National Institutes of Health, Department of Health and Human Services.

Conflict of interest

The author has no conflicts of interest to declare.

Author details

Ljubica Caldovic
Children's National Medical Center, Center for Genetic Medicine Research, The George Washington University, Institute for Biomedical Sciences, Washington, DC, USA

*Address all correspondence to: lcaldovic@childrensnational.org

References

[1] Waterlow JC. The mysteries of nitrogen balance. Nutrition Research Reviews. 1999;**12**:25-54

[2] Waterlow JC. The nature and significance of nutritional adaptation. European Journal of Clinical Nutrition. 1999;**53**(Suppl. 1):S2-S5

[3] Stewart PM, Batshaw M, Valle D, Walser M. Effects of arginine-free meals on ureagenesis in cats. The American Journal of Physiology. 1981;**241**(4):E310-E315

[4] Schimke RT. Adaptive characteristics of urea cycle enzymes in the rat. Journal of Biological Chemistry. 1962;**237**:459-468

[5] Ah Mew N, Simpson KL, Gropman AL, Lanpher BC, Chapman KA, Summar ML. Urea cycle disorders overview. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, LJH B, Stephens K, et al., editors. GeneReviews®. Seattle (WA): University of Washington; 1993

[6] Brusilow SW, Horwich AL. Urea cycle enzymes. In: Scriver CR, Beaudet AL, Sly WS, Valle D, editors. The Metabolic & Molecular Bases of Inherited Disease. Vol. 2. McGraw-Hill; 2001. pp. 1909-1963

[7] Caldovic L, Tuchman M. N-acetylglutamate and its changing role through evolution. Biochemical Journal. 2003;**372**(Pt 2):279-290

[8] Meijer AJ, Lamers WH, Chamuleau RAFM. Nitrogen metabolism and ornithine cycle function. Physiological Reviews. New-York. 1990;**70**(3):701-748

[9] Powers-Lee SG. N-Acetylglutamate synthase. Methods in Enzymology. 1985;**113**:27-35

[10] Shi D, Caldovic L, Tuchman M. Sources and fates of carbamyl phosphate: A labile energy-rich molecule with multiple facets. Biology. 2018;**7**(2). DOI: 10.3390/biology7020034. Basel

[11] Haussinger D, Lamers WH, Moorman AF. Hepatocyte heterogeneity in the metabolism of amino acids and ammonia. Enzyme. 1992;**46**(1-3):72-93

[12] Felipo V, Minana MD, Grisolia S. Control of urea synthesis and ammonia utilization in protein deprivation and refeeding. Archives of Biochemistry and Biophysics. 1991;**285**(2):351-356

[13] Hoogenkamp M, Gaemers IC, Schoneveld OJ, Das AT, Grange T, Lamers WH. Hepatocyte-specific interplay of transcription factors at the far-upstream enhancer of the carbamoylphosphate synthetase gene upon glucocorticoid induction. The FEBS Journal. 2007;**274**(1):37-45

[14] McLean P, Gurney MW. Effect of adrenalectomy and of growth hormone on enzymes concerned with urea synthesis in rat liver. Biochemical Journal. 1963;**87**:96-104

[15] Schimke RT. Studies on factors affecting the levels of urea cycle enzymes in rat liver. Journal of Biological Chemistry. 1963;**238**:1012-1018

[16] McLean P, Novello F. Influence of pancreatic hormones on enzymes concerned with urea synthesis in rat liver. Biochemical Journal. 1965;**94**:410-422

[17] Snodgrass PJ, Lin RC, Muller WA, Aoki TT. Induction of urea cycle enzymes of rat liver by glucagon. Journal of Biological Chemistry. 1978;**253**(8):2748-2753

[18] Gebhardt R, Mecke D. Permissive effect of dexamethasone on glucagon

induction of urea-cycle enzymes in perifused primary monolayer cultures of rat hepatocytes. European Journal of Biochemistry. 1979;**97**(1):29-35

[19] Lin RC, Snodgrass PJ, Rabier D. Induction of urea cycle enzymes by glucagon and dexamethasone in monolayer cultures of adult rat hepatocytes. Journal of Biological Chemistry. 1982;**257**(9):5061-5067

[20] Nebes VL, Morris SM Jr. Regulation of messenger ribonucleic acid levels for five urea cycle enzymes in cultured rat hepatocytes. Requirements for cyclic adenosine monophosphate, glucocorticoids, and ongoing protein synthesis. Molecular Endocrinology. 1988;**2**(5):444-451

[21] Ulbright C, Snodgrass PJ. Coordinate induction of the urea cycle enzymes by glucagon and dexamethasone is accomplished by three different mechanisms. Archives of Biochemistry and Biophysics. 1993;**301**(2):237-243

[22] Kitagawa Y, Sugimoto E. Interaction between glucocorticoids, 8-bromoadenosine 3′,5′-monophosphate, and insulin in regulation of synthesis of carbamoyl-phosphate synthetase I in Reuber hepatoma H-35. European Journal of Biochemistry. 1985;**150**(2):249-254

[23] Kitagawa Y, Ryall J, Nguyen M, Shore GC. Expression of carbamoyl-phosphate synthetase I mRNA in Reuber hepatoma H-35 cells. Regulation by glucocorticoid and insulin. Biochimica et Biophysica Acta. 1985;**825**(2):148-153

[24] Kitagawa Y. Hormonal regulation of carbamoyl-phosphate synthetase I synthesis in primary cultured hepatocytes and Reuber hepatoma H-35. Defective regulation in hepatoma cells. European Journal of Biochemistry. 1987;**167**(1):19-25

[25] Husson A, Renouf S, Fairand A, Buquet C, Benamar M, Vaillant R. Expression of argininosuccinate lyase mRNA in foetal hepatocytes. Regulation by glucocorticoids and insulin. European Journal of Biochemistry. 1990;**192**(3):677-681

[26] Christoffels VM, Grange T, Kaestner KH, Cole TJ, Darlington GJ, Croniger CM, et al. Glucocorticoid receptor, C/EBP, HNF3, and protein kinase A coordinately activate the glucocorticoid response unit of the carbamoylphosphate synthetase I gene. Molecular and Cellular Biology. 1998;**18**(11):6305-6315

[27] Goping IS, Shore GC. Interactions between repressor and anti-repressor elements in the carbamyl phosphate synthetase I promoter. Journal of Biological Chemistry. 1994;**269**(5):3891-3896

[28] Goping IS, Lagace M, Shore GC. Factors interacting with the rat carbamyl phosphate synthetase promoter in expressing and nonexpressing tissues. Gene. 1992;**118**(2):283-287

[29] Lagace M, Goping IS, Mueller CR, Lazzaro M, Shore GC. The carbamyl phosphate synthetase promoter contains multiple binding sites for C/EBP-related proteins. Gene. 1992;**118**(2):231-238

[30] Kimura A, Nishiyori A, Murakami T, Tsukamoto T, Hata S, Osumi T, et al. Chicken ovalbumin upstream promoter-transcription factor (COUP-TF) represses transcription from the promoter of the gene for ornithine transcarbamylase in a manner antagonistic to hepatocyte nuclear factor-4 (HNF-4). Journal of Biological Chemistry. 1993;**268**(15):11125-11133

[31] Murakami T, Nishiyori A, Takiguchi M, Mori M. Promoter and 11-kilobase upstream enhancer elements responsible for hepatoma cell-specific expression

of the rat ornithine transcarbamylase gene. Molecular and Cellular Biology. 1990;**10**(3):1180-1191

[32] Nishiyori A, Tashiro H, Kimura A, Akagi K, Yamamura K, Mori M, et al. Determination of tissue specificity of the enhancer by combinatorial operation of tissue-enriched transcription factors. Both HNF-4 and C/EBP beta are required for liver-specific activity of the ornithine transcarbamylase enhancer. Journal of Biological Chemistry. 1994;**269**(2):1323-1331

[33] Anderson GM, Freytag SO. Synergistic activation of a human promoter in vivo by transcription factor Sp1. Molecualr and Cellular Biology. 1991;**11**(4):1935-1943

[34] Boyce FM, Anderson GM, Rusk CD, Freytag SO. Human argininosuccinate synthetase minigenes are subject to arginine-mediated repression but not to trans induction. Molecualr and Cellular Biology. 1986;**6**(4):1244-1252

[35] Hooft van Huijsduijnen R, Li XY, Black D, Matthes H, Benoist C, Mathis D. Co-evolution from yeast to mouse: cDNA cloning of the two NF-Y (CP-1/CBF) subunits. The EMBO Journal. 1990;**9**(10):3119-3127

[36] Takiguchi M, Mori M. In vitro analysis of the rat liver-type arginase promoter. Journal of Biological Chemistry. 1991;**266**(14):9186-9193

[37] Inoue Y, Hayhurst GP, Inoue J, Mori M, Gonzalez FJ. Defective ureagenesis in mice carrying a liver-specific disruption of hepatocyte nuclear factor 4a (HNF4a). HNF4a regulates ornithine transcarbamylase *in vivo*. Journal of Biological Chemistry. 2002;**277**(28):25257-25265

[38] Kimura T, Christoffels VM, Chowdhury S, Iwase K, Matsuzaki H, Mori M, et al. Hypoglycemia-associated hyperammonemia caused by impaired expression of ornithine cycle enzyme genes in C/EBPa knockout mice. Journal of Biological Chemistry. 1998;**273**(42):27505-27510

[39] Kimura T, Chowdhury S, Tanaka T, Shimizu A, Iwase K, Oyadomari S, et al. CCAAT/enhancer-binding protein beta is required for activation of genes for ornithine cycle enzymes by glucocorticoids and glucagon in primary-cultured hepatocytes. FEBS Letters. 2001;**494**(1-2):105-111

[40] Felipo V, Butterworth RF. Neurobiology of ammonia. Progress in Neurobiology. 2002;**67**(4):259-279

[41] Martinez de Lapiscina EH, Aguirre ME, Giadas TC, Alvarez MT. 5-Fluorouracil induced hyperammonemic encephalophathy: Etiopathologic correlation. Canadian Journal of Neurological Sciences. 2012;**39**(4):553-554

[42] Nurmohamed S, Weenink A, Moeniralam H, Visser C, Bemelman F. Hyperammonemia in generalized *Mycobacterium genavense* infection after renal transplantation. American Journal of Transplantation. 2007;7(3):722-723. DOI: 10.1111/j.1600-6143.2006.01680.x

[43] Yeh KH, Cheng AL. High-dose 5-fluorouracil infusional therapy is associated with hyperammonaemia, lactic acidosis and encephalopathy. British Journal of Cancer. 1997;**75**(3):464-465

[44] Metzenberg RL, Hall LM, Marshall M, Cohen PP. Studies on the biosynthesis of carbamyl phosphate. Journal of Biological Chemistry. 1957;**229**(2):1019-1025

[45] Caldovic L, Morizono H, Gracia Panglao M, Gallegos R, Yu X, Shi D, et al. Cloning and expression of the human N-acetylglutamate synthase gene. Biochemical and

Biophysical Research Communications. 2002;**299**(4):581-586

[46] Caldovic L, Morizono H, Yu X, Thompson M, Shi D, Gallegos R, et al. Identification, cloning and expression of the mouse N-acetylglutamate synthase gene. Biochemical Journal. 2002;**364**(Pt 3):825-831

[47] Elpeleg O, Shaag A, Ben-Shalom E, Schmid T, Bachmann C. N-acetylglutamate synthase deficiency and the treatment of hyperammonemic encephalopathy. Annals of Neurology. 2002; **52**(6):845-849

[48] Haberle J, Schmidt E, Pauli S, Kreuder JG, Plecko B, Galler A, et al. Mutation analysis in patients with N-acetylglutamate synthase deficiency. Human Mutation. 2003;**21**(6):593-597

[49] Heibel SK, Ah Mew N, Caldovic L, Daikhin Y, Yudkoff M, Tuchman M. N-carbamylglutamate enhancement of ureagenesis leads to discovery of a novel deleterious mutation in a newly defined enhancer of the NAGS gene and to effective therapy. Human Mutation. 2011;**32**(10):1153-1160. DOI: 10.1002/humu.21553

[50] Heibel SK, Lopez GY, Panglao M, Sodha S, Marino-Ramirez L, Tuchman M, et al. Transcriptional regulation of N-acetylglutamate synthase. PLoS One. 2012;**7**(2):e29527. DOI: 10.1371/journal.pone.0029527

[51] Haberle J, Denecke J, Schmidt E, Koch HG. Diagnosis of N-acetylglutamate synthase deficiency by use of cultured fibroblasts and avoidance of nonsense-mediated mRNA decay. Journal of Inherited Metabolic Disease. 2003;**26**(6):601-605

[52] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. Journal of Molecular Biology. 1990;**215**(3):403-410

[53] Frith MC, Fu Y, Yu L, Chen JF, Hansen U, Weng Z. Detection of functional DNA motifs via statistical over-representation. Nucleic Acids Research. 2004;**32**(4):1372-1381. DOI: 10.1093/nar/gkh299 32/4/1372

[54] Nyunoya H, Broglie KE, Widgren EE, Lusty CJ. Characterization and derivation of the gene coding for mitochondrial carbamyl phosphate synthetase I of rat. Journal of Biological Chemistry. 1985;**260**(16):9346-9356

[55] Christoffels VM, Habets PE, Das AT, Clout DE, van Roon MA, Moorman AF, et al. A single regulatory module of the carbamoylphosphate synthetase I gene executes its hepatic program of expression. The Journal of Biological Chemistry. 2000;**275**(51):40020-40027. DOI: 10.1074/jbc.M007001200

[56] Christoffels VM, van den Hoff MJ, Lamers MC, van Roon MA, de Boer PA, Moorman AF, et al. The upstream regulatory region of the carbamoyl-phosphate synthetase I gene controls its tissue-specific, developmental, and hormonal regulation in vivo. Journal of Biological Chemistry. 1996;**271**(49):31243-31250

[57] van den Hoff MJ, van de Zande LP, Dingemanse MA, Das AT, Labruyere W, Moorman AF, et al. Isolation and characterization of the rat gene for carbamoylphosphate synthetase I. European Journal of Biochemistry. 1995;**228**(2):351-361

[58] Goping IS, Lamontagne S, Shore GC, Nguyen M. A gene-type-specific enhancer regulates the carbamyl phosphate synthetase I promoter by cooperating with the proximal GAG activating element. Nucleic Acids Research. 1995;**23**(10):1717-1721

[59] Lagace M, Howell BW, Burak R, Lusty CJ, Shore GC. Rat carbamyl-phosphate synthetase I gene. Promoter sequence and tissue-specific

transcriptional regulation in vitro. Journal of Biological Chemistry. 1987;**262**(22):10415-10418

[60] Christoffels VM, van den Hoff MJ, Moorman AF, Lamers WH. The far-upstream enhancer of the carbamoyl-phosphate synthetase I gene is responsible for the tissue specificity and hormone inducibility of its expression. Journal of Biological Chemistry. 1995;**270**(42):24932-24940

[61] Hoogenkamp M, Stallen JM, Lamers WH, Gaemers IC. In vivo footprinting of the carbamoylphosphate synthetase I cAMP-response unit indicates important roles for FoxA and PKA in formation of the enhanceosome. Biochimie. 2006;**88**(10):1357-1366. DOI: 10.1016/j.biochi.2006.06. 009

[62] Schoneveld OJ, Gaemers IC, Das AT, Hoogenkamp M, Renes J, Ruijter JM, et al. Structural requirements of the glucocorticoid-response unit of the carbamoyl-phosphate synthase gene. Biochemical Journal. 2004;**382**(Pt 2):463-470. DOI: 10.1042/BJ20040471

[63] Schoneveld OJ, Gaemers IC, Hoogenkamp M, Lamers WH. The role of proximal-enhancer elements in the glucocorticoid regulation of carbamoylphosphate synthetase gene transcription from the upstream response unit. Biochimie. 2005;**87**(11):1033-1040. DOI: 10.1016/j.biochi. 2005.02.015

[64] Chen Z, Tang N, Wang X, Chen Y. The activity of the carbamoyl phosphate synthase 1 promoter in human liver-derived cells is dependent on hepatocyte nuclear factor 3-beta. Journal of Cellular and Molecular Medicine. 2017;**21**(9):2036-2045. DOI: 10.1111/jcmm.13123

[65] Hata A, Tsuzuki T, Shimada K, Takiguchi M, Mori M, Matsuda I. Isolation and characterization of the human ornithine transcarbamylase gene: Structure of the 5′-end region. Journal of Biochemistry. 1986;**100**(3):717-725

[66] Luksan O, Jirsa M, Eberova J, Minks J, Treslova H, Bouckova M, et al. Disruption of OTC promoter-enhancer interaction in a patient with symptoms of ornithine carbamoyltransferase deficiency. Human Mutation. 2010;**31**(4):E1294-E1303. DOI: 10.1002/humu.21215

[67] Veres G, Craigen WJ, Caskey CT. The 5′ flanking region of the ornithine transcarbamylase gene contains DNA sequences regulating tissue-specific expression. Journal of Biological Chemistry. 1986;**261**(17):7588-7591

[68] Takiguchi M, Murakami T, Miura S, Mori M. Structure of the rat ornithine carbamoyltransferase gene, a large, X chromosome-linked gene with an atypical promoter. Proceedings of the National Academy of Sciences of the United States of America. 1987;**84**(17):6136-6140

[69] Murakami T, Takiguchi M, Inomoto T, Yamamura K, Mori M. Tissue- and developmental stage-specific expression of the rat ornithine carbamoyltransferase gene in transgenic mice. Developmental Genetics. 1989;**10**(5):393-401. DOI: 10.1002/dvg.1020100507

[70] Jones SN, Grompe M, Munir MI, Veres G, Craigen WJ, Caskey CT. Ectopic correction of ornithine transcarbamylase deficiency in sparse fur mice. Journal of Biological Chemistry. 1990;**265**(24):14684-14690

[71] Jang YJ, LaBella AL, Feeney TP, Braverman N, Tuchman M, Morizono H, et al. Disease-causing mutations in the promoter and enhancer of the ornithine transcarbamylase gene. Human Mutation. 2018;**39**(4):527-536. DOI: 10.1002/humu.23394

[72] Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of nonneutral substitution rates on mammalian phylogenies. Genome Research. 2010;**20**(1):110-121. DOI: 10.1101/gr.097857.109

[73] Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Research. 2005;**15**(8):1034-1050. DOI: 10.1101/gr.3715005

[74] Consortium EP. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;**489**(7414):57-74. DOI: 10.1038/nature11247

[75] Yue F, Cheng Y, Breschi A, Vierstra J, Wu W, Ryba T, et al. A comparative encyclopedia of DNA elements in the mouse genome. Nature. 2014;**515**(7527):355-364. DOI: 10.1038/nature13992

[76] Liu T, Ortiz JA, Taing L, Meyer CA, Lee B, Zhang Y, et al. Cistrome: An integrative platform for transcriptional regulation studies. Genome Biology. 2011;**12**(8):R83. DOI: 10.1186/gb-2011-12-8-r83

[77] Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD, Collins FS, et al. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. Proceedings of the National Academy of Sciences of the United States of America. 2002;**99**(26):16899-16903. DOI: 10.1073/pnas.242603899

[78] Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-based map of the human proteome. Science. 2015;**347**(6220):1260419. DOI: 10.1126/science.1260419

[79] Consortium GT, Laboratory DA, Coordinating Center–Analysis Working G, Statistical Methods groups-Analysis Working G, Enhancing Gg, Fund NIHC et al. Genetic effects on gene expression across human tissues. Nature. 2017;**550**(7675):204-213. DOI: 10.1038/nature24277

[80] Dawson MI, Xia Z. The retinoid X receptors and their ligands. Biochimica et Biophysica Acta. 2012;**1821**(1):21-56. DOI: 10.1016/j.bbalip.2011.09.014

Chapter 6

Recent Advances about Local Gene Delivery by Ultrasound

Zhiyi Chen, Meng Du and Fei Yan

Abstract

Gene therapy has been widely explored as a pharmacological approach, with a great potential to treat various diseases. Generally, many diseases have definite lesion's site, especially for tumors. This feature results in a great demand on the delivery of therapeutic gene to the local lesion's site. Ultrasound combined with microbubbles provides a promising platform to deliver gene in a spatiotemporally controlled way. Ultrasound beam can be positioned and targeted onto the deep-seated lesion's site of diseases by an external mobile transducer. Microbubbles can serve as vehicles for carrying genetic cargo and can be destructed by ultrasound, resulting in the local release of genetic payload. Meanwhile, sonoporation effect will occur upon which the bubbles are exposed to the appropriate ultrasonic energy, producing the transient small holes on the adjacent cell membrane and thus increasing the vascular and cellular permeability. In this chapter, we will review the recent advances about local gene delivery by ultrasound.

Keywords: ultrasound, gene therapy, microbubbles, sonoporation

1. Introduction

Gene therapy, designed to deliver nucleic acid into cells to compensate for abnormal genes, is now considered a promising treatment option for some human diseases [1]. With the development of modern medicine and precise medicine, there is an increasing trend to change the traditional gene delivery mode into local gene delivery. At present, there are mainly two gene delivery approaches, virus-mediated transfection and nonvirus-mediated transfection [2]. The former method has high-transfection efficiency, but the preparation procedure of recombinant viruses is sophisticated, and their clinical application is restricted due to biosafety concerns [3]. Nonviral vector approaches, such as liposome-mediated methods and electroporation techniques, are relatively safe. However, poor targeting and low-transfection efficiencies limit their widespread use [4]. It is a current research hotspot to look for an effective and safe method to mediate gene delivery for biomedical application.

Ultrasound is a widely used diagnostic technique in clinic, which possesses the advantages of safety, real-time monitoring, and low cost. Recently, with the development of ultrasound contrast agents, ultrasound has evolved from a diagnostic tool to a treatment application for delivering locally therapeutic substances into the lesion's sites. Ultrasound-targeted microbubble destruction (UTMD) provides a promising platform to deliver genes in a spatiotemporally controlled way.

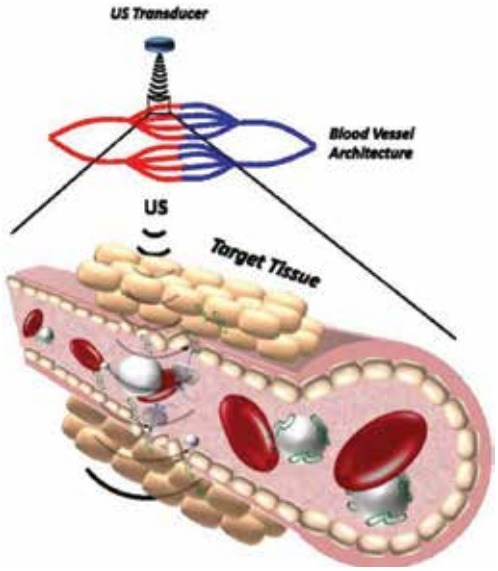

Figure 1.
Schematic model of ultrasound-mediated gene delivery. Bioeffect produced by ultrasound and microbubble interaction could enhance the permeability of vascular and promote the accumulation of gene (green) in tissue. (Quoted from: Sirsi and Borden [5]).

Microbubbles can serve as vehicles for carrying genetic cargo and can be destructed by ultrasound, resulting in the local release of genetic payload. Meanwhile, sonoporation effect will occur upon which the bubbles are exposed to the appropriate ultrasonic energy, producing the transient small holes on the adjacent cell membrane and thus increasing the vascular and cellular permeability **Figure 1**. In this chapter, we will briefly introduce the mechanism and review the recent advances about local gene delivery by ultrasound.

2. Mechanism of ultrasound-mediated gene transfection

2.1 Sonoporation

When ultrasound is irradiated locally with certain energy, the cavitation nuclei, such as ultrasound contrast agents and bubbles, could alternately occur expansion, contraction, splitting, fusion, and even rupture. This physical process is called cavitation effect. Accompanied by the cavitation effect, acoustic microstreaming, micro-jet, high temperature, and shockwave will occur in the medium, resulting in the formation of some temporary, reversible pores on the cell membrane, which is sonoporation [6, 7]. Generally, it is an accepted notion that the sonoporation from cavitation effect allows genes and drugs to enter cells [8].

There are a large number of studies, which have confirmed that sonoporation can increase the efficiency of gene delivery through enhancing the permeability of the cell membrane [9–12]. The number of pores, having a high impact on the gene delivery efficiency, can be affected by a lot of factors, such as acoustic pressure, irradiation duration time, and pulse repetition frequency [13–15]. Sonoporation pores trend to be larger along with the increase of acoustic pressure and irradiation time, which also enhance gene transfection efficiency [16]. However, excessive acoustic pressure or ultrasonic duration may reduce cell viability and even cause cell death, vascular rupture, and other side effects [17–19]. Therefore, to achieve a

high gene transfection efficiency and remain a cell viability as much as possible, it is important to optimize ultrasound irradiation parameters during gene transfection.

2.2 Endocytosis

In addition to sonoporation, cavitation effect can change the cell membrane structure through microstreaming and shear force. The mechanical force may cause cytoskeleton rearrangement and regulate various downstream cellular signaling pathways, helping the endocytosis of genetic cargo [20, 21]. Generally, there are three forms of endocytosis, including macropinocytosis, clathrin-mediated endocytosis, and caveolae-mediated endocytosis [22]. After ultrasound irradiation, the reactive oxygen species are produced to stimulate the calcium influx and induce the occurrence of endocytosis [23]. In addition, cavitation effect and shear force induced by ultrasound can change cell structure and influence endocytosis through mechanosensors and signaling cascade [24]. Meijering et al. demonstrated that endocytosis was involved in the uptake of the macromolecular substances, while small molecules enter cells mainly through the pores of the membrane surface [25].

2.3 Sonoprinting

Recently, Cock et al. put forward a new viewpoint on the mechanism of ultrasound-mediated gene delivery [26]. By using the real-time scanning confocal microscopy, they found that nanoparticle-loaded microbubbles could deposit the nanoparticles in patches onto the cell membrane during ultrasound irradiation and promote the particles that enter cell through the fluidity of the membrane. In their opinion, this method, termed sonoprinting, is neither the traditional sonoporation nor the material swallowing. The underlying mechanisms still need to be explored.

3. Type of ultrasound contrast agents as gene vector

Genes administrated by the intravenous route are easily be degraded. Conventionally, genes such as plasmids, mRNA, siRNA, and miRNA need to be protected from degradation by extracellular and intracellular barriers **Figure 2**. The ideal gene vectors should have the following characteristics: (1) safe and nontoxic, long cycle time in vivo, protecting the nucleic acid molecules from being destroyed by extracellular nucleic acid enzymes; (2) possessing the characteristics of a targeting ability and delivering the gene to target tissue or target cells; (3) high gene-carrying capacity; (4) promoting the gene to enter cytoplasmic or nucleus and stable expressing; (5) ensuring the controllability of gene function; and (6) noninvasive evaluation of gene delivery effectiveness. In the field of ultrasound-mediated gene delivery, many ultrasound contrast agents, including microbubbles, nanobubbles, nanodroplets, and some nanoparticles, are being developed into gene vectors in gene delivery mediated by ultrasound.

The gene vector may help them to avoid degradation by extracellular and intracellular barriers, including serum endonucleases, immune detection, and endosome (Quoted from: Yin et al. [2]).

3.1 Microbubbles

Microbubbles are small, gas-filled microspheres with the particle size of 1–3 μm. As gene vectors, they not only can protect the genes from nucleic acid enzyme degradation and from reticuloendothelial system clearance but may also enhance

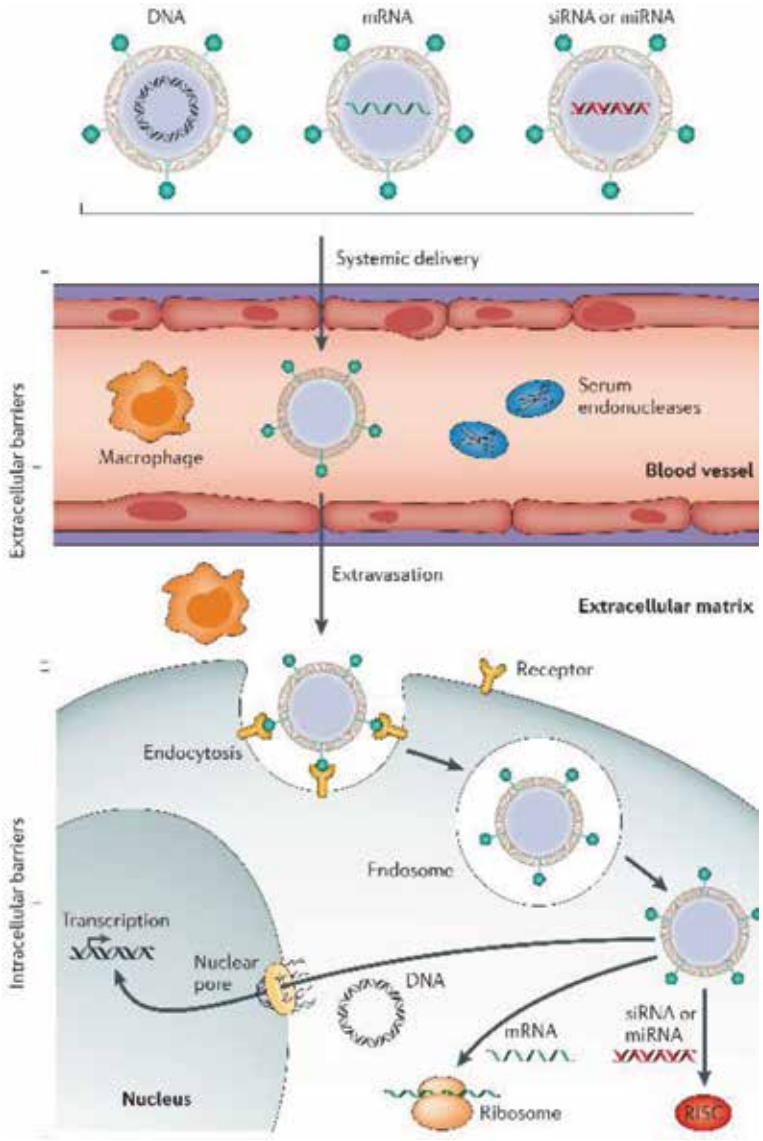

Figure 2.
Schematic model of transfection process of genes in carriers.

their local delivery through active and passive targeting. Traditional membrane materials consist of microbubbles, which include albumin, lipid, polymers, and surfactants. Different shell compositions have various characteristics. Albumin is commonly used in the preparation of commercial ultrasound contrast agents, but it is susceptible to degeneration due to temperature change. In addition, it is expensive and easy to cause immune response. The synthetic phospholipids are good and biocompatible, but their half life is short in vivo. Polymers are slightly inferior in biocompatibility, but it possesses better stability.

It has been proved that the application of ultrasound combined with commercial microbubbles and gene mixture could regulate gene expression and achieve therapeutic effect [27–30]. Wang et al. compared the effect of gene delivery by three kinds of typical commercial microbubbles—Optison, Sonovue, and Levovist. The mixture of microbubbles and plasmid DNA encoding green fluorescent protein was injected into tibialis anterior muscle of mice. After ultrasound irradiation,

the number of GFP-positive fibers was significantly increased in Optison- and Sonovue-treated groups, proving the efficiency of gene transfection by ultrasound combined with commercial microbubbles [31]. However, DNA is anionic molecules, and most microbubbles are negatively charged on the surface, which bring some difficulty for the formation of DNA/microbubble complexes. In order to address this issue, some cationic microbubbles are developed and applied as gene vector to enhance the gene-carrying capacity [32–37]. Wang et al. evaluated the difference of gene transfection rate between cationic microbubbles and neutral microbubbles in combination with ultrasound. Their results demonstrated that the expression of reporter gene in cationic microbubble group was 20-fold higher in vitro and 3-fold higher in tumor model than neutral bubbles [34]. Recently, Wei et al. applied Targesphere, a kind of commercial cationic microbubbles, as short hairpin (shRNA) vector for connective tissue growth factor (CTGF). It was showed that the expression of CTGF was decreased in renal fibrosis mouse model after ultrasound irradiation, which proved the great potential in gene delivery mediated by ultrasound combined with cationic microbubbles [38].

3.2 Nanoparticle, nanodroplet, and nanobubble

Nanoscale ultrasound contrast agents, with the particle size from 100 to 600 nm, are also developed in the recent years. Compared with traditional microbubbles, nanoscale contrast agents have smaller size and stronger penetrating ability. In addition, nanoscale contrast agents possess greater gene-carrying capacity due to their larger surface area. Common nanoscale ultrasound contrast agents include nanobubbles, solid nanoparticles, and liquid fluorocarbon nanoparticles. Most of the shell membrane of nanobubbles are lipid or polymer, and the core could be gas or liquid. Nanobubbles can cross through the blood vessels and aggregate in the tumors through the enhanced permeability and retention (EPR) effect [39]. It was proved that nanobubbles could achieve ideal gene transfection efficiency when combined with ultrasound [40, 41]. Horie et al. applied ultrasound combined with nanobubbles mediating tumor necrosis factor (TNF-α) DNA delivery to treat tumor-bearing mice and resulted in the decrease of the tumor vessel density and inhibition of tumor growth [42]. To enhance the gene-carrying capacity and local transfection efficiency, cationic nanobubble or targeted nanobubbles have been applied and showed excellent therapeutic effect in vitro and in vivo [43–45]. Yin et al. developed a new kind of siRNA-nanobubble, through a nanoparticle heteroassembly of siRNA-loaded polymeric micelles and liposomes, demonstrating their ideal therapeutic effect in cancer treatment [46]. Xie et al. used cell-permeable peptides (CPPs) to enhance the transferring rate of siRNA. They developed CPP-siRNA that targets oncogene c-myc and encapsulated it into nanobubbles. It was shown that the expression of c-myc mRNA was significantly decreased, and the growth of tumor was significantly inhibited after ultrasound irradiation [47].

Recently, liquid fluorocarbon nanodroplets have attracted wide attentions in the ultrasound-mediated gene delivery. These nanodroplets prepared from a lipid or a polymer shell can encapsulate liquid fluorocarbon emulsion (perfluoropentane, etc.). The liquid core would occur "acoustic droplet vaporization" (ADV) under ultrasound irradiation, which makes the nanodroplet transform into gas-containing microbubbles, greatly enhancing the cavitation effect of ultrasound **Figure 3**. Although nanodroplets have shown its therapeutic effect in high-intensity focused ultrasound (HIFU) and drug delivery, its application in gene delivery is still rare. Gao et al. synthesized a novel tumor-targeting cationic nanodroplet and applied it as gene vector to treat Her2-positive breast cancer. The results in their study

demonstrated that this nanodroplet could achieve better gene transfection efficiency, showing its potential in gene delivery by ultrasound [48, 49].

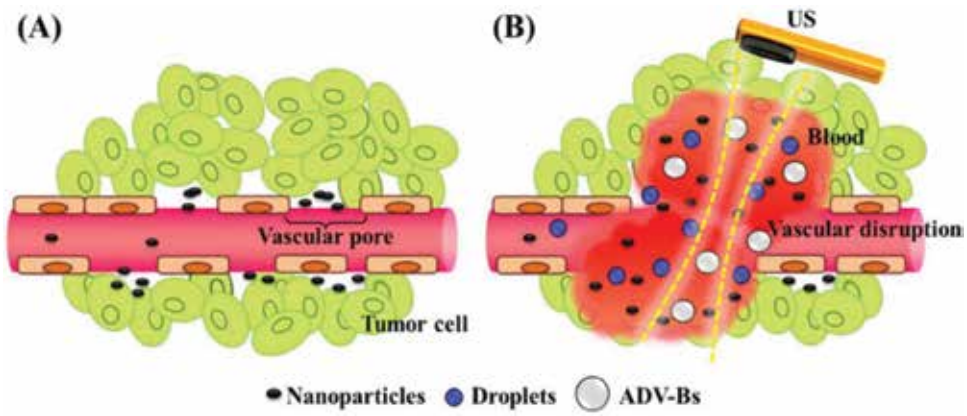

Figure 3.
Schematic model of acoustic droplet vaporization (ADV).

(A) Nanoparticles penetrate the tissue through the EPR effect; (B) droplets vaporize into microbubbles through ADV under certain acoustic pressure, which enhances the cavitation effect and changes the structure of tumor vessels. (Quoted from: Ho et al. [50]).

Nanoparticles commonly used in gene transfection include liposomes, polymer, and magnetic nanoparticles. Studies demonstrated that the cavitation effect produced by UTMD could increase the concentration of nanoparticles in targeted tissue and improved gene transfection efficiency. In the field of ultrasound-mediated gene delivery, liposome and polyethylenimine (PEI) are the most popular gene vectors.

Liposome is used as a nanocarrier for gene transfection, with high gene-carrying capacity and transfection efficiency. Taking advantages of UTMD, researchers have demonstrated that the accumulation of gene-carrying liposomes can be improved in targeting cells or tissue [51, 52]. Yoon et al. proved that ultrasound combined with microbubbles and gene-carrying liposomes could be a superior gene transfection system [53]. Recently, Chertok et al. modified heparin on the surface of liposome to increase the accumulation of gene in tumor site and reduce the off-target effect. Compared with nonheparinized DNA-carrying liposomes, modified liposomes combined with UTMD could significantly enhance the gene transfection rate in tumor in vivo [54].

PEI is another commonly used gene vector with high-density positive charge. It can form stable complex with genes through electrostatic adsorption. Also, utilization of PEI can avoid DNA degradation by nucleic acid enzyme and improve the stability and integrity of genes in vivo. Meanwhile, PEI can assist gene delivery into nuclei through proton sponge mechanism and endosomal escape, which will enhance the expression of targeting gene [55] **Figure 4**. However, the cell toxicity is inevitable because of its strongly positive charge. UTMD may function as an effective method to balance the cytotoxicity and transfection efficiency of PEI. UTMD could not only temporarily mediate the opening of cell membrane and promote the PEI-DNA complex entering the cell but also improve the level of intracellular calcium and PKC protein expression, which can enhance the effect of endocytosis. It was confirmed that UTMD combined with PEI or chemical modified PEI could be an effective and safety gene transfection strategy in vitro or in vivo [56, 57]. Dang et al.

demonstrated that UTMD combined with PEI could achieve the same transfection efficiency as Lipofectamine 2000 and lower cytotoxicity [58]. Deshpande et al. found that ultrasound combined with PEI could enhance the DNA transfection rate up to 200-fold than naked DNA plasmids [59]. Park et al. applied UTMD combined with PEI mediating the adenine nucleotide translocase-2 (ANT2) shRNA to successfully increase the survival rate of xenograft mice and induce the tumor regression [60].

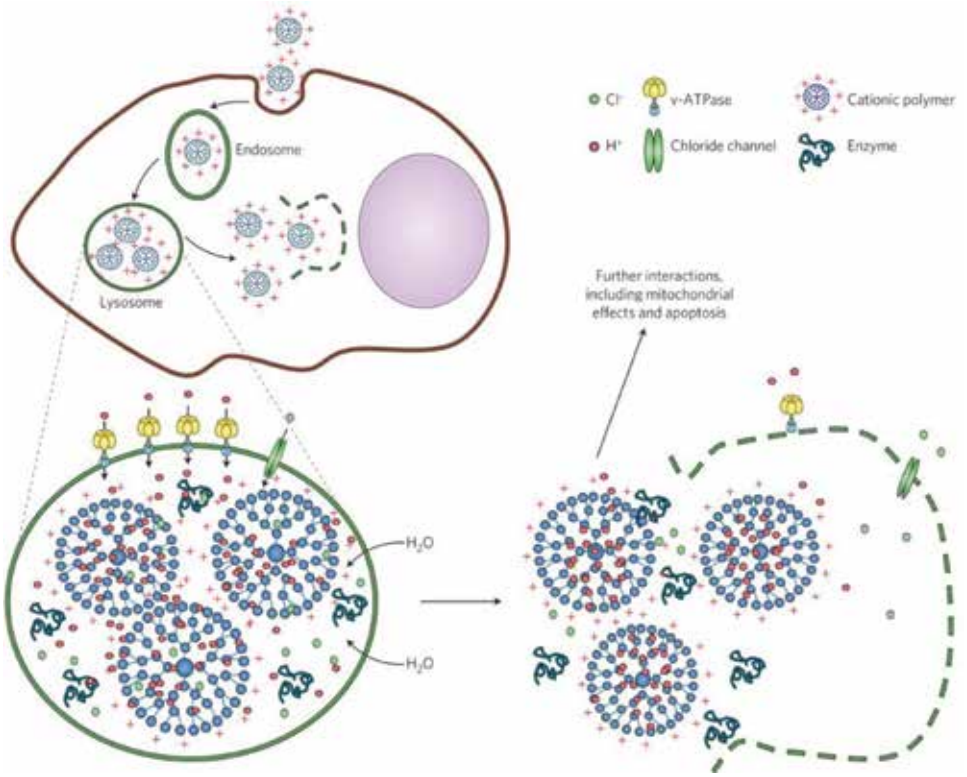

Figure 4.
Schematic model of the proton sponge effect by cationic nanoparticles.

PEI binds with cell membrane and is endocytosed. When they enter lysosome, the unsaturated amino groups are able to capture protons and cause the retention of Cl^- ion and water molecule, which will make lysosomal swelling and rupture, and then release the lysosomal content. (Quoted from: Nel et al. [55]).

4. Application of local gene delivery by ultrasound

4.1 Tumor

Tumor is a kind of genetically related disease. Its occurrence, development, and recurrence are closely related to the mutation and deletion of the gene. With the development of molecular biology, gene therapy has shown a great potential in cancer treatment. At present, the common strategy of gene therapy is to transfer tumor suppressor gene into tumor cells to restore normal phenotype of cells. Nande et al. applied UTMD to mediate tumor suppressor genes, including p53, Rb, p130, and significantly reduced tumor growth [61]. Chang et al. utilized the p53-loaded targeted microbubbles for ovarian cancer treatment and achieved higher transfection

efficiency than conventional nontargeted microbubbles [62]. Mishel et al. used ultrasound to mediate hSef-b delivery, another kind of human tumor suppressor gene, and demonstrated the efficacy of gene therapy mediated by ultrasound [63]. Recently, gene delivery by ultrasound was also applied in gene-directed enzyme prodrug therapy (GDEPT). The key process of GDEPT is effectively transferring gene encoding enzyme, which can convert a nontoxic prodrug into an activated cytotoxic agent [64]. Devulapally et al. used PEGylated-PLGA-PEI nanoparticles and mediated TK-NTR fusion gene delivery in tumor xenograft mice. Their results showed that the tumor size was reduced by 2.3-fold when compared with untreated mice [65].

In the field of tumor therapy, RNAi could selectively inhibit the expression of key genes in the development of cancer. It has been proved that co-delivery of siRNA and chemotherapy drugs by ultrasound could improve the therapeutic effect of tumor and reduce the dosage chemotherapy drugs [66–69]. Zhao et al. synthesized cationic porphyrin microbubbles for the delivery of FOXA1-siRNA, achieving an excellent therapeutic effect for breast cancer [70]. Cancer stem cells (CSCs), a group of tumor cells with self-renewal, multidirectional differentiation potential, are thought to be the key of tumor recurrence, metastasis, and drug resistance. Specific markers, such as CD133, are important targets for gene therapy. Liu et al. used UTMD to deliver shRNA-CD133 to liver CSCs and reversed the process of epithelial-mesenchymal transition [71].

4.2 Cardiovascular disease

Atherosclerosis is the main cause of coronary heart disease, cerebral infarction, and peripheral vascular disease. Studies demonstrated that ultrasound combined with microbubbles can deliver angiogenic genes to the ischemic region of the myocardium and enhance expression of angiogenesis-related factors and thus improve myocardial blood supply [72, 73]. Du et al. utilized UTMD and cationic microbubbles to mediate delivery of growth differentiation factor 11 (GDF11) plasmid to aged heart. Their results suggested that ultrasound could enhance GDF11 expression, increase the cardiac stem cell (CSC) proliferation, and rejuvenate the senescent heart from ischemic injury [74]. Castle et al. successfully enhanced the level of ApoA-I and high-density lipoprotein cholesterol (HDL-C) in vivo through delivering human apolipoprotein ApoA-I plasmids by ultrasound [75].

Heart failure, caused by various cardiac structures and functional disorders, will impair ventricular filling and ejection function and eventually cause cardiac output unable to meet body tissue metabolic needs. Lee et al. delivered survivin gene to cardiomyocyte by UTMD and observed its efficacy on cardiac function. The apoptosis rate of cadiomyocyte was significantly decreased, and the left ventricular systolic dysfunction was attenuated after 6 weeks, demonstrating that ultrasound-mediated gene delivery can be an effective treatment in heart failure [76].

4.3 Central nervous system diseases

Blood-brain barrier (BBB) is an important obstacle for central nervous system (CNS) diseases. BBB is mainly composed of cerebral capillary endothelial cells and their cells, matrix, astrocytes, and extracellular matrix [77]. To cross the BBB, researchers have tried various methods, including invasive surgery, hypertonic drugs, chemical modification of drugs to target delivery to brain, and micro-carriers [78–80]. Recently, ultrasound mediating BBB opening has attracted researchers' attention due to its characteristic of noninvasive, reversible, and targeted delivery. Hynymen et al. proved that microbubbles could be applied as cavitation nuclei to reduce the ultrasonic energy to open the BBB, reducing the risk of tissue damage and bleeding [81]. Based on this, numerous studies are exploring the therapeutic

effect of gene delivery mediated by ultrasound in CNS diseases [82, 83], such as glioma, Parkinson's disease, and Alzheimer's disease.

Glioma is the most common malignant tumor of the central nervous system. UTMD has a wide application prospect in the treatment of brain glioma. In 2016, Carpentier et al. developed an implantable ultrasonic irradiation system, named SonoCloud. They used this system to open the local area of BBB with microbubbles. In their study, 15 patients with recurrent brain glioma were selected to test the therapeutic effect of UTMD-mediated BBB opening. After intravenous administration of carboplatin and Sonovue combined with ultrasound treatment, it was proved that the BBB could be safely opened, and 9 of 15 patients showed no further tumor growth [84]. Fan et al. applied cationic microbubbles as therapeutic gene vectors and effectively mediated BBB opening for gene delivery in vitro and in vivo [85, 86]. Zhao et al. used targeted liposomes (NGR-liposomes) as vector for shRNA-Birc5 delivery and demonstrated the enhancement of local gene transfection and the inhibition of glioma progression [87].

Parkinson's disease (PD) is a common neurodegenerative disease of the nervous system due to the degeneration and death of dopaminergic neurons in substantia nigra and the significant decrease of dopamine content in striatum. Glial cell line-derived neurotrophic factor (GDNF) can protect the dopaminergic neurons and promote the regeneration of dopamine system in black striatum [88]. Fan et al. restored behavioral function in a PD animal model through delivering GDNF gene by transcranial focused ultrasound [89]. Lin et al. used the GDNF-loaded liposome-microbubble complexes and demonstrated the therapeutic effect of PD by using focused ultrasound-mediated BBB opening [90, 91].

In addition, ultrasound-mediated gene delivery was also applied in other CNS diseases. Song et al. developed PLGA nanobubbles for NGF delivery. NGF expression was significantly enhanced, and neuronal apoptosis in injured spinal cords was inhibited after ultrasound irradiation [92]. Wang et al. demonstrated that UTMD could successfully mediate VEGF gene delivery into brain and decreased infarct areas in a cerebral ischemic injury model [93].

4.4 Musculoskeletal disease

Arthritis is a common chronic inflammatory disease. Of these, the most common type is osteoarthritis and rheumatoid arthritis (RA). At present, the main treatment of arthritis is drug, including nonsteroidal anti-inflammatory drugs, cytotoxic drugs, and hormones. However, there are some drawbacks such as low local concentration and systemic side reaction. Ultrasound-mediated gene delivery has been proved to be effective in arthritis therapy. Xiang et al. applied UTMD-mediated enhanced green fluorescent protein (EGFP) gene delivery in antigen-induced arthritis rabbit model, and the significantly enhanced expression remained detectable for 40 days in the synovial pannus [94]. Tumor necrosis factor α (TNFα) secreted by synovial fibroblasts plays an important role in the progression of RA, which can cause bone destruction and joint dysfunction. Inue et al. transferred siRNA-TNFα to the articular synovial membrane of the rat through UTMD technique. They found that the expression of TNFα was inhibited, resulting in a significant remission of paw swallowing in comparison to control group [95].

In the field of fracture healing, bone morphogenetic protein-2 (BMP-2) is an ideal osteoinduction factor, which possesses the function of inducing cartilage and bone formation [96]. Some studies have confirmed that the transfection rate of BMP-2 gene in skeletal muscle cells and fibroblast cells could be enhanced by UTMD [97]. Osawa et al. delivered BMP-2 gene to the skeletal muscle in vivo, confirming the therapeutic effect of UTMD mediating gene transfection [98].

Tendon injury is a common disease in orthopedics with a significant impact on the quality of patient's life. Regulation of the expression of local cytokines in Achilles tendon by gene therapy is a potential therapeutic method to improve the prognosis of patient. Studies demonstrated that UTMD could increase the expression of genes in the Achilles tendon [99–101]. For example, Tang et al. transfected injured Achilles tendons of mice with insulin-like growth factor-1 (IGF-1) cDNA, showing that the maximum load, stiffness, and ultimate stress of treated Achilles tendons were higher than control group [102]. Bez et al. transferred BMP-6 encoding DNA by UTMD in Yucatan mini-pigs, showing the significantly enhanced osteointegration of all pigs after 8 weeks [103].

4.5 Ocular disease

For the treatment of ocular diseases, the most common method of drug delivery is surface administration or systemic administration. However, due to the unique structure of the eye, traditional drugs are difficult to enter the posterior eye segment, causing low bioavailability of drugs. As for ultrasound mediated gene delivery in ocular diseases, recent researches mainly focus on the cornea, retinoblastoma, and retinal neovascularization.

Cornea is a transparent tissue without blood vessels, which is an ideal target tissue for gene therapy because of its superficial position, transparent organization, and easy observation. Sonoda et al. confirmed that UTMD could mediate eGFP gene transfection to cornea epithelial cells of rabbit. In their study, they injected plasmid and microbubbles into the cornea of the rabbit and irradiated the eyes with ultrasound. They found that the corneal cells with GFP-positive expression were distributed around the injected region. No obvious tissue damage was observed in their study [104]. To optimize the gene transfection efficiency, Yamashita et al. developed a novel lipid microbubble, composed of polyethylene glycol (PEG) modified liposomes and perfluoropropane gas, and achieved a 27% gene transfection rate [105].

In the retina, there is a biological barrier similar to the blood-brain barrier named blood-retinal barrier (BRB), which is composed of tight connection between retinal endothelial cells and retinal pigment epithelial (RPE) cells. The presence of BRB prevents most systemically administered genes entering the retina, reducing the effectiveness of treatment. Park et al. demonstrated that UTMD could mediate BRB reversible opening without retinal damage [106]. Some studies have confirmed the effect of UTMD-mediated gene delivery into retinal in vitro and in vivo [107–110].

Retinoblastoma (RB) is a common ocular malignancy. Local treatment not only can retain part of the vision but also reduce the toxic side effects. Luo et al. applied wild-type 53 (wtp53) as a therapeutic gene. The in vitro experiment showed that the apoptosis rate of RB cells was higher (25.58%) than control group after ultrasound treatment [111]. To prove the therapeutic effect of gene delivery by ultrasound in vivo, Gao et al. transferred both wtp53 and Rb94 by UTMD to treat tumor-bearing mice. RB tumor growth was significantly inhibited, along with the decrease of the level of vascular endothelial factor and microvessel density [112].

Retinal neovascularization (RNV) is caused by hypoxic-ischemic ocular fundus diseases, characterized by retinal fibrous hyperplasia, retinal detachment, and even loss of vision. It has been reported that endostatin can be used for treating RNV because of its excellent antiangiogenic effect [113]. Xu et al. significantly enhanced the expression of endostatin by using cationic microbubbles to deliver endostatin gene under ultrasound irradiation. As a result, the growth of human retinal vascular endothelial cell was inhibited, suggesting that endostatin gene delivery mediated by UTMD may be a useful tool for RNV therapy [110].

4.6 Nephropathy

The blood flow of the normal kidney accounts for one-fourth to one-fifth of the total circulating blood volume. Therefore, a large number of microbubbles could enter the kidney blood vessels, which could be applied for ultrasound contrast imaging or targeted treatment. Based on this feature, some researchers applied UTMD to deliver genes to treat nephropathy, including diabetic nephropathy, hypertensive nephropathy, and renal fibrosis. Zhang et al. found that UTMD could increase the renal interstitial capillary permeability in diabetic nephropathy rat models [114]. Transforming growth factor β (TGF-β) is the key cytokine to promote the development of renal fibrosis. It can induce apoptosis of the podocytes on glomerular filtration membranes and promote the activation and proliferation of interstitial fibroblasts through TGF-β/SMAD signaling pathway. Lan et al. enhanced the expression of Smad7 in rat unilateral ureteral obstruction model by ultrasound as the gene delivery system, greatly attenuating tubulointerstitial fibrosis [115]. The therapeutic effect of UTMD-mediated Smad7 gene delivery in renal fibrosis was also proved in Smad7 gene knockout mice [116], diabetic nephropathy model mice [117], and angiotensin II-mediated hypertensive nephropathy [118].

In addition to the TGF-β/SMAD signal pathway, researchers have also explored the use of other signaling pathways in the treatment of renal fibrosis. RAP1 is a small molecule G protein that participates in the regulation of cell proliferation, differentiation, and intercellular adhesion [119]. Xiao et al. treated diabetic model rats with Rap1 gene delivery by ultrasound and microbubble (Optison). It was demonstrated that this treatment could protect the mitochondrial function of renal tubules and reduce the interstitial fibrosis [120]. In diabetic nephropathy, Yiu et al. confirmed the therapeutic effect of Kallistatin, which possesses the function of antioxidative and anti-inflammatory. The glomerulosclerosis and renal fibrosis were attenuated, and the renal function was improved after Kallistatin gene delivery by UTMD [121].

Recently, RNAi combined with UTMD therapy has been applied to the treatment of renal diseases. miR-29b is low expression in diabetes [122] and can function as a therapeutic targeting [123]. Chen et al. delivered miR-29b in diabetic mice by ultrasound combined with SonoVue. The results showed that this treatment could inhibit the inflammation induced by NF-κB/p63 and delay the progress of renal fibrosis [28]. Zhong et al. found that the level of miR-21 is highly associated with the development of renal fibrosis in diabetic mice and effectively improved renal fibrosis and inflammation by using UTMD-mediated miR-21 shRNA delivery [29]. Wei et al. applied UTMD combined with shRNA-CTGF to treat mouse models of renal fibrosis; the level of CTGF was significantly lower; and the renal fibrosis was attenuated, accompanied by the reduction of TGF-β and Type I collagen [38].

5. Conclusion and prospect

With the development of ultrasound contrast agents and the understanding of the biological effects of ultrasound, ultrasound-mediated gene delivery has been proven the great potential in the treatment of various diseases. Ultrasound contrast agents, including microbubbles, nanoparticles, and nanobubbles, can be used as gene vectors through intravenous or local injection into lesion site. With ultrasound irradiation at a certain level of acoustic intensity, the cavitation effect, sonoporation, and thermal effects occur, which can enhance the permeability of local tissue and promote the gene delivery into the pathological tissue.

Although ultrasound-mediated gene delivery has a broad application in animal study, there is still a long way for its application in human body. The main

problems, which need to be solved, may include the following aspects. First, many cationic materials are applied for the preparation of the ultrasound contrast agents. They have high gene-carrying capacity, but their biocompatibility is still doubtful. Second, ultrasound security is also an important concern. Unlike diagnostic ultrasound energy, the intensity applying in gene delivery is greater. Studies have shown that severe cavitation effects can lead to membrane rupture, DNA rupture, nuclear fragmentation, endothelial cell damage, microvascular leakage, hemolysis, myocardial injury, and even left ventricular function [19, 124]. More investigations need to be made to optimize the ultrasonic parameters so as to maximize the gene transfection efficiency and reduce the adverse side effects on the normal tissues and organs. In addition, the different types of ultrasound equipment used in various laboratories also bring some difficulties for the repeatability, which hinder the progress of ultrasound-mediated gene transfection technology to some degree. At the same time, it is believed that ultrasound will make more progress in gene delivery and bring about greater medical revolution in the future.

Acknowledgements

This work is supported by National Natural Science Foundation of China (Nos. 81671707, 81571701, and 81530056), Natural Science Foundation of Guangdong Province (Nos. 2017A030313651 and 2016A030311054), Research Projects of Guangzhou Science Technology and Innovation Commission (No. 201607010201), and Research Fund for Lin He's Academician Workstation of New Medicine and Clinical Translation.

Conflict of interest

The authors declare no competing financial interest.

Author details

Zhiyi Chen[1], Meng Du[1] and Fei Yan[2*]

1 Department of Ultrasound Medicine, Laboratory of Ultrasound Molecular Imaging, The Third Affiliated Hospital of Guangzhou Medical University, Guangzhou, P. R. China

2 Paul C. Lauterbur Research Center for Biomedical Imaging, Institute of Biomedical and Health Engineering, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, P. R. China

*Address all correspondence to: fei.yan@siat.ac.cn

References

[1] Dunbar CE, High KA, Joung JK, Kohn DB, Ozawa K, Sadelain M. Gene therapy comes of age. Science. 2018;**359**. DOI: 10.1126/science.aan4672

[2] Yin H, Kanasty RL, Eltoukhy AA, Vegas AJ, Dorkin JR, Anderson DG. Non-viral vectors for gene-based therapy. Nature Reviews. Genetics. 2014;**15**:541-555. DOI: 10.1038/nrg3763

[3] Thomas CE, Ehrhardt A, Kay MA. Progress and problems with the use of viral vectors for gene therapy. Nature Reviews. Genetics. 2003;**4**:346-358. DOI: 10.1038/nrg1066

[4] Glover DJ, Lipps HJ, Jans DA. Towards safe, non-viral therapeutic gene expression in humans. Nature Reviews. Genetics. 2005;**6**:299-310. DOI: 10.1038/nrg1577

[5] Sirsi SR, Borden MA. Advances in ultrasound mediated gene therapy using microbubble contrast agents. Theranostics. 2012;**2**:1208-1222. DOI: 10.7150/thno.4306

[6] Deng CX, Sieling F, Pan H, Cui J. Ultrasound-induced cell membrane porosity. Ultrasound in Medicine & Biology. 2004;**30**:519-526. DOI: 10.1016/j.ultrasmedbio.2004.01.005

[7] Lentacker I, De Cock I, Deckers R, De Smedt SC, Moonen CT. Understanding ultrasound induced sonoporation: Definitions and underlying mechanisms. Advanced Drug Delivery Reviews. 2014;**72**:49-64. DOI: 10.1016/j.addr.2013.11.008

[8] Yu H, Xu L. Cell experimental studies on sonoporation: State of the art and remaining problems. Journal of Controlled Release. 2014;**174**:151-160. DOI: 10.1016/j.jconrel.2013.11.010

[9] Park J, Fan Z, Deng CX. Effects of shear stress cultivation on cell membrane disruption and intracellular calcium concentration in sonoporation of endothelial cells. Journal of Biomechanics. 2011;**44**:164-169. DOI: 10.1016/j.jbiomech.2010.09.003

[10] Karshafian R, Bevan PD, Williams R, Samac S, Burns PN. Sonoporation by ultrasound-activated microbubble contrast agents: Effect of acoustic exposure parameters on cell membrane permeability and cell viability. Ultrasound in Medicine & Biology. 2009;**35**:847-860. DOI: 10.1016/j.ultrasmedbio.2008.10.013

[11] Mayer CR, Geis NA, Katus HA, Bekeredjian R. Ultrasound targeted microbubble destruction for drug and gene delivery. Expert Opinion on Drug Delivery. 2008;**5**:1121-1138. DOI: 10.1517/17425247.5.10.1121

[12] Suzuki R, Oda Y, Utoguchi N, Maruyama K. Progress in the development of ultrasound-mediated gene delivery systems utilizing nano- and microbubbles. Journal of Controlled Release. 2011;**149**:36-41. DOI: 10.1016/j.jconrel.2010.05.009

[13] Fan Z, Liu H, Mayer M, Deng CX. Spatiotemporally controlled single cell sonoporation. Proceedings of the National Academy of Sciences of the United States of America. 2012;**109**:16486-16491. DOI: 10.1073/pnas.1208198109

[14] Qin P, Xu L, Han T, Du L, Yu AC. Effect of non-acoustic parameters on heterogeneous sonoporation mediated by single-pulse ultrasound and microbubbles. Ultrasonics Sonochemistry. 2016;**31**:107-115. DOI: 10.1016/j.ultsonch.2015.12.001

[15] De Cock I, Zagato E, Braeckmans K, Luan Y, de Jong N, et al. Ultrasound and microbubble mediated drug delivery: Acoustic pressure as determinant

for uptake via membrane pores or endocytosis. Journal of Controlled Release. 2015;**197**:20-28. DOI: 10.1016/j.jconrel.2014.10.031

[16] Qiu Y, Zhang C, Tu J, Zhang D. Microbubble-induced sonoporation involved in ultrasound-mediated DNA transfection in vitro at low acoustic pressures. Journal of Biomechanics. 2012;**45**:1339-1345. DOI: 10.1016/j.jbiomech.2012.03.011

[17] Miller DL, Song J. Tumor growth reduction and DNA transfer by cavitation-enhanced high-intensity focused ultrasound in vivo. Ultrasound in Medicine & Biology. 2003;**29**:887-893

[18] Hwang JH, Brayman AA, Reidy MA, Matula TJ, Kimmey MB, Crum LA. Vascular effects induced by combined 1-MHz ultrasound and microbubble contrast agent treatments in vivo. Ultrasound in Medicine & Biology. 2005;**31**:553-564. DOI: 10.1016/j.ultrasmedbio.2004.12.014

[19] Vancraeynest D, Havaux X, Pasquet A, Gerber B, Beauloye C, et al. Myocardial injury induced by ultrasound-targeted microbubble destruction: Evidence for the contribution of myocardial ischemia. Ultrasound in Medicine & Biology. 2009;**35**:672-679. DOI: 10.1016/j.ultrasmedbio.2008.10.005

[20] Davies PF. Flow-mediated endothelial mechanotransduction. Physiological Reviews. 1995;**75**:519-560. DOI: 10.1152/physrev.1995.75.3.519

[21] Mizrahi N, Zhou EH, Lenormand G, Krishnan R, Weihs D, et al. Low intensity ultrasound perturbs cytoskeleton dynamics. Soft Matter. 2012;**8**:2438-2443. DOI: 10.1039/C2SM07246G

[22] Vercauteren D, Vandenbroucke RE, Jones AT, Rejman J, Demeester J, et al. The use of inhibitors to study endocytic pathways of gene carriers: Optimization and pitfalls. Molecular Therapy. 2010;**18**:561-569. DOI: 10.1038/mt.2009.281

[23] Lariccia V, Fine M, Magi S, Lin MJ, Yaradanakul A, et al. Massive calcium-activated endocytosis without involvement of classical endocytic proteins. The Journal of General Physiology. 2011;**137**:111-132. DOI: 10.1085/jgp.201010468

[24] Lionetti V, Fittipaldi A, Agostini S, Giacca M, Recchia FA, Picano E. Enhanced caveolae-mediated endocytosis by diagnostic ultrasound in vitro. Ultrasound in Medicine & Biology. 2009;**35**:136-143. DOI: 10.1016/j.ultrasmedbio.2008.07.011

[25] Meijering BD, Juffermans LJ, van Wamel A, Henning RH, Zuhorn IS, et al. Ultrasound and microbubble-targeted delivery of macromolecules is regulated by induction of endocytosis and pore formation. Circulation Research. 2009;**104**:679-687. DOI: 10.1161/CIRCRESAHA.108.183806

[26] De Cock I, Lajoinie G, Versluis M, De Smedt SC, Lentacker I. Sonoprinting and the importance of microbubble loading for the ultrasound mediated cellular delivery of nanoparticles. Biomaterials. 2016;**83**:294-307. DOI: 10.1016/j.biomaterials.2016.01.022

[27] Zolochevska O, Xia X, Williams BJ, Ramsay A, Li S, Figueiredo ML. Sonoporation delivery of interleukin-27 gene therapy efficiently reduces prostate tumor cell growth in vivo. Human Gene Therapy. 2011;**22**:1537-1550. DOI: 10.1089/hum.2011.076

[28] Chen HY, Zhong X, Huang XR, Meng XM, You Y, et al. MicroRNA-29b inhibits diabetic nephropathy in db/db mice. Molecular Therapy. 2014;**22**:842-853. DOI: 10.1038/mt.2013.235

[29] Zhong X, Chung AC, Chen HY, Dong Y, Meng XM, et al. miR-21 is a key therapeutic target for renal injury in a mouse model of type 2 diabetes. Diabetologia. 2013;**56**:663-674. DOI: 10.1007/s00125-012-2804-x

[30] Duvshani-Eshet M, Adam D, Machluf M. The effects of albumin-coated microbubbles in DNA delivery mediated by therapeutic ultrasound. Journal of Controlled Release. 2006;**112**:156-166. DOI: 10.1016/j.jconrel.2006.02.013

[31] Wang X, Liang HD, Dong B, Lu QL, Blomley MJ. Gene transfer with microbubble ultrasound and plasmid DNA into skeletal muscle of mice: Comparison between commercially available microbubble contrast agents. Radiology. 2005;**237**:224-229. DOI: 10.1148/radiol.2371040805

[32] Tlaxca JL, Anderson CR, Klibanov AL, Lowrey B, Hossack JA, et al. Analysis of in vitro transfection by sonoporation using cationic and neutral microbubbles. Ultrasound in Medicine & Biology. 2010;**36**:1907-1918. DOI: 10.1016/j.ultrasmedbio.2010.05.014

[33] Sun L, Huang CW, Wu J, Chen KJ, Li SH, et al. The use of cationic microbubbles to improve ultrasound-targeted gene delivery to the ischemic myocardium. Biomaterials. 2013;**34**:2107-2116. DOI: 10.1016/j.biomaterials.2012.11.041

[34] Wang DS, Panje C, Pysz MA, Paulmurugan R, Rosenberg J, et al. Cationic versus neutral microbubbles for ultrasound-mediated gene delivery in cancer. Radiology. 2012;**264**:721-732. DOI: 10.1148/radiol.12112368

[35] Panje CM, Wang DS, Pysz MA, Paulmurugan R, Ren Y, et al. Ultrasound-mediated gene delivery with cationic versus neutral microbubbles: Effect of DNA and microbubble dose on in vivo transfection efficiency. Theranostics. 2012;**2**:1078-1091. DOI: 10.7150/thno.4240

[36] Negishi Y, Endo-Takahashi Y, Matsuki Y, Kato Y, Takagi N, et al. Systemic delivery systems of angiogenic gene by novel bubble liposomes containing cationic lipid and ultrasound exposure. Molecular Pharmaceutics. 2012;**9**:1834-1840. DOI: 10.1021/mp200554c

[37] Zhang L, Sun Z, Ren P, You M, Zhang J, et al. Localized delivery of shRNA against PHD2 protects the heart from acute myocardial infarction through ultrasound-targeted cationic microbubble destruction. Theranostics. 2017;**7**:51-66. DOI: 10.7150/thno.16074

[38] Wei S, Xu C, Rychak JJ, Luong A, Sun Y, et al. Short hairpin RNA knockdown of connective tissue growth factor by ultrasound-targeted microbubble destruction improves renal fibrosis. Ultrasound in Medicine & Biology. 2016;**42**:2926-2937. DOI: 10.1016/j.ultrasmedbio.2016.07.022

[39] Guvener N, Appold L, de Lorenzi F, Golombek SK, Rizzo LY, et al. Recent advances in ultrasound-based diagnosis and therapy with micro- and nanometer-sized formulations. Methods. 2017;**130**:4-13. DOI: 10.1016/j.ymeth.2017.05.018

[40] Cavalli R, Bisazza A, Trotta M, Argenziano M, Civra A, et al. New chitosan nanobubbles for ultrasound-mediated gene delivery: Preparation and in vitro characterization. International Journal of Nanomedicine. 2012;**7**:3309-3318. DOI: 10.2147/IJN.S30912

[41] Zhang B, Chen M, Zhang Y, Chen W, Zhang L, Chen L. An ultrasonic nanobubble-mediated PNP/fludarabine suicide gene system: A new approach for the treatment of hepatocellular carcinoma. PLoS One. 2018;**13**:e196686. DOI: 10.1371/journal.pone.0196686

[42] Horie S, Watanabe Y, Ono M, Mori S, Kodama T. Evaluation of antitumor effects following tumor necrosis factor-alpha gene delivery using nanobubbles and ultrasound. Cancer Science. 2011;**102**:2082-2089. DOI: 10.1111/j.1349-7006.2011.02056.x

[43] Wu M, Zhao H, Guo L, Wang Y, Song J, et al. Ultrasound-mediated nanobubble destruction (UMND) facilitates the delivery of A10-3.2 aptamer targeted and siRNA-loaded cationic nanobubbles for therapy of prostate cancer. Drug Delivery. 2018;**25**:226-240. DOI: 10.1080/10717544.2017.1422300

[44] Song Z, Ye Y, Zhang Z, Shen J, Hu Z, et al. Noninvasive, targeted gene therapy for acute spinal cord injury using LIFU-mediated BDNF-loaded cationic nanobubble destruction. Biochemical and Biophysical Research Communications. 2018;**496**:911-920. DOI: 10.1016/j.bbrc.2018.01.123

[45] Xie X, Lin W, Li M, Yang Y, Deng J, et al. Efficient siRNA delivery using novel cell-penetrating peptide-siRNA conjugate-loaded nanobubbles and ultrasound. Ultrasound in Medicine & Biology. 2016;**42**:1362-1374. DOI: 10.1016/j.ultrasmedbio.2016.01.017

[46] Yin T, Wang P, Li J, Zheng R, Zheng B, et al. Ultrasound-sensitive siRNA-loaded nanobubbles formed by hetero-assembly of polymeric micelles and liposomes and their therapeutic effect in gliomas. Biomaterials. 2013;**34**:4532-4543. DOI: 10.1016/j.biomaterials.2013.02.067

[47] Xie X, Yang Y, Lin W, Liu H, Liu H, et al. Cell-penetrating peptide-siRNA conjugate loaded YSA-modified nanobubbles for ultrasound triggered siRNA delivery. Colloids and Surfaces. B, Biointerfaces. 2015;**136**:641-650. DOI: 10.1016/j.colsurfb.2015.10.004

[48] Gao D, Gao J, Xu M, Cao Z, Zhou L, et al. Targeted ultrasound-triggered phase transition nanodroplets for Her2-overexpressing breast cancer diagnosis and gene transfection. Molecular Pharmaceutics. 2017;**14**:984-998. DOI: 10.1021/acs.molpharmaceut.6b00761

[49] Gao D, Xu M, Cao Z, Gao J, Chen Y, et al. Ultrasound-triggered phase-transition cationic nanodroplets for enhanced gene delivery. ACS Applied Materials & Interfaces. 2015;**7**:13524-13537. DOI: 10.1021/acsami.5b02832

[50] Ho YJ, Chang YC, Yeh CK. Improving nanoparticle penetration in tumors by vascular disruption with acoustic droplet vaporization. Theranostics. 2016;**6**:392-403. DOI: 10.7150/thno.13727

[51] Bai WK, Zhang W, Hu B, Ying T. Liposome-mediated transfection of wild-type P53 DNA into human prostate cancer cells is improved by low-frequency ultrasound combined with microbubbles. Oncology Letters. 2016;**11**:3829-3834. DOI: 10.3892/ol.2016.4477

[52] Sun QF, Du F, Xu Y, Yang PS. Ultrasound-mediated microbubble destruction enhances bone morphogenetic protein-2 gene expression in mouse skeletal muscles. Hua Xi Kou Qiang Yi Xue Za Zhi. 2009;**27**:84-87

[53] Yoon YI, Kwon YS, Cho HS, Heo SH, Park KS, et al. Ultrasound-mediated gene and drug delivery using a microbubble-liposome particle system. Theranostics. 2014;**4**:1133-1144. DOI: 10.7150/thno.9945

[54] Chertok B, Langer R, Anderson DG. Spatial control of gene expression by nanocarriers using heparin masking and ultrasound-targeted microbubble destruction. ACS Nano. 2016;**10**:7267-7278. DOI: 10.1021/acsnano.6b01199

[55] Nel AE, Madler L, Velegol D, Xia T, Hoek EM, et al. Understanding biophysicochemical interactions at the nano-bio interface. Nature Materials. 2009;**8**:543-557. DOI: 10.1038/nmat2442

[56] Wang YJ, Zhou Q, Cao S, Hu B, Deng Q, et al. Efficient gene therapy with a combination of ultrasoundtargeted microbubble destruction and PEI/DNA/NLS complexes. Molecular Medicine Reports. 2017;**16**:7685-7691. DOI: 10.3892/mmr.2017.7510

[57] Arami S, Mahdavi M, Rashidi MR, Yekta R, Rahnamay M, et al. Apoptosis induction activity and molecular docking studies of survivin siRNA carried by Fe_3O_4-PEG-LAC-chitosan-PEI nanoparticles in MCF-7 human breast cancer cells. Journal of Pharmaceutical and Biomedical Analysis. 2017;**142**:145-154. DOI: 10.1016/j.jpba.2017.04.025

[58] Dang SP, Wang RX, Qin MD, Zhang Y, Gu YZ, et al. A novel transfection method for eukaryotic cells using polyethylenimine coated albumin microbubbles. Plasmid. 2011;**66**:19-25. DOI: 10.1016/j.plasmid.2011.03.003

[59] Deshpande MC, Prausnitz MR. Synergistic effect of ultrasound and PEI on DNA transfection in vitro. Journal of Controlled Release. 2007;**118**:126-135. DOI: 10.1016/j.jconrel.2006.12.010

[60] Park DH, Jung BK, Lee YS, Jang JY, Kim MK, et al. Evaluation of in vivo antitumor effects of ANT2 shRNA delivered using PEI and ultrasound with microbubbles. Gene Therapy. 2015;**22**:325-332. DOI: 10.1038/gt.2014.120

[61] Nande R, Greco A, Gossman MS, Lopez JP, Claudio L, et al. Microbubble-assisted p53, RB, and p130 gene transfer in combination with radiation therapy in prostate cancer. Current Gene Therapy. 2013;**13**:163-174

[62] Chang S, Guo J, Sun J, Zhu S, Yan Y, et al. Targeted microbubbles for ultrasound mediated gene transfection and apoptosis induction in ovarian cancer cells. Ultrasonics Sonochemistry. 2013;**20**:171-179. DOI: 10.1016/j.ultsonch.2012.06.015

[63] Mishel S, Shneyer B, Korsensky L, Goldshmidt-Tran O, Haber T, et al. Delivery of the gene encoding the tumor suppressor Sef into prostate tumors by therapeutic-ultrasound inhibits both tumor angiogenesis and growth. Scientific Reports. 2017;**7**:15060. DOI: 10.1038/s41598-017-12408-1

[64] Zhang J, Kale V, Chen M. Gene-directed enzyme prodrug therapy. The AAPS Journal. 2015;**17**:102-110. DOI: 10.1208/s12248-014-9675-7

[65] Devulapally R, Lee T, Barghava-Shah A, Sekar TV, Foygel K, et al. Ultrasound-guided delivery of thymidine kinase-nitroreductase dual therapeutic genes by PEGylated-PLGA/PIE nanoparticles for enhanced triple negative breast cancer therapy. Nanomedicine (London, England). 2018;**13**:1051-1066. DOI: 10.2217/nnm-2017-0328

[66] Bae YJ, Yoon YI, Yoon TJ, Lee HJ. Ultrasound-guided delivery of siRNA and a chemotherapeutic drug by using microbubble complexes: In vitro and in vivo evaluations in a prostate cancer model. Korean Journal of Radiology. 2016;**17**:497-508. DOI: 10.3348/kjr.2016.17.4.497

[67] Yin T, Wang P, Li J, Wang Y, Zheng B, et al. Tumor-penetrating codelivery of siRNA and paclitaxel with ultrasound-responsive nanobubbles hetero-assembled from polymeric micelles and liposomes. Biomaterials. 2014;**35**:5932-5943. DOI: 10.1016/j.biomaterials.2014.03.072

[68] Wang TY, Choe JW, Pu K, Devulapally R, Bachawal S, et al.

Ultrasound-guided delivery of microRNA loaded nanoparticles into cancer. Journal of Controlled Release. 2015;**203**:99-108. DOI: 10.1016/j.jconrel.2015.02.018

[69] Florinas S, Kim J, Nam K, Janat-Amsbury MM, Kim SW. Ultrasound-assisted siRNA delivery via arginine-grafted bioreducible polymer and microbubbles targeting VEGF for ovarian cancer treatment. Journal of Controlled Release. 2014;**183**:1-8. DOI: 10.1016/j.jconrel.2014.03.025

[70] Zhao R, Liang X, Zhao B, Chen M, Liu R, et al. Ultrasound assisted gene and photodynamic synergistic therapy with multifunctional FOXA1-siRNA loaded porphyrin microbubbles for enhancing therapeutic efficacy for breast cancer. Biomaterials. 2018;**173**:58-70. DOI: 10.1016/j.biomaterials.2018.04.054

[71] Liu YM, Li XF, Liu H, Wu XL. Ultrasound-targeted microbubble destruction-mediated downregulation of CD133 inhibits epithelial-mesenchymal transition, stemness and migratory ability of liver cancer stem cells. Oncology Reports. 2015;**34**: 2977-2986. DOI: 10.3892/or.2015.4270

[72] Zhang M, Yu WZ, Shen XT, Xiang Q, Xu J, et al. Advanced interfere treatment of diabetic cardiomyopathy rats by aFGF-loaded heparin-modified microbubbles and UTMD technique. Cardiovascular Drugs and Therapy. 2016;**30**:247-261. DOI: 10.1007/s10557-016-6639-4

[73] Zhao YZ, Tian XQ, Zhang M, Cai L, Ru A, et al. Functional and pathological improvements of the hearts in diabetes model by the combined therapy of bFGF-loaded nanoparticles with ultrasound-targeted microbubble destruction. Journal of Controlled Release. 2014;**186**:22-31. DOI: 10.1016/j.jconrel.2014.04.054

[74] Du GQ, Shao ZB, Wu J, Yin WJ, Li SH, et al. Targeted myocardial delivery of GDF11 gene rejuvenates the aged mouse heart and enhances myocardial regeneration after ischemia-reperfusion injury. Basic Research in Cardiology. 2017;**112**:7. DOI: 10.1007/s00395-016-0593-y

[75] Castle JW, Kent KP, Fan Y, Wallace KD, Davis CE, et al. Therapeutic ultrasound: Increased HDL-cholesterol following infusions of acoustic microspheres and apolipoprotein A-I plasmids. Atherosclerosis. 2015;**241**:92-99. DOI: 10.1016/j.atherosclerosis.2015.04.817

[76] Lee PJ, Rudenko D, Kuliszewski MA, Liao C, Kabir MG, et al. Survivin gene therapy attenuates left ventricular systolic dysfunction in doxorubicin cardiomyopathy by reducing apoptosis and fibrosis. Cardiovascular Research. 2014;**101**:423-433. DOI: 10.1093/cvr/cvu001

[77] Patel MM, Patel BM. Crossing the blood-brain barrier: Recent advances in drug delivery to the brain. CNS Drugs. 2017;**31**:109-133. DOI: 10.1007/s40263-016-0405-9

[78] Timbie KF, Mead BP, Price RJ. Drug and gene delivery across the blood-brain barrier with focused ultrasound. Journal of Controlled Release. 2015;**219**:61-75. DOI: 10.1016/j.jconrel.2015.08.059

[79] Ikeda M, Bhattacharjee AK, Kondoh T, Nagashima T, Tamaki N. Synergistic effect of cold mannitol and Na(+)/Ca(2+) exchange blocker on blood-brain barrier opening. Biochemical and Biophysical Research Communications. 2002;**291**:669-674. DOI: 10.1006/bbrc.2002.6495

[80] Lee HJ, Zhang Y, Pardridge WM. Blood-brain barrier disruption following the internal carotid arterial perfusion of alkyl glycerols. Journal of

Drug Targeting. 2002;**10**:463-467. DOI: 10.1080/1061186021000038337

[81] Hynynen K, McDannold N, Vykhodtseva N, Jolesz FA. Noninvasive MR imaging-guided focal opening of the blood-brain barrier in rabbits. Radiology. 2001;**220**:640-646. DOI: 10.1148/radiol.2202001804

[82] Huang Q, Deng J, Wang F, Chen S, Liu Y, et al. Targeted gene delivery to the mouse brain by MRI-guided focused ultrasound-induced blood-brain barrier disruption. Experimental Neurology. 2012;**233**:350-356. DOI: 10.1016/j.expneurol.2011.10.027

[83] Negishi Y, Yamane M, Kurihara N, Endo-Takahashi Y, Sashida S, et al. Enhancement of blood-brain barrier permeability and delivery of antisense oligonucleotides or plasmid DNA to the brain by the combination of bubble liposomes and high-intensity focused ultrasound. Pharmaceutics. 2015;**7**:344-362. DOI: 10.3390/pharmaceutics7030344

[84] Carpentier A, Canney M, Vignot A, Reina V, Beccaria K, et al. Clinical trial of blood-brain barrier disruption by pulsed ultrasound. Science Translational Medicine. 2016;**8**:342r-343r. DOI: 10.1126/scitranslmed.aaf6086

[85] Fan CH, Chang EL, Ting CY, Lin YC, Liao EC, et al. Folate-conjugated gene-carrying microbubbles with focused ultrasound for concurrent blood-brain barrier opening and local gene delivery. Biomaterials. 2016;**106**:46-57. DOI: 10.1016/j.biomaterials.2016.08.017

[86] Chang EL, Ting CY, Hsu PH, Lin YC, Liao EC, et al. Angiogenesis-targeting microbubbles combined with ultrasound-mediated gene therapy in brain tumors. Journal of Controlled Release. 2017;**255**:164-175. DOI: 10.1016/j.jconrel.2017.04.010

[87] Zhao G, Huang Q, Wang F, Zhang X, Hu J, et al. Targeted shRNA-loaded liposome complex combined with focused ultrasound for blood brain barrier disruption and suppressing glioma growth. Cancer Letters. 2018;**418**:147-158. DOI: 10.1016/j.canlet.2018.01.035

[88] Fan CH, Lin CY, Liu HL, Yeh CK. Ultrasound targeted CNS gene delivery for Parkinson's disease treatment. Journal of Controlled Release. 2017;**261**:246-262. DOI: 10.1016/j.jconrel.2017.07.004

[89] Fan CH, Ting CY, Lin CY, Chan HL, Chang YC, et al. Noninvasive, targeted, and non-viral ultrasound-mediated GDNF-plasmid delivery for treatment of Parkinson's disease. Scientific Reports. 2016;**6**:19579. DOI: 10.1038/srep19579

[90] Lin CY, Hsieh HY, Chen CM, Wu SR, Tsai CH, et al. Non-invasive, neuron-specific gene therapy by focused ultrasound-induced blood-brain barrier opening in Parkinson's disease mouse model. Journal of Controlled Release. 2016;**235**:72-81. DOI: 10.1016/j.jconrel.2016.05.052

[91] Lin CY, Hsieh HY, Pitt WG, Huang CY, Tseng IC, et al. Focused ultrasound-induced blood-brain barrier opening for non-viral, non-invasive, and targeted gene delivery. Journal of Controlled Release. 2015;**212**:1-9. DOI: 10.1016/j.jconrel.2015.06.010

[92] Song Z, Wang Z, Shen J, Xu S, Hu Z. Nerve growth factor delivery by ultrasound-mediated nanobubble destruction as a treatment for acute spinal cord injury in rats. International Journal of Nanomedicine. 2017;**12**:1717-1729. DOI: 10.2147/IJN.S128848

[93] Wang HB, Yang L, Wu J, Sun L, Wu J, et al. Reduced ischemic injury

after stroke in mice by angiogenic gene delivery via ultrasound-targeted microbubble destruction. Journal of Neuropathology and Experimental Neurology. 2014;**73**:548-558. DOI: 10.1097/NEN.0000000000000077

[94] Xiang X, Tang Y, Leng Q, Zhang L, Qiu L. Targeted gene delivery to the synovial pannus in antigen-induced arthritis by ultrasound-targeted microbubble destruction in vivo. Ultrasonics. 2016;**65**:304-314. DOI: 10.1016/j.ultras.2015.09.011

[95] Inoue H, Arai Y, Kishida T, Shin-Ya M, Terauchi R, et al. Sonoporation-mediated transduction of siRNA ameliorated experimental arthritis using 3 MHz pulsed ultrasound. Ultrasonics. 2014;**54**:874-881. DOI: 10.1016/j.ultras.2013.10.021

[96] King GN, Hughes FJ. Bone morphogenetic protein-2 stimulates cell recruitment and cementogenesis during early wound healing. Journal of Clinical Periodontology. 2001;**28**:465-475

[97] Sun QF, Liu Y, Yang PS, Du F. Ultrasound-mediated microbubble destruction enhances exogenous gene expression in NIH3T3 cells in vitro. Hua Xi Kou Qiang Yi Xue Za Zhi. 2008;**26**:198-200, 205

[98] Osawa K, Okubo Y, Nakao K, Koyama N, Bessho K. Osteoinduction by microbubble-enhanced transcutaneous sonoporation of human bone morphogenetic protein-2. The Journal of Gene Medicine. 2009;**11**:633-641. DOI: 10.1002/jgm.1331

[99] Qiu L, Zhang L, Wang L, Jiang Y, Luo Y, et al. Ultrasound-targeted microbubble destruction enhances naked plasmid DNA transfection in rabbit Achilles tendons in vivo. Gene Therapy. 2012;**19**:703-710. DOI: 10.1038/gt.2011.165

[100] Delalande A, Bureau MF, Midoux P, Bouakaz A, Pichon C. Ultrasound-assisted microbubbles gene transfer in tendons for gene therapy. Ultrasonics. 2010;**50**:269-272. DOI: 10.1016/j.ultras.2009.09.035

[101] Delalande A, Bouakaz A, Renault G, Tabareau F, Kotopoulis S, et al. Ultrasound and microbubble-assisted gene delivery in Achilles tendons: Long lasting gene expression and restoration of fibromodulin KO phenotype. Journal of Controlled Release. 2011;**156**:223-230. DOI: 10.1016/j.jconrel.2011.08.020

[102] Tang Y, Leng Q, Xiang X, Zhang L, Yang Y, Qiu L. Use of ultrasound-targeted microbubble destruction to transfect IGF-1 cDNA to enhance the regeneration of rat wounded Achilles tendon in vivo. Gene Therapy. 2015;**22**:610-618. DOI: 10.1038/gt.2015.32

[103] Bez M, Kremen TJ, Tawackoli W, Avalos P, Sheyn D, et al. Ultrasound-mediated gene delivery enhances tendon allograft integration in mini-pig ligament reconstruction. Molecular Therapy. 2018;**26**:1746-1755. DOI: 10.1016/j.ymthe.2018.04.020

[104] Sonoda S, Tachibana K, Uchino E, Okubo A, Yamamoto M, et al. Gene transfer to corneal epithelium and keratocytes mediated by ultrasound with microbubbles. Investigative Ophthalmology & Visual Science. 2006;**47**:558-564. DOI: 10.1167/iovs.05-0889

[105] Yamashita T, Sonoda S, Suzuki R, Arimura N, Tachibana K, et al. A novel bubble liposome and ultrasound-mediated gene transfer to ocular surface: RC-1 cells in vitro and conjunctiva in vivo. Experimental Eye Research. 2007;**85**:741-748. DOI: 10.1016/j.exer.2007.08.006

[106] Park J, Zhang Y, Vykhodtseva N, Akula JD, McDannold NJ. Targeted and reversible blood-retinal barrier disruption via focused ultrasound and microbubbles. PLoS One. 2012;**7**: e42754. DOI: 10.1371/journal.pone.0042754

[107] Li HL, Zheng XZ, Wang HP, Li F, Wu Y, Du LF. Ultrasound-targeted microbubble destruction enhances AAV-mediated gene transfection in human RPE cells in vitro and rat retina in vivo. Gene Therapy. 2009;**16**:1146-1153. DOI: 10.1038/gt.2009.84

[108] Wan C, Qian J, Li F, Li H. Ultrasound-targeted microbubble destruction enhances polyethylenimine-mediated gene transfection in vitro in human retinal pigment epithelial cells and in vivo in rat retina. Molecular Medicine Reports. 2015;**12**:2835-2841. DOI: 10.3892/mmr.2015.3703

[109] Li H, Qian J, Yao C, Wan C, Li F. Combined ultrasound-targeted microbubble destruction and polyethylenimine-mediated plasmid DNA delivery to the rat retina: Enhanced efficiency and accelerated expression. The Journal of Gene Medicine. 2016;**18**:47-56. DOI: 10.1002/jgm.2875

[110] Xu Y, Xie Z, Zhou Y, Zhou X, Li P, et al. Experimental endostatin-GFP gene transfection into human retinal vascular endothelial cells using ultrasound-targeted cationic microbubble destruction. Molecular Vision. 2015;**21**:930-938

[111] Luo J, Zhou X, Diao L, Wang Z. Experimental research on wild-type p53 plasmid transfected into retinoblastoma cells and tissues using an ultrasound microbubble intensifier. The Journal of International Medical Research. 2010;**38**:1005-1015. DOI: 10.1177/147323001003800327

[112] Gao R, Zhou X, Yang Y, Wang Z. Transfection of wtp53 and Rb94 genes into retinoblastomas of nude mice by ultrasound-targeted microbubble destruction. Ultrasound in Medicine & Biology. 2014;**40**:2662-2670. DOI: 10.1016/j.ultrasmedbio.2014.05.012

[113] Bai YJ, Huang LZ, Zhou AY, Zhao M, Yu WZ, Li XX. Antiangiogenesis effects of endostatin in retinal neovascularization. Journal of Ocular Pharmacology and Therapeutics. 2013;**29**:619-626. DOI: 10.1089/jop.2012.0225

[114] Zhang Y, Ye C, Xu Y, Dong X, Li J, et al. Ultrasound-mediated microbubble destruction increases renal interstitial capillary permeability in early diabetic nephropathy rats. Ultrasound in Medicine & Biology. 2014;**40**:1273-1281. DOI: 10.1016/j.ultrasmedbio.2013.12.006

[115] Lan HY, Mu W, Tomita N, Huang XR, Li JH, et al. Inhibition of renal fibrosis by gene transfer of inducible Smad7 using ultrasound-microbubble system in rat UUO model. Journal of the American Society of Nephrology. 2003;**14**:1535-1548

[116] Chen HY, Huang XR, Wang W, Li JH, Heuchel RL, et al. The protective role of Smad7 in diabetic kidney disease: Mechanism and therapeutic potential. Diabetes. 2011;**60**:590-601. DOI: 10.2337/db10-0403

[117] Ka SM, Yeh YC, Huang XR, Chao TK, Hung YJ, et al. Kidney-targeting Smad7 gene transfer inhibits renal TGF-beta/MAD homologue (SMAD) and nuclear factor kappaB (NF-kappaB) signalling pathways, and improves diabetic nephropathy in mice. Diabetologia. 2012;**55**:509-519. DOI: 10.1007/s00125-011-2364-5

[118] Liu GX, Li YQ, Huang XR, Wei LH, Zhang Y, et al. Smad7 inhibits

AngII-mediated hypertensive nephropathy in a mouse model of hypertension. Clinical Science (London, England). 2014;**127**:195-208. DOI: 10.1042/CS20130706

[119] Dube N, Kooistra MR, Pannekoek WJ, Vliem MJ, Oorschot V, et al. The RapGEF PDZ-GEF2 is required for maturation of cell-cell junctions. Cellular Signalling. 2008;**20**:1608-1615. DOI: 10.1016/j.cellsig.2008.05.006

[120] Xiao L, Zhu X, Yang S, Liu F, Zhou Z, et al. Rap1 ameliorates renal tubular injury in diabetic nephropathy. Diabetes. 2014;**63**:1366-1380. DOI: 10.2337/db13-1412

[121] Yiu WH, Wong DW, Wu HJ, Li RX, Yam I, Chan LY, Leung JC, Lan HY, Lai KN, Tang SC. Kallistatin protects against diabetic nephropathy in db/db mice by suppressing AGE-RAGE-induced oxidative stress. Kidney International. 2016;**89**:386-398. DOI: 10.1038/ki.2015.331

[122] Du B, Ma LM, Huang MB, Zhou H, Huang HL, et al. High glucose down-regulates miR-29a to increase collagen IV production in HK-2 cells. FEBS Letters. 2010;**584**:811-816. DOI: 10.1016/j.febslet.2009.12.053

[123] Roderburg C, Urban GW, Bettermann K, Vucur M, Zimmermann H, et al. Micro-RNA profiling reveals a role for miR-29 in human and murine liver fibrosis. Hepatology. 2011;**53**:209-218. DOI: 10.1002/hep.23922

[124] Miller DL. Overview of experimental studies of biological effects of medical ultrasound caused by gas body activation and inertial cavitation. Progress in Biophysics and Molecular Biology. 2007;**93**:314-330. DOI: 10.1016/j.pbiomolbio.2006.07.027

Chapter 7

Gene Regulation in Ruminants: A Nutritional Perspective

Johan S. Osorio and Sonia J. Moisa

Abstract

This chapter will focus on cellular regulatory programs implemented by the ruminant physiology in order to respond to external stimuli such as nutrition as well as important physiological events such as parturition. The increasing adoption of "omics" technologies and bioinformatics in nutrition and physiology in ruminant research have allowed us to delineate a clearer picture on what regulates major biological process at a molecular level such as milk synthesis and meat quality and fatty acid composition as well as pathological conditions such as ketosis, mastitis, and heat stress. The assembly of such plethora of information in a blend among nutritional research, molecular biology, and novel tools to study the response of the genome to nutrition has led to emerging disciplines such as nutritional genomics or "nutrigenomics."

Keywords: nutrigenomics, transcription factors, ruminants

1. Introduction

The increasing adoption of molecular biology techniques and bioinformatics in nutrition and physiology in ruminant research has provided a wealth of knowledge on regulatory mechanisms of major biological processes related to milk synthesis and meat quality and marbling at a cellular level. This body of knowledge has prompted a compelling case for a change in the paradigm in ruminant nutrition, where nutrients in ruminant diets can act as bioactive molecules and exert alterations in molecular mechanisms depending on the animal physiological state. Such alterations can be carried out through gene regulation mechanisms, also known as nutrigenomics. The continuous accumulation of nutrient-gene interactions in ruminant research will eventually lead to practical applications where nutritional interventions may be made in order to improve performance and efficiency in milk yield or skeletal muscle.

1.1 Ruminant model for nutrient-gene interactions

The adoption of advanced molecular technologies in basic nutritional research in ruminants has led to a more robust notion of how nutrients can affect the animal at the cellular level. Then, this body of knowledge has led toward a general notion in ruminant nutrition, where nutrients in the diet can no longer be considered only as: (1) the building blocks for cells, tissue, and organs or (2) energy source for cell metabolism and basic cell function, but rather a new alternative concept for nutrient is that they (3) act as bioactive molecules that can regulate fundamental molecular mechanisms depending on the animal physiological stage.

Because of the inherited gastrointestinal differences between ruminants and monogastrics, the final effect of a nutrient at the molecular level will differ primarily based on how susceptible such a nutrient is to rumen fermentation. Therefore, rumen fermentation and kinetics play an important role in the context of nutrient-gene interaction in ruminants (**Figure 1**). Then, from a nutrigenomic standpoint, a given nutrient in a ruminant diet will likely be fermented or bypass the rumen. If fermented in the rumen, this nutrient will become either part of the microbial biomass or an intermediate metabolite such as volatile fatty acids (VFA) which can be absorbed through the rumen wall and enter the metabolism of ruminants. In the case of nutrients bypassing the rumen, these can be converted to intermediate metabolites [1], produce a signal transduction cascade [2, 3], or directly bind and activate specialized cellular proteins called transcription factors (or nuclear receptors) [4, 5] which are responsible for carrying out the final change in gene expression by binding to specific sections in the DNA upstream of the target gene. Some transcription factors can create a secondary wave of change in gene expression by upregulating the transcription of subsequent transcription factor [6], and previously, it has been proposed that transcription factors may work in an orchestrated fashion creating a network of transcription factors that respond to dietary effects [7]. An alternative effect from intermediate metabolites is the production of DNA or histone modifications by changing the available information in the DNA [8], also known as epigenetic effects. A potential epigenetic mechanism mediated by transcription factors is the increased transcription of noncoding RNAs such as microRNAs [9], which upon transcription these small RNAs will target coding RNA prior to their translation into proteins.

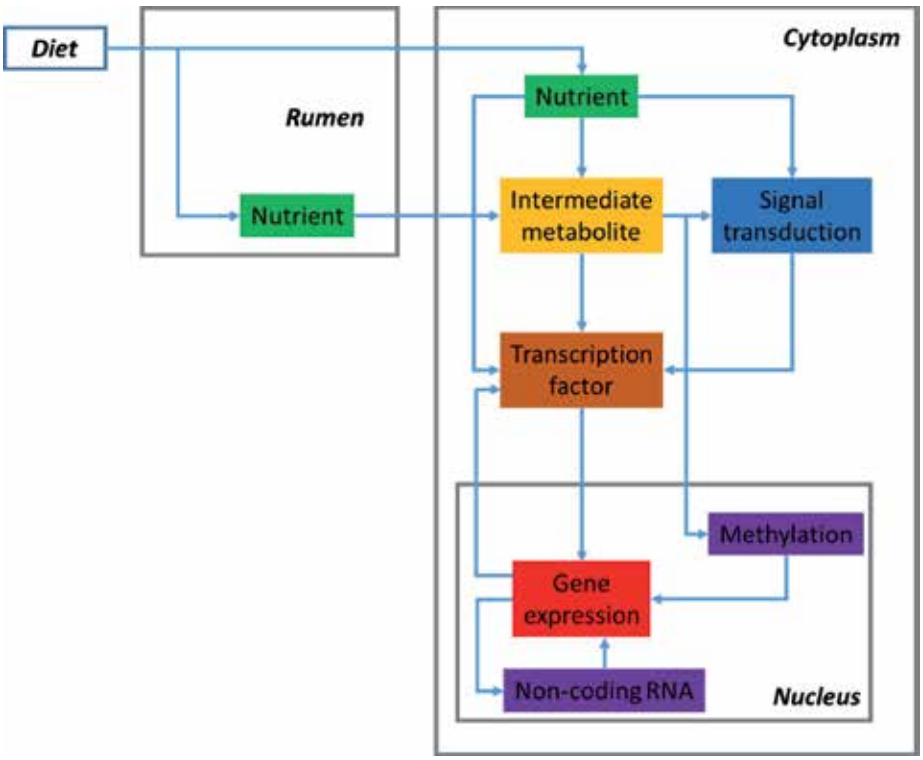

Figure 1.
Proposed ruminant model for gene expression regulation of dietary nutrients through transcription factor activation and epigenetic mechanism (i.e., DNA or histone methylation and noncoding RNA).

1.2 Mediators of nutrient-gene interactions in ruminants

Nutrients and bioactive compounds in regular ruminant diets will mainly interact indirectly with the genome through mediators in the form of specialized molecular proteins such as transcription factors, DNA methyltransferases, histone methyltransferases, among others. Here, we provide a brief overview of the major mediators of nutrient-gene interactions in ruminants known to date. Based on the ruminant model for nutrient-gene interactions (**Figure 1**), the specific transcription factors, enzymes, or cellular mechanisms with *in vivo* or *in vitro* data in ruminants are presented in **Table 1**. Among the known transcription factors with nutrigenomic potential, the peroxisome proliferator-activated receptors (PPARs) have been well studied in ruminants [10]. These transcription factors belong to the ligand-dependent nuclear receptors (LdNR) family [11], and their importance for nutrigenomic interventions in ruminants relies on their ability to bind and be activated by long-chained fatty acids (LCFA) commonly present in ruminant diets. The PPAR isotypes (e.g., α, γ, and δ) play multiple roles across several tissues in mammals, for instance, PPARγ has been observed to regulate adipogenesis and insulin sensitivity [12, 13], while PPARα has a crucial role in hepatic fatty acid catabolism [14]. In contrast to PPARγ and PPARα, PPARδ has been studied to a lesser extent; however, it is known for its role in fatty acid catabolism in skeletal muscle [15] and regulation of glucose uptake [7]. Additionally, Bionaz and collaborators [10] proposed a model for the concomitant and orchestrated regulation of major physiological adaptations by the three isotypes of PPARs in dairy cows going from late pregnancy into lactation. These effects are exerted across several tissues (e.g., liver, skeletal muscle, mammary gland, adipose, immune cells, etc.) where PPARs have a strong effect, and their ability to be activated by dietary fatty acids makes them a strong candidate for nutrigenomic effects in ruminants. The PPARs exert similar effects as observed in dairy cows in other ruminants, for instance, PPARγ has been associated with adipogenic effects in beef, cows [16, 17], and goats [18] as well as fatty acid oxidation by PPARα in transition dairy goats [19].

Similar to the PPARs, the liver X receptor (LXR) belongs to the LdNR family and has a prominent role in controlling cholesterol synthesis [15]. The LXR is known to be activated by oxysterols and derivatives from cholesterol metabolism, and fatty acids [15]. From the two known isoforms of LXR (e.g., α and β), the LXRα presents interesting characteristics including the potential control of sterol regulatory element-binding transcription factor 1 (*SREBF1*) gene expression [20], which is a major transcription factor associated with the regulation of milk fat synthesis [21]. The ability of LXRα to regulate *SRBEF1* expression confers this TF a strong potential to enhance milk fat synthesis in ruminants; however, most of the current data on LXRα activity have been conducted with synthetic agonist [20, 22]. Therefore, a stronger case for the nutrigenomic potential of this TF could be made by future research including its activation by common fatty acids present in ruminant diets.

Retinoids are metabolites derived from vitamin A, and they can regulate gene expression through two classes of receptors: retinoic acid receptors and retinoid X receptors (RXR). The latter can form homodimers and be activated in the presence of the retinoid 9-cis-retinoic acid and consequently activating specific target genes [23]. From the two isoforms (i.e., α and β) of RXR, the RXRα has been the most evaluated in ruminants, primarily because it can form heterodimers with most LdNR including PPAR, LXR, and VDR [24]. Although the latter confers RXRα a tremendous biological significance, there are limited data in ruminants on the potential nutrigenomic effects of vitamin A and derivative retinoids such as 9-cis-retinoic acid through RXRα.

Common name	Protein symbol	Gene symbol	Agonist	Main function	Ruminant[1]	Reference
Transcription factors						
Peroxisome proliferator-activated receptor α	PPARα	*PPARA*	Fatty acids	Fatty acid metabolism, inflammation, and tissue regeneration	B, D, G, and S	[10, 24], [126, 127]
Peroxisome proliferator-activated receptor γ	PPARγ	*PPARG*	Fatty acids	Adipogenesis, insulin sensitivity, and lipogenesis	B, D, G, and S	[10, 16–18, 24], [126]
Peroxisome proliferator-activated receptor β	PPARβ	*PPARD*	Fatty acids	Fatty acid metabolism, tissue regeneration, and glucose uptake in mammary tissue	B, D, G, and S	[10], [127–130]
Liver X receptor α	LXRα	*NR1H3*	Oxysterols/fatty acids	Cholesterol homeostasis, macrophage functions, and inflammation	B, D, and G	[22, 62], [131]
Retinoic X receptor α	RXRα	*RXRA*	9-cis-retinoic acid	Forming heterodimers with other LdNR and neutrophil differentiation	B and D	[132–134]
Sterol regulatory element-binding protein 1	SREBP1	*SREBF1*	N/A	Cholesterol and fatty acid synthesis	B, D, G, and S	[25, 26], [135]
DNA methyltransferases						
DNA methyltransferase 1	DNMT1	*DNMT1*	N/A	Maintenance of methylation patterns	B and D	[72, 105]
DNA methyltransferase 3 α	DNMT3a	*DNMT3A*	N/A	Creates de novo methylation patterns. Present in cytoplasm and nucleus	B and D	[72], [136]
DNA methyltransferase 3 β	DNMT3b	*DNMT3B*	N/A	Creates de novo methylation patterns restricted to nucleus	B and D	[137, 138]
Noncoding RNA						
MicroRNA 33		miR33b		Regulates lipogenesis	D	[30]
MicroRNA 192		miR192		Regulates myogenesis	S	[139]

[1]*Ruminant as B = beef cows, D = dairy cows, G = goats, and S = sheep.*

Table 1.
Important mediators associated with nutrient-gene interactions in ruminants via transcriptional regulatory factors (transcription factors) and epigenetic factors (DNA methyltransferases and noncoding RNAs).

The study of milk fat depression revealed the importance of SREBP1 in milk fat synthesis [25], which resulted in a deep understanding how *t*10,*c*12 CLA, and a milk fat-depressing diet consistently downregulate *SREBF1* in bovine mammary tissue [21]. Consequently, this effect downregulates the expression of genes associated with milk fat synthesis such as fatty acid synthase (*FASN*), lipoprotein lipase (*LPL*), and insulin-induced gene 1 (*INSIG1*) [21]. The importance of this TF for nutrigenomics in ruminants cannot be overstated since this was the first nutrigenomic effect documented. And the importance of this TF in the regulation of fat synthesis has also been observed in beef cattle and translated in marbling and meat quality [26]. A section on this commonality between dairy and beef cattle is dedicated at the end of this chapter.

Epigenetic mechanisms play a significant role as mediators of nutrient-gene interactions in ruminants and the ramifications of these effects in ruminant nutrition and physiology are only beginning to be uncovered, and they add another layer of complexity to our model (**Figure 1**). From a nutrigenomic standpoint, methyl donors present in common diets fed to ruminants such as folate, vitamin B (e.g., 2, 6, 12), choline, and methionine can regulate epigenetic modifications through the one-carbon metabolism where the intermediate s-adenosylmethionine (SAM) is produced and subsequently used as the universal methyl donor for DNA and histone methylations [27]. These effects are carried out through specialized enzymes such as DNA and histone methyltransferases. While the effect of dietary methyl donors on DNA methyltransferases has been evaluated in ruminants, the effect of methyl donors on histone methyltransferases in ruminants remains unknown. Similarly, other histone modifications such as acetylation and phosphorylation have not been investigated within the context of nutrigenomics in ruminants.

Among the various epigenetic mechanisms, noncoding RNA and specifically microRNAs have received a lot of notoriety in recent years [28], and, in contrast, long noncoding RNAs (LncRNAs) are only beginning to be evaluated in ruminants [29]. Examples of microRNAs with a potential application to improving milk and meat quality are the miR33b and miR192 (**Table 1**). The former has been previously associated with lipogenesis in the mammary gland of dairy cows as well as having the greatest upregulation from pregnancy into lactation [30]. In the case of miR192, it has been observed to influence muscle development through myogenesis in sheep. Since the interactions between microRNAs and coding mRNAs are one-to-many, meaning that a single microRNA can regulate the translation of several coding mRNAs, special caution should be applied when interpreting this type of data. The LncRNAs are relatively new in the context of ruminant nutrition and physiology and are commonly characterized by containing >200 nt that are not translated into proteins [31]. This work provided nuances on the role of LncRNAs in the mammary gland in terms of mastitis and milk quality and production.

The importance of understanding these multiple mediators of nutrient-gene interactions cannot be overstated. The authors envision that the continuous accumulation of this wealth of knowledge will lead to accurate and consistent manipulation of the ruminant genome to access or unlock the full genetic potential with the aim to produce ruminant products more efficiently, with a targeted effect on human health, and with a lesser cost for the environment.

2. Gene regulation in dairy cattle

Milk is one of the most nutritious foods known to man, and milk from dairy cattle has been part of the human diet from approximately 9000 years ago [32]. And, currently, the consumption of milk and milk-derived products around the world is expected to

increase, primarily due to an increase in world population and increased consumption in countries where milk has not traditionally been considered popular [33]. Until now, the demand of milk worldwide has been supplied by a large increase in milk yield per cow, which has been a product primarily from the selection and enhancement of management practices, including improved nutrition. However, because of the ever-increasing demand for milk and milk-derived products as stated above, there is a need to continue increasing milk production efficiency.

Milk and milk products are an excellent source of macronutrients such as fat, protein, and carbohydrates, and contain a variety of bioactive molecules associated with health benefits, for instance, conjugated linolenic acid (CLA). The CLA has been associated with reductions in cancer development [25]. Because of its ability to contain bioactive molecules, milk has been considered a functional food. However, our ability to understand and yet manipulate the cow genome through nutrigenomic approaches to enrich specific bioactive molecules in milk is in its infancy. This calls for a continuous development of a wealth of knowledge around the various complex nutrient-gene interactions in dairy cows as well as development of nutritional models that can account for both traditional aspects of ruminant nutrition and more novel molecular regulation of nutrient metabolism.

2.1 Gene regulation and milk biosynthesis

During the lactation, the mammary gland is in charge of the final biosynthesis of milk using preformed elements from other organs (e.g., glucose synthesized in the liver) or compounds synthesized within the mammary epithelial cells (e.g., *de novo* fatty acids). The biosynthesis of milk in the mammary gland is highly regulated for several factors including nutrient supply (e.g., glucose, AA, and fatty acids) and hormones primarily related to hormonal changes during the onset and at the decline of the lactation.

2.1.1 Lactogenesis and the mammary gland transcriptome

Lactogenesis is the hallmark of mammalians, and as such, this biological process conveys a tremendous impact in gene regulation with the objective to induce the mammary gland to lactate and coincides with the formation of colostrum and occurs in coordination with parturition. The strong effects of lactogenesis on the mammary gland at the cellular level have been consistently recorded through transcriptomic analysis (i.e., microarray or RNA sequencing) across several mammals including mouse [34–36], rats [37, 38], bovine [39–41], sheep [42], goat [43], human [44–46], pig [47], kangaroo [48], and seal [49]. The extreme changes in the transcriptome from pregnancy to lactation underscore the importance of such change in the transcriptome in the mammary gland to initiate and maintain milk synthesis. In the case of bovine, a closer look at such transcriptional changes due to the onset and throughout the lactation was reported previously in terms of fat [50] and protein [51] synthesis. During the lactation, the milk fat biosynthesis in the bovine mammary gland had a marked upregulation in the expression of genes associated with FA uptake from blood, intracellular transport/channeling, and key transcription factors associated with lipogenesis (i.e., PPARG and SREBF2) [50]. Then, in terms of milk protein synthesis, it was observed that cell membrane transporters, especially for AA and glucose, played an essential role along with insulin signaling through mTOR for the regulation of protein synthesis in the bovine mammary gland [51]. At a greater scale, the impact of lactogenesis in the mammary gland transcriptome has been associated with epigenetic changes that result in alterations in the DNA structure and consequently the available genetic information

for transcription [39]. Bionaz [39] observed chromatin changes (i.e., euchromatin or active transcribed chromatin and heterochromatin or tightly packed and transcriptionally unavailable chromatin) associated with lactogenesis, where a decrease of euchromatin status was observed as lactation begins and followed by an increase in euchromatin status as milk yield decreases during late lactation. The latter could be associated with natural response to lactation in the mammary gland to inhibit further epigenetic effects during the onset of lactation and to maintain a consistent transcriptome until milk synthesis declines in late lactation.

2.1.2 Milk biosynthesis in the dairy cattle from a nutrigenomic approach

Nutrigenomics is a coined term to refer to the interactions between nutrients and the genome. However, the term nutrigenomics does not refer to the effect of nutrients on the sequence of DNA, but rather the nutrient-gene interactions through the intermediate action of transcription factors (TFs) in the short to medium term and epigenetic factors in the medium to long term. Bioactive compounds with potential nutrigenomic effects can be found in regular diets of dairy cattle, and such compounds can directly or indirectly activate or repress the activity of TF.

Nutrigenomics is a product of the postgenome era and its impact on human, companion animals, and livestock species has gained more attention in recent years [52–55]. In dairy cattle, and overall in ruminants, the field can be considered in its infancy but, as argued by Coffey [56], holds great potential to improve health and productivity.

2.1.3 Transcription factors with nutrigenomic potential in dairy cattle

The TFs are fundamental to the study of nutrigenomics; they can act as intermediaries between dietary nutrients and the final alteration in gene expression. TFs that respond to nutrient effects can be activated directly or indirectly by nutrients, and upon activation, they translocate from the cytoplasm to the nucleus where they alter the transcription of specific target genes. Transcription factors can bind specific short DNA sequences (i.e., 6–12 nucleotides) called response elements located in the enhancer regions upstream of the actual gene sequence [57]. The ability of TF to modulate gene expression upon activation by nutrients confers these proteins a central stage in the field of nutrigenomics. Therefore, accurate identification and characterization of TF responding to specific nutrients and to what extent these TF can be manipulated through dietary effects should be the focus of future research in nutrigenomics in dairy cattle.

Although between 2000 and 3000 TFs with sequence-specific DNA-binding domains in the human genome were estimated, only ~100 have been experimentally verified for their DNA-binding and regulatory functions [58]. The AnimalTFDB website had collected information for almost 1300 TF and almost 400 transcription cofactors for *Bos taurus* [59]. Rather than TF working independently, a regulatory network of TF has been suggested, which is essential to coordinate the response to external stimuli and translate this into changes in gene expression [60].

Among the known TF, the nuclear receptor superfamily of TF, with 48 members in the human genome, is the most important group of nutrient sensors [61]. From this superfamily, a short list of TFs has been identified as ligand-dependent nuclear receptors (LdNR), which can bind and be activated by macro- and micronutrients [55]. Recently, the main LdNR with a potential role in nutrigenomics with an emphasis in large [62] and small [24] ruminants has been reviewed. Among the LdNR associated with macronutrients such as fatty acids are the peroxisome proliferator-activated receptors (PPAR), liver X receptors (LXR), and hepatic

nuclear factor 4 (HNF4) [63]. The LdNR associated with micronutrients are primarily vitamin-specific and include retinoid X receptors (RXR) and retinoic acid receptors (RAR) activated by retinoic acids or metabolites of vitamin A [64], vitamin D receptor (VDR), and pregnane X receptor activated by vitamin E [65].

The PPAR belongs to the LdNR group of TF and requires to form heterodimers with RXR in order to be functional. The main characteristic of PPAR is to have a prominent role in controlling expression of genes involved in lipid metabolism and inflammation. The potential nutrigenomic role of PPAR in ruminants has been reviewed at length previously [10]. This review discussed the role of PPARα in controlling lipid metabolism and inflammation in liver, the potential role of PPARβ/δ in controlling glucose uptake in mammary tissue, and the potential role of PPARγ in controlling milk fat synthesis and mastitis [10]. An interesting feature of PPAR is their capacity for binding and be activated by LCFA in both monogastrics and ruminants [10, 66]. However, in the case of ruminants, data indicate that activation of the PPAR isotypes PPARα and PPARγ is more consistent with saturated LCFA, primarily palmitate and stearate, than unsaturated LCFA [10].

2.1.4 Nutriepigenomics in dairy cattle

Environment factors such as diet and ambient conditions can not only affect the short- and medium-term gene expression, but there is also a medium- to long-term regulation of genes. The latter is primarily carried out through changes in the availability of gene sequences to be transcribed into mRNA. This concept is referred to epigenetics, where "epi" is a Greek-derived term meaning "over" then, epigenetics is commonly referred as "on-top-of genetics." The implications of epigenetics indicate that there could be a set of inherited characteristics, phenotypes, and chemical entities that are superimposed on the DNA and do not follow basic Mendelian laws. Every individual will have a set of epigenetic marks throughout the genome, which is also known as the epigenome.

Epigenetic modifications are carried out through several biological processes including DNA methylation, histone modifications (e.g., methylation and acetylation), noncoding RNA (e.g., micro- and long-RNA). And, when these biological processes respond to nutrients and compounds in the diet, it is associated with nutriepigenomic effects. In ruminants, such effects could serve important physiological adaptations during the onset of lactation including increasing the availability of gene sequences through alterations in DNA methylation for the transcription of essential proteins such as caseins in the mammary gland of dairy cows [67]. This new spinoff of nutrigenomics (i.e., nutriepigenomics) will provide essential information to our understanding of how nutrients can affect the biology of ruminants at a molecular level. However, at the same time, nutriepigenomics will add another layer of complexity to our field, where such interactions have to be fully understood, and in time, manipulated through dietary interventions.

Methylation is a major route for epigenetic modifications, through DNA and histone methylation. Therefore, methyl donors (e.g., choline, methionine, folic acid, etc.) found in the diets of dairy cattle can have a nutriepigenomic effect. These dietary methyl donors will likely increase the synthesis of SAM, which is the major biological methyl donor in the body [68]. The essential role of SAM within the context of the transition cow relies on the multiple biological processes that require this methyl donor, including transsulfuration, polyamine biosynthesis, DNA methylation [69], and histone methylation [70]. Among these, the epigenetic modifications caused by DNA and histone methylation are particularly important in order to understand the potential nutriepigenomic alterations caused by dietary methyl donor (e.g., methionine) supplementation.

DNA methylation occurs through specialized enzymes called DNA methyltransferases, which utilized the methyl group provided by SAM to methylate cytosines within a Cyt-phosphate-Gua (CpG) region ("island") in the DNA and eventually creating methylated CpG patterns in the mammalian genome [71]. In ruminants, supplementation of rumen-protected to dairy cows resulted in a prepartal upregulation of *DNMT3A*, a gene that encodes for a DNA methyltransferase in charge of the de novo methylation of the DNA [72]. And, more recently, the significance of these findings was confirmed by observing alterations due to methionine supplementation in the liver of transition dairy cows in terms of global DNA methylation and specific region methylation of an important TF, the peroxisome proliferator-activated receptor alpha PPARα [8]. The characteristics and uniqueness of this TF within the context of the transition dairy cow were initially presented by Drackley [73], and since then, this nuclear receptor has become an exciting area of research in dairy cattle nutrigenomics (i.e., nutrient-gene interaction) [10]. Therefore, the connection between Met and PPARα upregulation through DNA methylation during the transition period is another suitable mechanism to explain the consistent improvements in performance (e.g., milk yield and DMI) observed in transition dairy cows supplemented with Met [74].

In the cellular nuclei, the DNA is normally packed in condensed structures called chromatins, consisting primarily of histone proteins, which serve as spools where the DNA winds around. Then, the genetic information contained in the DNA exists in two states: unavailable or wind around histone proteins, and available or unwound. Chromatin remodeling is the main mechanism by which DNA is wind or unwound from histones and these dynamic modifications occur by enzymatic modifications including acetylation, phosphorylation, ubiquitination, and methylation [75]. The latter being a potential mechanism through which Met can alter gene expression in dairy cows. Currently, the limited amount of data on histone methylation in dairy cows has been conducted in immune cells [76] primarily related to subclinical mastitis [77]. This work has provided nuances on the interactions between mastitis-related pathogens and histone methylation; however, dietary effects on histone methylation have not been investigated.

3. Gene regulation in beef cattle

3.1 Mutations

In order to understand how changes in a single gene could significantly alter the body structure and physiology of beef cattle, we need to focus in mutations that occur in specific areas in the bovine genome. For example, a 11-nucleotide deletion in the Myostatin gene (MSTN) determines double muscling in Belgian Blue cattle, and a single nucleotide change produces the same effect on Piedmontese cattle [78]. Furthermore, the leptin gene (LEP) presents several single nucleotide polymorphisms (SNP), like AJ236854:c.73 T > C, which induces a cysteine-to-arginine amino acid substitution that could affect protein functionality [79]. The c.73C allele of LEP is associated with higher average daily gain (ADG), lower dressing percentage, and higher marbling scores [79], which are desirable characteristics in a beef carcass. Finally, intramuscular fat deposition could also be affected by a mutation in the fatty acid synthase gene (FASN), contributing to the characteristic fatty acid composition of Japanese Black beef [80]. However, not all mutations have beneficial effects on the productivity of meat-producing animals. There are also mutations that are considered lethal, affecting, for example, the reproductive performance of females through early embryonic loss [81], or mutations that produce genetic

disorders in beef cattle [82]. The Arachnomelia syndrome in Simmental cattle [83] that produce malformations of the skeleton mainly affecting the legs, the spinal column, and the skull is an example of these genetic disorders. These types of inborn errors can be prevented nowadays with techniques like genome editing [83]. Although the implementation of this technique in the animal production industry might generate controversy, it will offer tremendous potential for breeding animals with desirable traits.

Undesired mutations present in the bovine genome are difficult to avoid when they occur. Through the implementation of selection plans and genotyping the herds in order to avoid the reproduction of carrying individuals is the most commonly utilized strategy. Similarly, selection of animals that carry biomarkers in their genomes that will make them improve their meat production and marbling efficiency is the general goal of researchers passionate about beef production. Currently, identification of relevant genetic and genomic markers is ongoing, especially for tenderness—a top priority quality attribute [84].

3.2 Nutrigenomics in beef cattle

In spite of the emergence of new alternatives to beef production (e.g., cultivated meat) in order to meet the growing world population's food demand, researchers in beef production are also focusing on techniques to regulate the bovine genome with a more natural approach (i.e., nutrigenomics). Nutrigenomics study the interactions between nutrients and genes [85], that is how the nutrients present in the diet can affect gene expression. In beef animals, nutrigenomics was widely studied [26]; following, a more detailed description of how specific nutrients regulate the expression of genes related to muscle growth and intramuscular fat deposition will be addressed.

The composition of adipose tissue produced in a meat-producing animal can be "manipulated" by diet, with some variability between breeds. For example, high silage-based diets produce less proportion of the fatty acid 18:2 n-6, with the consequent decrease in the amount of total polyunsaturated fatty acids (PUFA), as compared to low silage-based diets in subcutaneous adipose tissue [86]. This variability in fatty acids composition was attributed to greater activity of fatty acid binding protein 4 (FABP4), lipoprotein lipase (LPL), and stearoyl CoA desaturase (SCD) genes in subcutaneous adipose tissue of animals fed a low silage-based diet [86].

In the same way, we are able to regulate the fatty acid profile composition of a specific fat depot by dietary changes, researchers are trying to prioritize the growth and development of the intramuscular fat that will lead to greater marbling. Intramuscular fat starts to accumulate in the late stages of growth, as compared to other adipose tissue depots that normally develop first (i.e., visceral, intermuscular, and subcutaneous fat). The ultimate goal is to improve marbling scores that will lead to premium prices for the carcasses that classify as prime or choice according to the USDA carcass grading scale [87].

When comparing dairy and beef cattle breeds, subcutaneous fatty acid profile presents several differences in terms of fatty acids profile, probably due to the difference in the degree of fatness, which is always lower for dairy cattle [88]. The important variability in fat composition between breeds could be explained by the variability in relative SCD1 expression in subcutaneous fat [88]. SCD1 seems to have a role in a depot-specific fashion.

3.3 How management decisions can affect gene expression

The combination of early weaning and high dietary starch leads to a strong programming effect in skeletal muscle tissue, with PPARγ and CCAAT

enhancer-binding protein alpha (CEBPA) as the central coordinators of this response. The implementation of early weaning (e.g., 2 months of age) in beef calves provides a different type of diet as compared to a calf weaned normally at 6–7 months of age. Therefore, the early administration of starch in a beef producing animal produces a precocious and sustained activation of the PPARG and its target genes, leading to greater intramuscular fat deposition and consequently more carcasses grading as "Choice" [89].

Castration increases intramuscular fat (IMF) deposition, improving beef quality in cattle. In a transcriptome analysis performed in longissimus muscle (LM) samples, when comparing bulls and steers, castration showed to upregulate transcription of genes for lipogenesis, fatty acid oxidation, and also genes coding for enzymes associated with the tricarboxylic acid cycle and oxidative phosphorylation [90]. Therefore, castration shifts the transcriptome of lipid/energy metabolism to favor intramuscular fat deposition in the LM following castration.

Another beef producer decision that can affect the expression of genes is the selected calving season (i.e., different temperature/humidity index = THI). For example, two groups of pregnant cows that calved during thermoneutral temperature conditions (THI = 67.3) and cows that calved in summer season (THI = 79.9) were bleed during their transition period (i.e., cows between 3 weeks before and 3 weeks after calving) for RNA extraction of white blood cells [91]. Results showed that expression of CASP-3, BCL-2, BAK, P53, and ratio of BAX/BCL-2 in white blood cells increased during summer as compared to thermoneutral conditions, suggesting the susceptibility of these cells to apoptosis or cell death [91]. Consequently, cows calving in two different calving seasons will present differences in their inflammatory response, affecting the maternal recognition of the fetus during early pregnancy [92] or also will have a negative impact in the cow's milking ability postcalving [93].

3.4 Feedlot versus pasture

The beef industry has two main ways to produce beef: pasture-base, and grain-based or feedlot. Consuming energy above requirements helps to increase the intramuscular fat deposition in beef cattle. Feeding high concentration of cereal grains is the way to reach surplus energy that can be utilized in the rumen and the small intestine to produce volatile fatty acids for glucose and energy production. The starch present in the grain is fermented by microbes in the rumen, producing propionate (a glucose precursor), which will be the signal received by the membrane receptors present in the cell activating a cascade of events that will end up with the activation of genes related to the process of adipogenesis. This type of diet is more commonly administered during the so-called finishing or fattening phase.

When the diet is mainly based on the forage available, the rumen population consists of microbes that produce a greater proportion of acetic acid which increases the activation of 5′-prime-AMP-activated protein kinase alpha (AMPK) phosphorylation, reducing the transcriptional activity of the sterol regulatory element-binding protein 1c (SREBP1C) and the carbohydrate responsive element-binding protein (MLXIPL), which decreased the expression of lipogenic genes [94]. In beef cattle finished under a forage-based diet, the fatty acid profile varies considerably as compared to animals finished under feedlot diets. Beef finished under forage-based diets presents greater concentration of polyunsaturated fatty acids (PUFAs), especially the fatty acids with nutraceutical value (20:5 or EPA and 22:6 or DHA). These types of beef products are in the eye of consumers who care about eating healthy foods. These PUFAs mentioned above are upstream regulators of genes related to fatty acid synthesis and transport. FABP4, FASN, and PPARG are particularly activated by these additional PUFAs

generated due to the administration of forage-based diets [95]. Even though we could expect a greater proportion of IMF due to the mentioned additional PUFAs, there is an overall lower fat content in a grass-fed beef product. Furthermore, grass-fed beef is known to have a different flavor and aroma as compared to grain-fed beef when cook on the grill [96].

3.5 Fetal programming in beef cattle

The fetal programming concept is related to the important physiological changes that can occur due to environmental/nutritional events during prenatal development. For example, nutrient restriction of 85% as compared to 140% of total metabolizable energy requirements during the second half of gestation can alter in fetal muscle the expression of both, myogenic and adipogenic genes, without apparent differences in fetal phenotype [97].

The canonical Wnt pathway, a β-catenin-dependent signaling pathway called the Wnt/β-catenin signaling pathway is key in establishing the fate of the undifferentiated stem cells; hence, β-catenin plays an essential role in the regulation of embryonic, postnatal, and oncogenic growth of many tissues. If the β-catenin pathway is blocked, the total number of myocytes will be reduced, and the differentiation of mesenchymal stem cells into mature adipocytes will be potentiated [98]. In the same way, adipogenesis is initiated around mid-gestation in ruminant animals; therefore, a strategic maternal nutritional plan in order to enhance the number of mesenchymal cells committed to adipogenesis will increase the number of intramuscular adipocytes in a process known as hyperplasia; therefore, this outcome will be translated as more marbling in the offspring postnatally. PPARG alone can stimulate adipocyte differentiation [99], although the continuation of this process is regulated by many PPARG target genes [100].

Bioinformatics analysis revealed a pseudoinflammatory process in early-weaned beef calves during their growing phase [101], which it is associated with the activation of the innate immune system presumably due to macrophage infiltration of intramuscular fat [101], which is a typical obesity symptom. These results could be considered as a biological sign of a precocious beginning of the adipogenic metabolic machinery in young beef calves.

3.6 Epigenetics regulations in beef cattle

Changes caused by chromatin (the complex of DNA and histone proteins that forms chromosomes within the nucleus) modification due to environmental factors is called epigenetics [102]. Another epigenetics regulation approach is through microRNAs, which are endogenous noncoding small RNA molecules (20–24 base pairs) that prevent the production of a particular protein by binding to and destroying the messenger RNA that would have produced the protein [103].

Epigenetics regulation is based on chromatin remodeling rather than alteration of the DNA code. With the aim to identify methylated genes affecting bovine growth, an elegant study provides a genome-wide landscape of DNA methylomes and their relationship with mRNA and miRNA for fetal and adult muscle of Chinese Qinchuan beef cattle [104]. Presence of SNPs in epigenetic-related genes was analyzed in different beef breeds. Interestingly, three DNA (cytosine-5)-methyltransferases (DNMTs), DNMT1, DNMT3a, and DNMT3b were found significantly associated with beef quality parameters on the carcass. In particular, DNMT3b presented five SNPs related to carcass traits, becoming a potential candidate gene for beef quality improvement [105].

In the bovine genome, no microRNAs were identified on chromosome Y, while microRNA related to adipose tissue are expressed in chromosome X [106]. This

could be a reason that explains the sexual differences in fat metabolism in mammals. Furthermore, there are highly expressed microRNAs for beef adipose tissue, miR-378 which was found upregulated in steers with high levels of subcutaneous fat [107] and miR-2478 [108] which potentially targets ELOVL fatty acid elongase 6 (ELOVL6) and stearoyl CoA desaturase (SCD), is bovine specific and had higher expression in grass-fed as compared to grain-fed cattle.

A study in bovine skeletal muscle development used next-generation small RNA sequencing, a total of 512 miRNAs were identified [109]. Thirty-six miRNAs were differentially expressed between prenatal and postnatal stages of muscle development including several myomiRs (miR-1, miR-206, and let-7 families). Compared to miRNA expression between different muscle tissues, 14 miRNAs were upregulated and 22 miRNAs were downregulated in the muscle of postnatal stage [109]. Furthermore, a genome-wide landscape of DNA methylomes and their relationship with mRNA and miRNA for bovine fetal and adult muscle recently discovered will provide a solid basis for exploring the epigenetic mechanisms of muscle growth and development [104].

Muscling in cattle influenced by genetic background, ultimately affecting beef yield is of major interest to the beef industry. The best alternative to promote muscle development is through satellite cell proliferation [110]. In fact, myoblast or satellite cells are utilized for the proliferation and differentiation of cultured meat [111]. The transcription factor Sp1, an activator of myosin D (MyoD) and a suppressor of cyclin-dependent kinase inhibitor 1A (CDKN1A), plays an important role in bovine muscle cell proliferation and differentiation. This transcription factor is a target of miR-128 and, if this microRNA is overexpressed, it inhibits muscle satellite cell proliferation and differentiation [112]. Furthermore, miR-1 and miR-206 facilitate bovine skeletal muscle satellite cell myogenic differentiation by restricting the expression of their target genes, and that inhibition of miR-1 and miR-206 increased the paired box 7 (Pax7) and histone deacetylase 4 (HDAC4) protein levels enhancing satellite cell proliferation [113].

In Biceps femoris muscle of Japanese Shorthorn cattle, a grazing period up to 4 months increased the expression of miR-208b, which has a muscle fiber type-associated role. Furthermore, a target for miR-208b, MyoD, a myogenic regulatory factor associated with the shifting of muscle property to the fast type, had lower expression in the grazed cattle after 4 months of grazing, as compared to feedlot cattle. During skeletal muscle adaptation to grazing, miR-206 expression remained higher as compared to housed animals in which it decreases [114]. MiR-206 is known as the skeletal muscle-specific myomiRNA [115].

Offspring's health depends on the maternal body condition at mid-gestation, which will make them be more predisposed to develop obesity at an early age, which in beef production is desirable. Fetal intramuscular adipogenesis was enhanced at mid-gestation due to alteration of microRNA expression; downregulation of let-7g was the main cause for this outcome [116]. This microRNA inhibits the mRNA expression of PPARG and CEBPA, both important regulators of adipogenesis [117]. Furthermore, cow plane of nutrition during the last third of gestation showed epigenetic effects on the offspring's skeletal muscle through the downregulation of miR-34a that has a role on the activation of cell cycle arrest by suppressing SIRT1 expression, which promotes adipocyte differentiation [118].

4. Molecular nutrition in ruminants

The importance of ruminants to the world food security is reflected in their contribution to the demand for animal protein around the world and particularly

in developing countries, and such demand is expected to increase in the future [119]. To face this demand, advancements in ruminant nutrition and physiology will require improvements on feed efficiency and development of novel functional foods from ruminants by enriching specific compounds associated with health benefits in humans. The latter will need a deep understanding and wealth of knowledge of molecular regulatory mechanisms in response to physiological conditions and nutrition. In this context, this vast amount of multilayered data in terms of mRNA, proteins, metabolites, and phenotypes can only be undertaken with powerful tools such as omics technologies and bioinformatics. In fact, these are the foundations of modern system biology, a field of study with the aim to enhance the understanding of complex biological models and interactions occurring within cells and tissues. Understanding this complexity and the outcomes of nutritional interventions and physiological conditions will allow the formulation of novel theories and ideas to enhance feed efficiency, development of new functional foods derived from ruminant products, and reduce carbon footprint.

Even though the outcome is different, there are similarities in dairy and beef cattle from a nutrigenomic perspective. For instance, both the synthesis of milk fat in dairy cows and the synthesis of intramuscular fat in beef steers are regulated by a similar network of TF. Nutrients or stimulus received with the diet (PUFAs, insulin, etc.), activates PPARα in the liver of the dairy cow and PPARγ in the intramuscular preadipocyte of a beef steer. The activated PPARs form a heterodimer with retinoic X receptor alpha (RXRA), leading to the upregulation of their lipogenesis-related target genes (**Figure 2**). Furthermore, in the same way, the activation of the PI3K/Akt/mTOR signaling pathway will lead to the synthesis of milk protein in dairy cows [120], the activation of the same metabolic pathway might lead to muscle hypertrophy in beef cattle, but this is a concept that has not been completely elucidated [121]. It is also worth to mention the importance of fatty acid binding proteins (FABPs) in ruminants, which bind and transport LCFA. FABP4 affects milk yield and milk protein content, both economically important traits in the dairy industry [122], and FABP4 also presents gene polymorphisms that have been associated with meat quality traits in beef cattle [123].

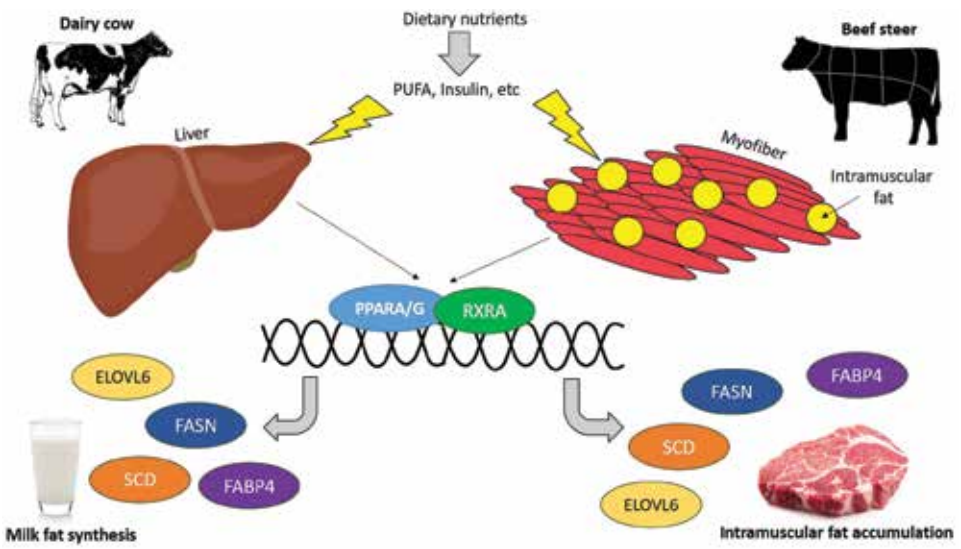

Figure 2.
Example of nutrigenomics linkage between beef and dairy cattle.

In a study aiming to use the fibroblast model to explore differences between a dairy breed (Holstein) and a beef breed (Angus) in their innate responses to LPS exposure, several immune-associated differentially expressed genes between breeds were found [124]. Within them, TLR4, which is the extracellular receptor responsible for recognition of LPS presented higher level of expression in Holstein cows as compared to Angus, suggests the Holstein animals will detect and respond to Gram-negative bacteria more vigorously than Angus animals.

Finally, epigenetic differences between beef and dairy cattle could also be observed mainly because of the different environments the offspring are exposed after birth. While a beef calf usually stays with the dam approximately until 6–7 months old, a dairy calf is separated from its mother as soon as it finishes consuming colostrum or earlier. Although there are studies that started to analyze the epigenetic differences between breeds [124], this is a promising area that needs to be studied in deep.

5. Conclusions

The general nutrigenomic model for ruminants needs to be updated based on emerging nuance information with the ever-growing pace in ruminant nutrition research with "omics" technologies. The dissection of what intermediate components or processes such as intermediate metabolites, signal transduction, TF, etc., are utilized by specific nutrients will allow for accurate predictions of the nutrigenomic outcomes of such nutrients in a practical setting. However, the multilayered and multifactorial nature of the nutrigenomic model will require the implementation of additional tools such as system biology and network theory in order to have a more holistic approach to understand how nutrients regulate milk synthesis or skeletal muscle gain and marbling.

One of the greatest challenges in ruminant nutrigenomics is to account for the final products from rumen fermentation, where several factors such as rate of passage, intake, particle size of the diet can affect rumen fermentation and kinetics. The latter can be avoided by feeding nutrients encapsulated or protected from ruminal degradation; however, this does not eliminate the need to account for the substantial impact the rumen fermentation and its products may have on the overall nutrigenomic effect from a particular diet. Because of this reason, the resurgence of the field of the microbiome in ruminant nutrition research promise to add valuable information on rumen microbes response to nutrients in the diet and correlate this with final nutrigenomic responses at the whole animal level.

Our understanding of the impact of nutrition on regulatory mechanisms at the cellular level in the ruminant animal has grown an accelerated pace over the last decades. As pointed out by Drackley 12 year ago [125], the marriage between "omics" technologies with measurements of tissue metabolism and the final performance (e.g., milk yield and skeletal muscle gain) has been enlightening and essential to identify key responses to nutritional changes and physiology. However, there is still too much to learn in the complex nutrient-gene interactions in the context of the ruminant animal. The future of nutrigenomics in ruminants is to develop technologies and algorithms to predict the final molecular outcomes of nutrients and diets fed to ruminants in a practical setting. This monumental task can only be accomplished by generating a wealth of knowledge in several orders of magnitude of what we currently have on nutrigenomics in ruminants.

Author details

Johan S. Osorio[1*] and Sonia J. Moisa[2]

1 South Dakota State University, Brookings, SD, USA

2 Auburn University, Auburn, AL, USA

*Address all correspondence to: johan.osorio@sdstate.edu

References

[1] Nobel S, Abrahmsen L, Oppermann U. Metabolic conversion as a pre-receptor control mechanism for lipophilic hormones. European Journal of Biochemistry. 2001;**268**(15):4113-4125. DOI: 10.1046/j.1432-1327.2001.02359.x

[2] Clarke SD. Nutrient regulation of gene and protein expression. Current Opinion in Clinical Nutrition and Metabolic Care. 1999;**2**(4):287-289

[3] Eastwood MA. A molecular biological basis for the nutritional and pharmacological benefits of dietary plants. Q JM. 2001;**94**(1):45-48

[4] Dauncey MJ, White P, Burton KA, Katsumata M. Nutrition-hormone receptor-gene interactions: Implications for development and disease. The The Proceedings of the Nutrition Society 2001;**60**(1):63-72

[5] Jacobs MN, Lewis DF. Steroid hormone receptors and dietary ligands: A selected review. The Proceedings of the Nutrition Society. 2002;**61**(1):105-122

[6] Rakhshandehroo M, Knoch B, Muller M, Kersten S. Peroxisome proliferator-activated receptor alpha target genes. PPAR Research. 2010;**2010**. DOI: 10.1155/2010/612089

[7] Osorio JS, Lohakare J, Bionaz M. Biosynthesis of milk fat, protein, and lactose: Roles of transcriptional and posttranscriptional regulation. Physiological Genomics. 2016;**48**(4):231-256. DOI: 10.1152/physiolgenomics.00016.2015

[8] Osorio JS, Jacometo CB, Zhou Z, Luchini D, Cardoso FC, Loor JJ. Hepatic global DNA and peroxisome proliferator-activated receptor alpha promoter methylation are altered in peripartal dairy cows fed rumen-protected methionine. Journal of Dairy Science. 2016;**99**(1):234-244. DOI: 10.3168/jds.2015-10157

[9] Li R, Beaudoin F, Ammah AA, Bissonnette N, Benchaar C, Zhao X, et al. Deep sequencing shows microRNA involvement in bovine mammary gland adaptation to diets supplemented with linseed oil or safflower oil. BMC Genomics. 2015;**16**:884. DOI: 10.1186/s12864-015-1965-7

[10] Bionaz M, Chen S, Khan MJ, Loor JJ. Functional role of PPARs in ruminants: Potential targets for fine-tuning metabolism during growth and lactation. PPAR Research. 2013;**2013**:684159. DOI: 10.1155/2013/684159

[11] Burris TP, Solt LA, Wang Y, Crumbley C, Banerjee S, Griffett K, et al. Nuclear receptors and their selective pharmacologic modulators. Pharmacological Reviews. 2013;**65**(2):710-778. DOI: 10.1124/pr.112.006833

[12] Escher P, Wahli W. Peroxisome proliferator-activated receptors: Insight into multiple cellular functions. Mutation Research. 2000;**448**(2):121-138

[13] Olefsky JM, Saltiel AR. PPAR gamma and the treatment of insulin resistance. Trends in Endocrinology and Metabolism. 2000;**11**(9):362-368

[14] Desvergne B, Wahli W. Peroxisome proliferator-activated receptors: Nuclear control of metabolism. Endocrine Reviews. 1999;**20**(5):649-688. DOI: 10.1210/edrv.20.5.0380

[15] Desvergne B, Michalik L, Wahli W. Transcriptional regulation of metabolism. Physiological Reviews. 2006;**86**(2):465-514. DOI: 10.1152/physrev.00025.2005

[16] Sevane N, Armstrong E, Cortes O, Wiener P, Wong RP, Dunner S, et al. Association of bovine meat quality traits with genes included in the PPARG and PPARGC1A networks. Meat Science. 2013;**94**(3):328-335. DOI: 10.1016/j.meatsci.2013.02.014

[17] Moisa SJ, Shike DW, Meteer WT, Keisler D, Faulkner DB, Loor JJ. Yin yang 1 and adipogenic gene network expression in longissimus muscle of beef cattle in response to nutritional management. Gene Regulation and Systems Biology. 2013;**7**:71-83. DOI: 10.4137/GRSB.S11783

[18] Faulconnier Y, Bernard L, Boby C, Domagalski J, Chilliard Y, Leroux C. Extruded linseed alone or in combination with fish oil modifies mammary gene expression profiles in lactating goats. Animal: An International Journal of Animal Bioscience. 2018;**12**(8):1564-1575. DOI: 10.1017/S1751731117002816

[19] Agazzi A, Invernizzi G, Campagnoli A, Ferroni M, Fanelli A, Cattaneo D, et al. Effect of different dietary fats on hepatic gene expression in transition dairy goats. Small Ruminant Research. 2010;**93**(1):31-40. DOI: 10.1016/j.smallrumres.2010.04.027

[20] McFadden JW, Corl BA. Activation of liver X receptor (LXR) enhances de novo fatty acid synthesis in bovine mammary epithelial cells. Journal of Dairy Science. 2010;**93**(10):4651-4658. DOI: 10.3168/jds.2010-3202

[21] Harvatine KJ, Bauman DE. SREBP1 and thyroid hormone responsive spot 14 (S14) are involved in the regulation of bovine mammary lipid synthesis during diet-induced milk fat depression and treatment with CLA. The Journal of Nutrition. 2006;**136**(10):2468-2474

[22] Wang W, Luo J, Zhong Y, Lin XZ, Shi HB, Zhu JJ, et al. Goat liver X receptor alpha, molecular cloning, functional characterization and regulating fatty acid synthesis in epithelial cells of goat mammary glands. Gene. 2012;**505**(1):114-120. DOI: 10.1016/j.gene.2012.05.028

[23] Yang YZ, Subauste JS, Koenig RJ. Retinoid X receptor alpha binds with the highest affinity to an imperfect direct repeat response element. Endocrinology. 1995;**136**(7):2896-2903. DOI: 10.1210/endo.136.7.7789315

[24] Osorio JS, Vailati-Riboni M, Palladino A, Luo J, Loor JJ. Application of nutrigenomics in small ruminants: Lactation, growth, and beyond. Small Ruminant Research. 2017;**154**:29-44. DOI: 10.1016/j.smallrumres.2017.06.021

[25] Bauman DE, Harvatine KJ, Lock AL. Nutrigenomics, rumen-derived bioactive fatty acids, and the regulation of milk fat synthesis. Annual Review of Nutrition. 2011;**31**:299-319. DOI: 10.1146/annurev.nutr.012809.104648

[26] Ladeira MM, Schoonmaker JP, Gionbelli MP, Dias JC, Gionbelli TR, Carvalho JR, et al. Nutrigenomics and beef quality: A review about lipogenesis. International Journal of Molecular Sciences. 2016;**17**(6). DOI: 10.3390/ijms17060918

[27] Xu W, Wang F, Yu Z, Xin F. Epigenetics and cellular metabolism. Genetics & Epigenetics. 2016;**8**:43-51. DOI: 10.4137/GEG.S32160

[28] Wang X, Gu Z, Jiang H. MicroRNAs in farm animals. Animal: An International Journal of Animal Bioscience. 2013;**7**(10):1567-1575. DOI: 10.1017/S1751731113001183

[29] Koufariotis LT, Chen YP, Chamberlain A, Vander Jagt C, Hayes BJ. A catalogue of novel bovine long noncoding RNA across 18 tissues. PLoS One. 2015;**10**(10):e0141225. DOI: 10.1371/journal.pone.0141225

[30] Wang M, Moisa S, Khan MJ, Wang J, Bu D, Loor JJ. MicroRNA expression patterns in the bovine mammary gland are affected by stage of lactation. Journal of Dairy Science. 2012;**95**(11):6529-6535. DOI: 10.3168/jds.2012-5748

[31] Tong C, Chen Q , Zhao L, Ma J, Ibeagha-Awemu EM, Zhao X. Identification and characterization of long intergenic noncoding RNAs in bovine mammary glands. BMC Genomics. 2017;**18**(1):468. DOI: 10.1186/s12864-017-3858-4

[32] Evershed RP, Payne S, Sherratt AG, Copley MS, Coolidge J, Urem-Kotsu D, et al. Earliest date for milk use in the near east and southeastern Europe linked to cattle herding. Nature. 2008;**455**(7212):528-531. DOI: 10.1038/nature07180

[33] Wiley AS. The globalization of Cow's milk production and consumption: Biocultural perspectives. Ecology of Food and Nutrition. 2007;**46**(3-4):281-312. DOI: 10.1080/03670240701407657

[34] Anderson SM, Rudolph MC, McManaman JL, Neville MC. Key stages in mammary gland development. Secretory activation in the mammary gland: It's not just about milk protein synthesis! Breast Cancer Research. 2007;**9**(1):204. DOI: 10.1186/bcr1653

[35] Lemay DG, Neville MC, Rudolph MC, Pollard KS, German JB. Gene regulatory networks in lactation: Identification of global principles using bioinformatics. BMC Systems Biology. 2007;**1**:56. DOI: 10.1186/1752-0509-1-56

[36] Zhou Y, Gong W, Xiao J, Wu J, Pan L, Li X, et al. Transcriptomic analysis reveals key regulators of mammogenesis and the pregnancy-lactation cycle. Science China. Life Sciences. 2014;**57**(3):340-355. DOI: 10.1007/s11427-013-4579-9

[37] Anantamongkol U, Charoenphandhu N, Wongdee K, Teerapornpuntakit J, Suthiphongchai T, Prapong S, et al. Transcriptome analysis of mammary tissues reveals complex patterns of transporter gene expression during pregnancy and lactation. Cell Biology International. 2010;**34**(1):67-74. DOI: 10.1042/CBI20090023

[38] Rodriguez-Cruz M, Coral-Vazquez RM, Hernandez-Stengele G, Sanchez R, Salazar E, Sanchez-Munoz F, et al. Identification of putative ortholog gene blocks involved in gestant and lactating mammary gland development: A rodent cross-species microarray transcriptomics approach. International Journal of Genomics. 2013;**2013**:624681. DOI: 10.1155/2013/624681

[39] Bionaz M, Periasamy K, Rodriguez-Zas SL, Everts RE, Lewin HA, Hurley WL, et al. Old and new stories: Revelations from functional analysis of the bovine mammary transcriptome during the lactation cycle. PLoS One. 2012;**7**(3):e33268. DOI: 10.1371/journal.pone.0033268

[40] Finucane KA, McFadden TB, Bond JP, Kennelly JJ, Zhao FQ. Onset of lactation in the bovine mammary gland: Gene expression profiling indicates a strong inhibition of gene expression in cell proliferation. Functional & Integrative Genomics. 2008;**8**(3):251-264. DOI: 10.1007/s10142-008-0074-y

[41] Wickramasinghe S, Rincon G, Islas-Trejo A, Medrano JF. Transcriptional profiling of bovine milk using RNA sequencing. BMC Genomics. 2012;**13**:45. DOI: 10.1186/1471-2164-13-45

[42] Singh M, Thomson PC, Sheehy PA, Raadsma HW. Comparative transcriptome analyses reveal conserved and distinct mechanisms in ovine and bovine lactation. Functional & Integrative Genomics. 2013;**13**(1):115-131. DOI: 10.1007/s10142-012-0307-y

[43] Faucon F, Rebours E, Bevilacqua C, Helbling JC, Aubert J, Makhzami S, et al. Terminal differentiation of goat mammary tissue during pregnancy requires the expression of genes involved in immune functions. Physiological Genomics. 2009;**40**(1):61-82. DOI: 10.1152/physiolgenomics.00032.2009

[44] Lemay DG, Ballard OA, Hughes MA, Morrow AL, Horseman ND, Nommsen-Rivers LA. RNA sequencing of the human milk fat layer transcriptome reveals distinct gene expression profiles at three stages of lactation. PLoS One. 2013;**8**(7):e67531. DOI: 10.1371/journal.pone.0067531

[45] Maningat PD, Sen P, Rijnkels M, Sunehag AL, Hadsell DL, Bray M, et al. Gene expression in the human mammary epithelium during lactation: The milk fat globule transcriptome. Physiological Genomics. 2009;**37**(1):12-22. DOI: 10.1152/physiolgenomics.90341.2008

[46] Mohammad MA, Haymond MW. Regulation of lipid synthesis genes and milk fat production in human mammary epithelial cells during secretory activation. American Journal of Physiology. Endocrinology and Metabolism. 2013;**305**(6):E700-E716. DOI: 10.1152/ajpendo.00052.2013

[47] Shu DP, Chen BL, Hong J, Liu PP, Hou DX, Huang X, et al. Global transcriptional profiling in porcine mammary glands from late pregnancy to peak lactation. OMICS. 2012;**16**(3):123-137. DOI: 10.1089/omi.2011.0116

[48] Lefevre CM, Digby MR, Whitley JC, Strahm Y, Nicholas KR. Lactation transcriptomics in the Australian marsupial, *Macropus eugenii*: Transcript sequencing and quantification. BMC Genomics. 2007;**8**:417. DOI: 10.1186/1471-2164-8-417

[49] Menzies KK, Lefevre C, Sharp JA, Macmillan KL, Sheehy PA, Nicholas KR. A novel approach identified the FOLR1 gene, a putative regulator of milk protein synthesis. Mammalian Genome. 2009;**20**(8):498-503. DOI: 10.1007/s00335-009-9207-4

[50] Bionaz M, Loor JJ. Gene networks driving bovine milk fat synthesis during the lactation cycle. BMC Genomics. 2008;**9**:366. DOI: 10.1186/1471-2164-9-366

[51] Bionaz M, Loor JJ. Gene networks driving bovine mammary protein synthesis during the lactation cycle. Bioinformatics and Biology Insights. 2011;**5**:83-98. DOI: 10.4137/bbi.s7003

[52] Dawson KA. Nutrigenomics: Feeding the genes for improved fertility. Animal Reproduction Science. 2006;**96**(3-4):312-322. DOI: 10.1016/j.anireprosci.2006.08.009

[53] Fekete SG, Brown DL. Veterinary aspects and perspectives of nutrigenomics: A critical review. Acta Veterinaria Hungarica. 2007;**55**(2):229-239. DOI: 10.1556/AVet.55.2007.2.9

[54] de Godoy MR, Swanson KS. Companion animals symposium: Nutrigenomics: Using gene expression and molecular biology data to understand pet obesity. Journal of Animal Science. 2013;**91**(6):2949-2964. DOI: 10.2527/jas.2012-5860

[55] Muller M, Kersten S. Nutrigenomics: Goals and strategies. Nature Reviews. Genetics. 2003;**4**(4):315-322. DOI: 10.1038/nrg1047

[56] Coffey SG. Prospects for improving the nutritional quality of dairy and meat products. Forum of Nutrition. 2007;**60**:183-195. DOI: 10.1159/0000107195

[57] Shlyueva D, Stampfel G, Stark A. Transcriptional enhancers: From

properties to genome-wide predictions. Nature Reviews. Genetics. 2014;**15**(4):272-286. DOI: 10.1038/nrg3682

[58] Vaquerizas JM, Kummerfeld SK, Teichmann SA, Luscombe NM. A census of human transcription factors: Function, expression and evolution. Nature Reviews. Genetics. 2009;**10**(4):252-263. DOI: 10.1038/nrg2538

[59] Zhang HM, Liu T, Liu CJ, Song S, Zhang X, Liu W, et al. AnimalTFDB 2.0: A resource for expression, prediction and functional study of animal transcription factors. Nucleic Acids Research. 2015;**43**(Database issue):D76-D81. DOI: 10.1093/nar/gku887

[60] Cheatle Jarvela AM, Hinman VF. Evolution of transcription factor function as a mechanism for changing metazoan developmental gene regulatory networks. EvoDevo. 2015;**6**(1):3. DOI: 10.1186/2041-9139-6-3

[61] Chawla A, Repa JJ, Evans RM, Mangelsdorf DJ. Nuclear receptors and lipid physiology: Opening the X-files. Science. 2001;**294**(5548):1866-1870. DOI: 10.1126/science.294.5548.1866

[62] Bionaz M, Osorio JS, Loor JJ. TRIENNIAL LACTATION SYMPOSIUM: Nutrigenomics in dairy cows: Nutrients, transcription factors, and techniques. Journal of Animal Science. 2015;**93**(12):5531-5553. DOI: 10.2527/jas.2015-9192

[63] Khan SA, Vanden Heuvel JP. Role of nuclear receptors in the regulation of gene expression by dietary fatty acids (review). The Journal of Nutritional Biochemistry. 2003;**14**(10):554-567

[64] Minucci S, Leid M, Toyama R, Saint-Jeannet JP, Peterson VJ, Horn V, et al. Retinoid X receptor (RXR) within the RXR-retinoic acid receptor heterodimer binds its ligand and enhances retinoid-dependent gene expression. Molecular and Cellular Biology. 1997;**17**(2):644-655

[65] Landes N, Pfluger P, Kluth D, Birringer M, Ruhl R, Bol GF, et al. Vitamin E activates gene expression via the pregnane X receptor. Biochemical Pharmacology. 2003;**65**(2):269-273

[66] Nakamura MT, Yudell BE, Loor JJ. Regulation of energy metabolism by long-chain fatty acids. Progress in Lipid Research. 2014;**53**:124-144. DOI: 10.1016/j.plipres.2013.12.001

[67] Vanselow J, Yang W, Herrmann J, Zerbe H, Schuberth HJ, Petzl W, et al. DNA-remethylation around a STAT5-binding enhancer in the alphaS1-casein promoter is associated with abrupt shutdown of alphaS1-casein synthesis during acute mastitis. Journal of Molecular Endocrinology. 2006;**37**(3):463-477. DOI: 10.1677/jme.1.02131

[68] Martinov MV, Vitvitsky VM, Banerjee R, Ataullakhanov FI. The logic of the hepatic methionine metabolic cycle. Biochimica et Biophysica Acta. 2010;**1804**(1):89-96. DOI: 10.1016/j.bbapap.2009.10.004

[69] Lu SC, Mato JM. S-adenosylmethionine in liver health, injury, and cancer. Physiological Reviews. 2012;**92**(4):1515-1542. DOI: 10.1152/physrev.00047.2011

[70] Shima H, Matsumoto M, Ishigami Y, Ebina M, Muto A, Sato Y, et al. S-adenosylmethionine synthesis is regulated by selective N(6)-adenosine methylation and mRNA degradation involving METTL16 and YTHDC1. Cell Reports. 2017;**21**(12):3354-3363. DOI: 10.1016/j.celrep.2017.11.092

[71] Kass SU, Pruss D, Wolffe AP. How does DNA methylation repress

transcription? Trends in Genetics. 1997;**13**(11):444-449

[72] Osorio JS, Ji P, Drackley JK, Luchini D, Loor JJ. Smartamine M and MetaSmart supplementation during the peripartal period alter hepatic expression of gene networks in 1-carbon metabolism, inflammation, oxidative stress, and the growth hormone-insulin-like growth factor 1 axis pathways. Journal of Dairy Science. 2014;**97**(12):7451-7464. DOI: 10.3168/jds.2014-8680

[73] Drackley JK. ADSA Foundation Scholar AwardBiology of dairy cows during the transition period: The final frontier? Journal of Dairy Science. 1999;**82**(11):2259-2273

[74] Osorio JS, Ji P, Drackley JK, Luchini D, Loor JJ. Supplemental Smartamine M or MetaSmart during the transition period benefits postpartal cow performance and blood neutrophil function. Journal of Dairy Science. 2013;**96**(10):6248-6263. DOI: 10.3168/jds.2012-5790

[75] Singh K, Erdman RA, Swanson KM, Molenaar AJ, Maqbool NJ, Wheeler TT, et al. Epigenetic regulation of milk production in dairy cows. Journal of Mammary Gland Biology and Neoplasia. 2010;**15**(1):101-112. DOI: 10.1007/s10911-010-9164-2

[76] He Y, Yu Y, Zhang Y, Song J, Mitra A, Zhang Y, et al. Genome-wide bovine H3K27me3 modifications and the regulatory effects on genes expressions in peripheral blood lymphocytes. PLoS One. 2012;**7**(6):e39094. DOI: 10.1371/journal.pone.0039094

[77] He Y, Song M, Zhang Y, Li X, Song J, Zhang Y, et al. Whole-genome regulation analysis of histone H3 lysin 27 trimethylation in subclinical mastitis cows infected by *Staphylococcus aureus*. BMC Genomics. 2016;**17**:565. DOI: 10.1186/s12864-016-2947-0

[78] McPherron AC, Lee SJ. Double muscling in cattle due to mutations in the myostatin gene. Proceedings of the National Academy of Sciences of the United States of America. 1997;**94**(23):12457-12461

[79] Di Stasio L, Brugiapaglia A, Galloni M, Destefanis G, Lisa C. Effect of the leptin c.73T>C mutation on carcass traits in beef cattle. Animal Genetics. 2007;**38**(3):316-317. DOI: 10.1111/j.1365-2052.2007.01595.x

[80] Abe T, Saburi J, Hasebe H, Nakagawa T, Misumi S, Nade T, et al. Novel mutations of the FASN gene and their effect on fatty acid composition in Japanese black beef. Biochemical Genetics. 2009;**47**(5-6):397-411. DOI: 10.1007/s10528-009-9235-5

[81] Hoff JL, Decker JE, Schnabel RD, Taylor JF. Candidate lethal haplotypes and causal mutations in Angus cattle. BMC Genomics. 2017;**18**(1):799. DOI: 10.1186/s12864-017-4196-2

[82] Ciepłoch A, Rutkowska K, Oprządek J, Poławska E. Genetic disorders in beef cattle: A review. Genes & Genomics. 2017;**39**:461-471

[83] Bhat SA, Malik AA, Ahmad SM, Shah RA, Ganai NA, Shafi SS, et al. Advances in genome editing for improved animal breeding: A review. Veterinary World. 2017;**10**(11):1361-1366. DOI: 10.14202/vetworld.2017.1361-1366

[84] Cassar-Malek I, Picard B. Expression marker-based strategy to improve beef quality. ScientificWorldJournal. 2016;**2016**:2185323. DOI: 10.1155/2016/2185323

[85] Neeha VS, Kinth P. Nutrigenomics research: A review. Journal of Food Science and Technology. 2013;**50**(3):415-428. DOI: 10.1007/s13197-012-0775-z

[86] da Costa AS, Pires VM, Fontes CM, Mestre Prates JA. Expression of genes controlling fat deposition in two genetically diverse beef cattle breeds fed high or low silage diets. BMC Veterinary Research. 2013;**9**:118. DOI: 10.1186/1746-6148-9-118

[87] Hale DS, Goodson, K, Savell, JW. USDA Beef Quality and Yield Grades [Internet]. College Station, TX 77843-24712013. Available from: https://meat.tamu.edu/beefgrading/

[88] Gamarra D, Aldai N, Arakawa A, Barron LJR, Lopez-Oceja A, de Pancorbo MM, et al. Distinct correlations between lipogenic gene expression and fatty acid composition of subcutaneous fat among cattle breeds. BMC Veterinary Research. 2018;**14**. DOI: 10.1186/s12917-018-1481-5

[89] Moisa SJ, Shike DW, Faulkner DB, Meteer WT, Keisler D, Loor JJ. Central role of the PPARgamma gene network in coordinating beef cattle intramuscular adipogenesis in response to weaning age and nutrition. Gene Regulation and Systems Biology. 2014;**8**:17-32. DOI: 10.4137/GRSB.S11782

[90] Jeong J, Bong J, Kim GD, Joo ST, Lee HJ, Baik M. Transcriptome changes favoring intramuscular fat deposition in the longissimus muscle following castration of bulls. Journal of Animal Science. 2013;**91**(10):4692-4704. DOI: 10.2527/jas.2012-6089

[91] Somal A, Aggarwal A, Upadhyay RC. Effect of thermal stress on expression profile of apoptosis related genes in peripheral blood mononuclear cells of transition Sahiwal cow. Iranian Journal of Veterinary Research. 2015;**16**(2):137-143

[92] Oliveira LJ, Barreto RS, Perecin F, Mansouri-Attia N, Pereira FT, Meirelles FV. Modulation of maternal immune system during pregnancy in the cow. Reproduction in Domestic Animals. 2012;**47**(Suppl 4):384-393. DOI: 10.1111/j.1439-0531.2012.02102.x

[93] Froidmont E, Mayeres P, Picron P, Turlot A, Planchon V, Stilmant D. Association between age at first calving, year and season of first calving and milk production in Holstein cows. Animal. 2013;**7**(4):665-672. DOI: 10.1017/S1751731112001577

[94] Li XW, Chen H, Guan Y, Li XB, Lei LC, Liu JX, et al. Acetic acid activates the AMP-activated protein kinase signaling pathway to regulate lipid metabolism in bovine hepatocytes. PLoS One. 2013;**8**(7):e67880. DOI: 10.1371/journal.pone.0067880

[95] Buchanan JW, Garmyn AJ, Hilton GG, VanOverbeke DL, Duan Q , Beitz DC, et al. Comparison of gene expression and fatty acid profiles in concentrate and forage finished beef. Journal of Animal Science. 2013;**91**(1):1-9. DOI: 10.2527/jas.2012-5154

[96] Daley CA, Abbott A, Doyle PS, Nader GA, Larson S. A review of fatty acid profiles and antioxidant content in grass-fed and grain-fed beef. Nutrition Journal. 2010;**9**:10. DOI: 10.1186/1475-2891-9-10

[97] Paradis F, Wood KM, Swanson KC, Miller SP, McBride BW, Fitzsimmons C. Maternal nutrient restriction in mid-to-late gestation influences fetal mRNA expression in muscle tissues in beef cattle. BMC Genomics. 2017;**18**(1):632. DOI: 10.1186/s12864-017-4051-5

[98] Du M, Tong J, Zhao J, Underwood KR, Zhu M, Ford SP, et al. Fetal programming of skeletal muscle development in ruminant animals. Journal of Animal Science. 2010;**88**(13 Suppl):E51-E60. DOI: 10.2527/jas.2009-2311

[99] Siersbaek R, Nielsen R, Mandrup S. PPARgamma in adipocyte

differentiation and metabolism--novel insights from genome-wide studies. FEBS Letters. 2010;**584**(15):3242-3249. DOI: 10.1016/j.febslet.2010.06.010

[100] Lee JE, Ge K. Transcriptional and epigenetic regulation of PPARgamma expression during adipogenesis. Cell & Bioscience. 2014;**4**:29. DOI: 10.1186/2045-3701-4-29

[101] Moisa SJ, Shike DW, Shoup L, Rodriguez-Zas SL, Loor JJ. Maternal plane of nutrition during late gestation and weaning age alter Angus x Simmental offspring longissimus muscle transcriptome and intramuscular fat. PLoS One. 2015;**10**(7):e0131478. DOI: 10.1371/journal.pone.0131478

[102] Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. Cell Research. 2011;**21**(3):381-395. DOI: 10.1038/cr.2011.22

[103] Wang H, Zheng Y, Wang G, Li H. Identification of microRNA and bioinformatics target gene analysis in beef cattle intramuscular fat and subcutaneous fat. Molecular BioSystems. 2013;**9**(8):2154-2162. DOI: 10.1039/c3mb70084d

[104] Huang YZ, Sun JJ, Zhang LZ, Li CJ, Womack JE, Li ZJ, et al. Genome-wide DNA methylation profiles and their relationships with mRNA and the microRNA transcriptome in bovine muscle tissue (Bos taurine). Scientific Reports-UK. 2014;**4**:1-17. DOI: ARTN 6546/10.1038/srep06546

[105] Liu X, Usman T, Wang Y, Wang Z, Xu X, Wu M, et al. Polymorphisms in epigenetic and meat quality related genes in fourteen cattle breeds and association with beef quality and carcass traits. Asian-Australasian Journal of Animal Sciences. 2015;**28**(4):467-475. DOI: 10.5713/ajas.13.0837

[106] Romao JM, Jin W, He M, McAllister T, Guan le L. MicroRNAs in bovine adipogenesis: Genomic context, expression and function. BMC Genomics. 2014;**15**:137. DOI: 10.1186/1471-2164-15-137

[107] Jin W, Dodson MV, Moore SS, Basarab JA, Guan LL. Characterization of microRNA expression in bovine adipose tissues: A potential regulatory mechanism of subcutaneous adipose tissue development. BMC Molecular Biology. 2010;**11**:29. DOI: 10.1186/1471-2199-11-29

[108] Muroya S, Shibata M, Hayashi M, Oe M, Ojima K. Differences in circulating microRNAs between grazing and grain-fed wagyu cattle are associated with altered expression of intramuscular microRNA, the potential target PTEN, and lipogenic genes. PLoS One. 2016;**11**(9):e0162496. DOI: 10.1371/journal.pone.0162496

[109] Sun J, Sonstegard TS, Li C, Huang Y, Li Z, Lan X, et al. Altered microRNA expression in bovine skeletal muscle with age. Animal Genetics. 2015;**46**(3):227-238. DOI: 10.1111/age.12272

[110] Dayton WR, White ME. MEAT SCIENCE AND MUSCLE BIOLOGY SYMPOSIUM—Role of satellite cells in anabolic steroid-induced muscle growth in feedlot steers. Journal of Animal Science. 2014;**92**(1):30-38. DOI: 10.2527/jas.2013-7077

[111] Sadkowski T, Ciecierska A, Oprzadek J, Balcerek E. Breed-dependent microRNA expression in the primary culture of skeletal muscle cells subjected to myogenic differentiation. BMC Genomics. 2018;**19**(1):109. DOI: 10.1186/s12864-018-4492-5

[112] Dai Y, Zhang WR, Wang YM, Liu XF, Li X, Ding XB, et al. MicroRNA-128

regulates the proliferation and differentiation of bovine skeletal muscle satellite cells by repressing Sp1. Molecular and Cellular Biochemistry. 2016;**414**(1-2):37-46. DOI: 10.1007/s11010-016-2656-7

[113] Dai Y, Wang YM, Zhang WR, Liu XF, Li X, Ding XB, et al. The role of microRNA-1 and microRNA-206 in the proliferation and differentiation of bovine skeletal muscle satellite cells. In Vitro Cellular & Developmental Biology. Animal. 2016;**52**(1):27-34. DOI: 10.1007/s11626-015-9953-4

[114] Horikawa A, Ogasawara H, Okada K, Kobayashi M, Muroya S, Hojito M. Grazing-induced changes in muscle microRNA-206 and -208b expression in association with myogenic gene expression in cattle. Animal Science Journal. 2015;**86**(11):952-960. DOI: 10.1111/asj.12381

[115] McCarthy JJ. MicroRNA-206: The skeletal muscle-specific myomiR. Biochimica et Biophysica Acta. 2008;**1779**(11):682-691. DOI: 10.1016/j.bbagrm.2008.03.001

[116] Yan X, Huang Y, Zhao JX, Rogers CJ, Zhu MJ, Ford SP, et al. Maternal obesity downregulates microRNA let-7g expression, a possible mechanism for enhanced adipogenesis during ovine fetal skeletal muscle development. International Journal of Obesity. 2013;**37**(4):568-575. DOI: 10.1038/ijo.2012.69

[117] Yan X, Zhu M-J, Du M. MicroRNA has-let-7g inhibits proliferation, adipogenic differentiation and inflammation in C3H10T½ cells. The FASEB Journal. 2011;**25**(Supplement lb):493

[118] Moisa SJ, Shike DW, Shoup L, Loor JJ. Maternal plane of nutrition during late-gestation and weaning age alter steer calf longissimus muscle adipogenic microRNA and target gene expression. Lipids. 2016;**51**(1):123-138. DOI: 10.1007/s11745-015-4092-y

[119] Wattiaux MA. Dairy sector across the world: National trends and opportunities for sustainable growth. In: Beede DK, editor. Large Dairy Herd Management. 3rd ed. Champaign, IL: American Dairy Science Association; 2017. pp. 3-19

[120] Wang M, Xu B, Wang H, Bu D, Wang J, Loor JJ. Effects of arginine concentration on the in vitro expression of casein and mTOR pathway related genes in mammary epithelial cells from dairy cattle. PLoS One. 2014;**9**(5):e95985. DOI: 10.1371/journal.pone.0095985

[121] Underwood KR, Means WJ, Zhu MJ, Ford SP, Hess BW, Du M. AMP-activated protein kinase is negatively associated with intramuscular fat content in longissimus dorsi muscle of beef cattle. Meat Science. 2008;**79**(2):394-402. DOI: 10.1016/j.meatsci.2007.10.025

[122] Zhou H, Cheng L, Azimu W, Hodge S, Edwards GR, Hickford JG. Variation in the bovine FABP4 gene affects milk yield and milk protein content in dairy cows. Scientific Reports. 2015;**5**:10023. DOI: 10.1038/srep10023

[123] Goszczynski DE, Papaleo-Mazzucco J, Ripoli MV, Villarreal EL, Rogberg-Munoz A, Mezzadra CA, et al. Genetic variation in FABP4 and evaluation of its effects on beef cattle fat content. Animal Biotechnology. 2017;**28**(3):211-219. DOI: 10.1080/10495398.2016.1262868

[124] Benjamin AL, Green BB, Crooker BA, McKay SD, Kerr DE. Differential responsiveness of Holstein and Angus dermal fibroblasts to LPS

challenge occurs without major differences in the methylome. BMC Genomics. 2016;**17**:258. DOI: 10.1186/s12864-016-2565-x

[125] Drackley JK, Donkin SS, Reynolds CK. Major advances in fundamental dairy cattle nutrition. Journal of Dairy Science. 2006;**89**(4):1324-1336. DOI: 10.3168/jds.S0022-0302(06)72200-7

[126] Gonzalez-Calvo L, Joy M, Alberti C, Ripoll G, Molino F, Serrano M, et al. Effect of finishing period length with alpha-tocopherol supplementation on the expression of vitamin E-related genes in the muscle and subcutaneous fat of light lambs. Gene. 2014;**552**(2):225-233. DOI: 10.1016/j.gene.2014.09.037

[127] Coyne GS, Kenny DA, Childs S, Sreenan JM, Waters SM. Dietary n-3 polyunsaturated fatty acids alter the expression of genes involved in prostaglandin biosynthesis in the bovine uterus. Theriogenology. 2008;**70**(5):772-782. DOI: 10.1016/j.theriogenology.2008.05.048

[128] Lohakare J, Osorio JS, Bionaz M. Peroxisome proliferator-activated receptor beta/delta does not regulate glucose uptake and lactose synthesis in bovine mammary epithelial cells cultivated in vitro. The Journal of Dairy Research. 2018:1-8. DOI: 10.1017/S0022029918000365

[129] Shi HB, Zhang CH, Xu ZA, Lou GG, Liu JX, Luo J, et al. Peroxisome proliferator-activated receptor delta regulates lipid droplet formation and transport in goat mammary epithelial cells. Journal of Dairy Science. 2018;**101**(3):2641-2649. DOI: 10.3168/jds.2017-13543

[130] Khan M, Couturier A, Kubens JF, Most E, Mooren FC, Kruger K, et al. Niacin supplementation induces type II to type I muscle fiber transition in skeletal muscle of sheep. Acta Veterinaria Scandinavica. 2013;**55**:85. DOI: 10.1186/1751-0147-55-85

[131] Gruffat D, Cherfaoui M, Bonnet M, Thomas A, Bauchart D, Durand D. Breed and dietary linseed affect gene expression of enzymes and transcription factors involved in n-3 long chain polyunsaturated fatty acids synthesis in longissimus thoracis muscle of bulls. Journal of Animal Science. 2013;**91**(7):3059-3069. DOI: 10.2527/jas.2012-6112

[132] Goszczynski DE, Mazzucco JP, Ripoli MV, Villarreal EL, Rogberg-Munoz A, Mezzadra CA, et al. Genetic characterisation of PPARG, CEBPA and RXRA, and their influence on meat quality traits in cattle. Journal of Animal Science and Technology. 2016;**58**:14. DOI: 10.1186/s40781-016-0095-3

[133] Qu Y, Elsasser TH, Kahl S, Garcia M, Scholte CM, Connor EE, Schroeder GF, Moyes KM. The effects of feeding mixed tocopherol oil on whole-blood respiratory burst and neutrophil immunometabolic-related gene expression in lactating dairy cows. Journal of Dairy Science. 2018;**101**(5):4332-4342. DOI: 10.3168/jds.2017-13902

[134] Akbar H, Schmitt E, Ballou MA, Correa MN, Depeters EJ, Loor JJ. Dietary lipid during late-pregnancy and early-lactation to manipulate metabolic and inflammatory gene network expression in dairy cattle liver with a focus on PPARs. Gene Regulation and Systems Biology. 2013;**7**:103-123. DOI: 10.4137/GRSB.S12005

[135] Carreno D, Hervas G, Toral PG, Castro-Carrera T, Frutos P. Fish oil-induced milk fat depression and associated downregulation of mammary lipogenic genes in dairy ewes. Journal of Dairy Science. 2016. DOI: 10.3168/jds.2016-11019

[136] Wang X, Lan X, Radunz AE, Khatib H. Maternal nutrition during

pregnancy is associated with differential expression of imprinted genes and DNA methyltranfereases in muscle of beef cattle offspring. Journal of Animal Science. 2015;**93**(1):35-40. DOI: 10.2527/jas.2014-8148

[137] Jacometo CB, Zhou Z, Luchini D, Trevisi E, Correa MN, Loor JJ. Maternal rumen-protected methionine supplementation and its effect on blood and liver biomarkers of energy metabolism, inflammation, and oxidative stress in neonatal Holstein calves. Journal of Dairy Science. 2016;**99**(8):6753-6763. DOI: 10.3168/jds.2016-11018

[138] Liu X, Guo XY, Xu XZ, Wu M, Zhang X, Li Q, et al. Novel single nucleotide polymorphisms of the bovine methyltransferase 3b gene and their association with meat quality traits in beef cattle. Genetics and Molecular Research. 2012;**11**(3):2569-2577. DOI: 10.4238/2012.June.29.1

[139] Zhao Q, Kang Y, Wang HY, Guan WJ, Li XC, Jiang L, et al. Expression profiling and functional characterization of miR-192 throughout sheep skeletal muscle development. Scientific Reports. 2016;**6**:30281. DOI: 10.1038/srep30281

Section 4

Basic Studies in Transcription Toward a New Era of Molecular Biology

Chapter 8

Transcriptional Initiation in Ribosomal Protein Genes in the Fission Yeast *Schizosaccharomyces pombe*

Diego A. Rojas, Sandra Moreira-Ramos, Fabiola Urbina and Edio Maldonado

Abstract

Transcription of class II genes in eukaryotic organisms is carried out by the multi-subunit enzyme RNA polymerase II (RNA pol II) and includes the general transcription factors and the mediator. The region inside the promoters, which recruits and specifies the transcriptional machinery, is called "core promoter" and contains sub-regions called "core promoter elements," which are necessary for transcription initiation, where the most studied and classic element is the TATA-box. Ribosome protein gene (RPG) promoters do not possess a TATA-box (TATA-less promoters), and those, in particular, in the fission yeast *Schizosaccharomyces pombe* have a TATA-box analog called the HomolD-box. The transcription of RPG promoters is dependent on the RNA pol II transcription system and the HomolD-box is recognized by the transcription factor Rrn7. In this chapter, the authors will describe the general mechanisms associated to the transcription of TATA-less promoters in eukaryotic organisms and how the transcription initiation is carried out in the RPG promoters from those organisms, particularly in *Schizosaccharomyces pombe*. Finally, the authors will analyze the role of the HomolD-box and the transcription factor Rrn7 in the coordination of transcription initiation from RPG promoters and other ribosome-related genes and the presence of transcriptional modules in their promoters, which could be coordinated and regulated by a discrete number of transcription factors.

Keywords: transcription, ribosomal protein gene (RPG), RNA polymerase, TATA-less promoter, *Schizosaccharomyces pombe*

1. Introduction

Protein synthesis in eukaryotic organisms includes several steps and requires many regulatory events [1, 2]. One of these critical steps is ribosome biogenesis, which includes ribosomal protein gene (RPG) transcription and rRNA synthesis. As in many other central events in the cell, ribosome biosynthesis must be a regulated and coordinated process. A typical coordinated regulation of gene and protein expression is the presence of common DNA elements in the promoters of related genes, which are co-regulated by a discrete number of transcription factors. Those genes under the control of a common DNA element form a transcriptional

module (regulon). In this chapter, the authors will describe the state of the art of several topics associated to the transcription initiation from TATA-less promoters in eukaryotic organisms, such as the transcriptional regulation of RPGs in metazoan cells and the description of a novel mechanism of regulation present in the RPG of the fission yeast *Schizosaccharomyces pombe*.

2. General features of transcription initiation in the eukaryotic organisms and TATA-less CPEs

Transcription in eukaryotic organisms is carried out by RNA polymerases (RNA pols), which are enzymatic complexes composed by at least 12 subunits. In general, eukaryotic genes are classified as class I, II, and III, where class I genes codify rRNAs; class II codify pre-mRNAs; and class III codify 5S rRNA, tRNAs, and snRNAs, respectively. Transcription of each class of genes is carried out by a different RNA pol. Class I genes are transcribed by RNA pol I, class II genes are transcribed by RNA pol II, and class III genes are transcribed by RNA pol III, respectively. This specific transcription is based on the recognition of specific DNA sequences in the promoters of each class of genes by different transcription factors (TFs) that are able to recruit each specific RNA pol. These sequences are named "core promoter elements" (CPEs) and are located inside the region of the promoter named "core promoter" (CP) that is able to direct the formation of a pre-initiation complex (PIC) and initiate specific transcription of the gene. The CPEs are recognized by TFs specific to each RNA pol, which are called "general transcription factors" (GTFs). In summary, each RNA pol has a set of specific GTFs and these protein factors are able to recognize the CPEs associated to each class' gene promoters.

RNA pol II has been widely studied due to the enzyme that transcribes protein-coding genes. One of the first CPEs described in the promoters of class II genes was the so-called TATA-box [3–5]. This CPE is distributed in the promoters of most eukaryotic organisms and is located 25–40 bp upstream from the transcription initiation site. The formation of a PIC on the promoters containing a TATA-box has been extensively studied and characterized [6–8]. The formation of a PIC on the TATA-box starts with the recognition and binding of the transcription factor TATA-binding protein (TBP) to the TATA-box which in turn recruits the other GTFs and RNA pol II to form the PIC, which is able to initiate transcription upon the addition of the ribonucleotides [9, 10] (**Figure 1**). As it can be seen from the model, RNA pol II is integrated into the PIC in association with TFIIF when the promoter-TBP-TFIIB complex is formed. On the other hand, a fraction of RNA pol II can be purified from cell extracts in association with TFIIF and the mediator, and since those complexes are preformed inside the nucleus, a fast recruitment of the PIC to the promoter could be produced [11–13]. The multi-subunit complex named mediator (a general transcriptional coactivator) is also necessary for the transcription *in vivo*, in crude cell extracts, of class II genes [14]. In addition, recently, it has been demonstrated that another protein complex is recruited *in vivo* at most of the class II gene promotes in *S. cerevisiae*, where it plays a fundamental role in transcription. This multi-protein complex is named SAGA and is composed of several subunits including Gcn5, which have histone acetyltransferase (HAT) activity Spt gene products and TBP-associated factors (TAFs) that are shared with the complex TFIID [15].

In metazoan cells, the transcription factor TBP is tightly associated to TAFs and the TBP-TAF complex is named TFIID [16]. The role of TAFs seems to be the recognition of certain CPEs such as the Inr, motif ten element (MTE), and downstream promoter element (DPE) (see below). However, in yeast, this complex seems to be unstable, since it is possible to purify TBP free of TAFs from yeast cell extracts. Although TAFs are required for *in vivo* transcription of *S. cerevisiae* genes, the exact

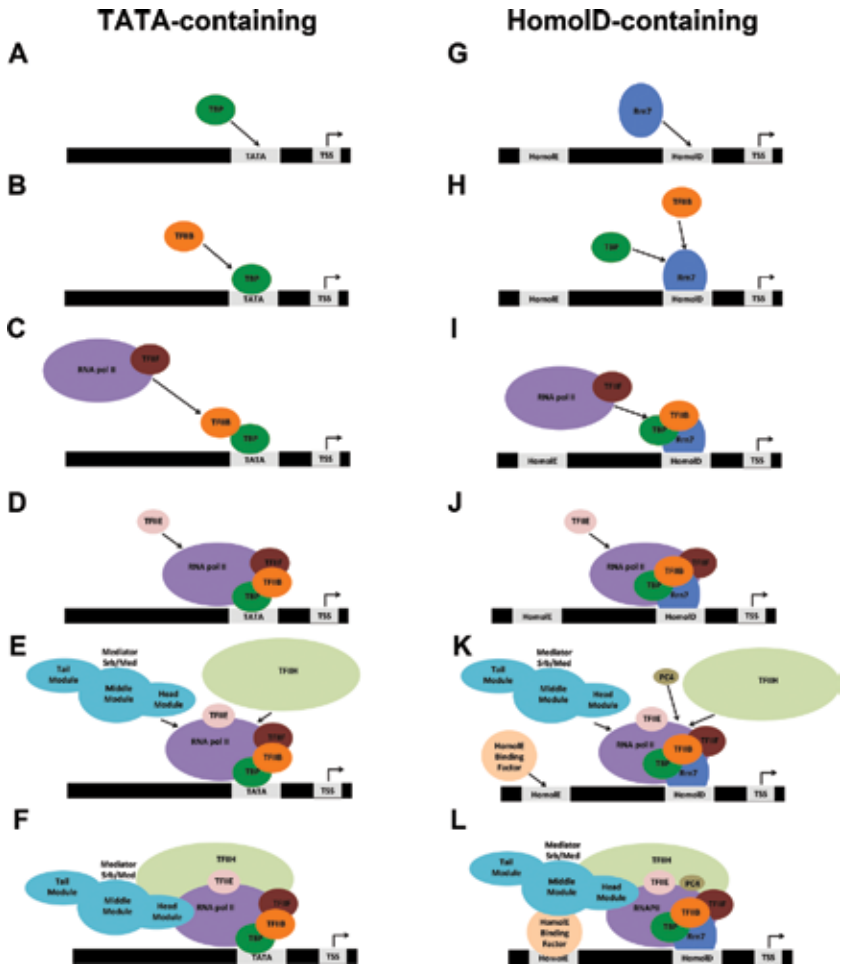

Figure 1.
PIC formation on TATA-containing and HomolD-containing promoters. *Classical PIC formation on a TATA-containing promoter is outlined in A–F. First, TBP binds to the TATA-box and then TFIIB is recruited to the promoter-TBP complex. This allows the RNA pol II-TFIIF complex to be incorporated into the promoter-TBP-TFIIB complex (C). Once that RNA pol II-TFIIF is loaded onto the complex, the transcription factor TFIIE is incorporated (D) followed by the binding of TFIIH (E). The mediator complex might be incorporated into the complex after the binding of RNA pol II-TFIIF and TFIIE (D). At step F, the complete PIC is formed and it is competent for transcription initiation. A competent PIC is formed on HomolD-box containing promoters, such as RPG promoters, to initiate RNA pol II-dependent transcription (G–L). The first step is the binding of the transcription factor Rrn7 to the HomolD-box sequence (G). Then, transcription factors TBP and TFIIB bind to Rrn7 (H). This DNA-protein complex is recognized by RNA pol II-TFIIF (I) and TFIIE (J). This complex is competent to initiate HomolD-box-dependent transcription. However, coactivators such as the mediator, PC4, and the HomolE-binding factor would be necessary to modulate transcription initiation (K and L). TSS: transcription start site. Note that steps C to F for PIC formation on TATA-box-containing promoters might be common with steps I to L on PIC formation on HomolD-box-containing promoters.*

mechanism of their function has not been identified yet. Interestingly, using an *in vitro* approach using TFIID-depleted yeast cell extracts, it was found that transcription from both TATA-containing and TATA-less promoters is dependent on TFIID, but isolated recombinant TBP can only rescue the transcription of TATA-containing promoters, indicating that additional interactions are necessary to efficiently transcribe TATA-less promoters [17].

However, our vision of transcription initiation on TATA-box-containing promoters cannot explain the mechanisms of transcription initiation on all the class II genes, because the analysis of several other class II gene promoter sequences showed that

in most of them the TATA-box is absent. Different studies have determined that only 10–15% of mammal core promoters contain a TATA-box element [18–20]. Those promoters that do not contain a TATA-box were named TATA-less promoters and they have also been studied and different CPEs have been characterized.

After the identification of TATA-box sequence, other conserved promoter elements were identified. One of them is the initiator element (Inr), identified as a conserved DNA element in the region near to the transcription start site [21]. This element can not only direct transcription initiation by itself if other CPEs are not present, but also act synergistically in the presence of a TATA-box [22]. The proteins $TAF_{II}150$ and $TAF_{II}250$ have been identified as the transcription factors that are able to recognize the Inr and allow the formation of the PIC in Inr-containing promoters [23, 24]. However, other Inr-containing promoters might be able to direct transcription initiation in a TAFs-independent manner. In those promoters, a few proteins have been identified as Inr-binding factors, such as TF_{II}-I and YY1 [25, 26]. Also, in other reports, transcription initiation from the human DNA beta polymerase promoter and from the human dihydrofolate reductase (DHFR) promoter, both TATA-less and Inr-containing promoters, has been achieved using solely TBP, IIB, IIE, IIF, IIH, and RNA pol II [26, 27]. This suggests that in some TATA-less promoters, the formation of a functional PIC might follow a common pathway with those TATA-containing promoters.

Another CPE that has been described in TATA-less promoters is the downstream promoter element (DPE), identified first in *Drosophila melanogaster* [28]. This element is widely distributed in metazoan organisms and is located 28–32 bp downstream from the transcription start site and can be contained in the context of a TATA-box and/or an Inr. Studies in *Drosophila* have shown that proteins $TAF_{II}40$ and $TAF_{II}60$ might bind to the DPE to improve transcription initiation [29, 30]. Similar elements have not been found in yeast yet.

Several other CPEs have been identified in TATA-less promoters but their contribution to transcription initiation is still poorly understood. Such is the case of motif ten element (MTE) [31]; TFIIB recognition element (BRE) [32]; X core promoter element 1 and 2 (XCPE1 and 2) [33, 34], both of which are able to direct transcription initiation; and the poly-pyrimidine initiator motif (TCT) motif [35]. The TCT motif element will be described in another section of this chapter.

However, using the information from the sequencing of the genomes of other organisms and the new bioinformatics technologies, it is expected that novel conserved CPEs will be identified and characterized and the transcription initiation mechanisms of TATA-less promoters will be revealed. Such is the case of the ribosomal protein genes (RPGs) in the fission yeast *Schizosaccharomyces pombe*, whose promoters do not contain a TATA-box; instead they possess a conserved sequence, acting as a TATA-analog to direct transcription initiation in those genes. In the next section, the RPG promoter of the fission yeast will be described and the transcription initiation mechanism will be discussed.

3. Characterization of ribosomal protein gene (RPG) promoters of *Schizosaccharomyces pombe* and their transcription initiation mechanism

3.1 The *Schizosaccharomyces pombe* RPG transcriptional module: the HomolD-box

The characterization of the promoter sequences of 14 RPGs from the fission yeast *Schizosaccharomyces pombe* showed discrete conserved modules, which were

named Homol A, B, C, D, and E (**Table 1**) [36–38]. These homology regions were completely different from those described in promoters of genes from other yeasts and mammals, such as TATA-box, Inr, or DPEs. The function of each Homol element was studied using a promoter-deletion mutant approach [37]. This work showed that the role of Homol A, B, C, and E is associated to the regulation of transcription initiation, and that they might have a upstream activation sequence (UAS)-like function. Only the HomolD sequence was able to function as an element that could direct transcription initiation in the same way as the TATA-box [36]. The conserved sequence of the HomolD-box is the octamer CAGTCACA/G; however, in several sequences, this element is found in the inverted form as TGTGACTG. The HomolD-box is located 39–52 bp upstream of the transcription start site in the same position as the TATA-box in the fission yeast promoters. In an *in vivo* approach, using reporter-gene assays in *S. pombe* cells, it was shown that the HomolD-box is necessary to direct and initiate transcription from the RPG and was postulated to act as a TATA-box analog; in the same work, using an electrophoretic mobility shift assay (EMSA), a novel protein complex that binds to the HomolD-box was identified [36]. In other studies using an *in vitro* approach, it was shown that point mutations in the HomolD-box sequence abolish completely the ability of this element to direct transcription initiation from the RPG [39].

Currently, we know that the genome of *Schizosaccharomyces pombe* contains 141 RPGs encoding the full set of 79 ribosomal proteins. Interestingly, the analysis of the promoter sequences showed that 140 RPGs contained a highly conserved HomolD-box in the region 49–104 bp upstream of the ATG start codon [40]. Additionally, other 59 non-RPGs also showed the presence of the HomolD-box in their promoters. In addition, using promoter databases, it was possible to find HomolD-box sequences in several promoters from other eukaryotic organisms, such as humans and plants, indicating the broad distribution of this novel CPE. Moreover, a functional HomolD-box was found in the human ATPV1H gene where RECQL/DDB1 complex binds to this sequence and is required for *in vitro* transcription [41].

Interestingly, HomolD-boxes in RPG promoters are broadly distributed in the *Ascomycota* fungus phylum [42]. However, in those organisms closely related to the yeast *Saccharomyces cerevisiae,* other CPEs, in the same position as the HomolD-box, are present in RPG promoters. These elements are named Rap1 and bind the transcription factor Rap1p [43]. It seems likely that Rap1 replaced the HomolD-box of *Schizosaccharomyces pombe* in *Saccharomyces cerevisiae* during evolution. Moreover, several other yeast species share both HomolD-box and Rap1 promoter elements [42]. Taking all those observations together, we suggest that RPGs from *S. pombe*, *S. cerevisiae*, *Drosophila,* and mammals form a transcriptional module that is under the control of the HomolD-box, Rap1-box, and TCT motif (*Drosophila* promoter element), respectively.

Homol	Consensus	Binding TF	Function	Reference
HomolA	TCAGTAACGAA	Unknown	UAS-like	[48]
HomolB	AAAAGCTATG	Unknown	UAS-like	[48]
HomolC	AAGAGTAAAATCT	Unknown	UAS-like	[48]
HomolD	CAGTCACA/G	Rrn7 (*S. pombe*) RECQL/DDB1 (Human, *S. pombe*)	Transcription initiation and regulation of RPG expression	[36, 39, 48]
HomolE	AGGGTAGGGT	Unknown	UAS-like	[37, 48]

Table 1.
Homol sequences identified in RPG promoters in S. pombe.

3.2 The role of Rrn7 and CK2 in RPG transcription initiation in *Schizosaccharomyces pombe*

The HomolD-box present in the RPG promoters of the fission yeast is the target of a DNA-binding protein with biochemical features different from TBP. The identification of the HomolD-box-binding protein was achieved using DNA affinity chromatography with double-stranded tandem HomolD-boxes covalently attached to a resin. Proteins bound to the resin were eluted and analyzed by mass spectrometry. The result was that the transcription factor Rrn7 was identified in the protein DNA-bound fraction [39]. This factor is a member of the RNA pol I transcriptional machinery and its function is to transcribe rDNA in the nucleolus. In the rDNA promoter, this factor is able to bind to a conserved box, which is similar to a HomolD-box. Rrn7 showed a specific HomolD-box-binding activity and it is required for the specific transcription of RPGs containing a HomolD-box [39]. Moreover, the GTFs and RNA pol II were required for accurate transcription initiation of a HomolD-box-containing promoter.

Rrn7 is part of the Zn-ribbon protein family related to TFIIB, including the mammalian ortholog TAF1B [44]. It possesses a Zn-ribbon domain in the N-terminal region and two cyclin-like domains in the carboxy-terminal region, displaying domain conservation with the TFIIB family members [44]. Recently, it has been demonstrated that *Schizosaccharomyces pombe* Rrn7 is able to interact with casein kinase 2 (CK2) both *in vitro* and *in vivo*, leading to a functional phosphorylation of threonine 67 in the N-terminal domain. This modification modulates negatively the transcriptional activity of Rrn7, affecting HomolD-directed transcription and DNA-binding activity [45]. Studies in *S. pombe* cell cultures using the specific CK2 inhibitor 4,5,6,7-tetrabromobenzotriazole (TBB) have shown the potentiation of RPG expression during CK2 inhibition. Moreover, using chromatin immunoprecipitation assays, it has been found that CK2 is associated with RPG promoters, suggesting that this kinase has a role in the modulation of ribosomal protein abundance [45].

3.3 Preinitiation complex (PIC) formation on the RPG promoters in *Schizosaccharomyces pombe*

As stated before, RPGs that contain a HomolD-box are transcribed by the RNA pol II transcription apparatus [39]. The formation of the PIC on a HomolD-box-containing promoter was recently described [45, 46]. The first step in the formation of a PIC on these promoters is the binding of Rrn7 to the HomolD-box. As mentioned previously, this step in the PIC establishment might be regulated by phosphorylation of Rrn7 via CK2 protein kinase [45]. Upon the binding of Rrn7 to the HomolD-box, the general transcription factors TBP and TFIIB are able to recognize this DNA-protein complex [46]. After the binding of TBP/TFIIB to the complex, the RNA pol II/TFIIF complex is recruited, which in turn allows the TFIIE factor to be incorporated into the complex [46]. Finally, the mediator and the coactivator PC4 may be incorporated into the PIC and might modulate basal transcription through a putative HomolE-binding factor in those promoters that contain this DNA element. All the steps describing the pathway of complex formation are summarized in **Figure 1**.

4. Regulation of RPG expression in *Schizosaccharomyces pombe*

The expression of genes containing the HomolD-box in their promoters is almost unknown. However, data from analysis of the RPG expression profiles during

several biological processes in *S. pombe*, for example, the switch from vegetative to meiotic growth and growth under stress conditions, have revealed a tightly coordinated expression for all 141 RPGs. For example, during the switch from vegetative to meiotic growth, transcription of RPG is down-regulated, but then, within a short time, strong reactivation of RPG expression is observed at the beginning of meiosis. The same co-regulation profile is observed in 32 of the 59 non-RPGs that contain a HomolD-box in their core promoter [40]. Many, but not all, of these non-RPGs encode components whose homologs in other organisms are involved in protein biosynthesis and signal transduction [40].

Several promoters of *Schizosaccharomyces pombe* RPGs have been isolated and characterized [47–50], showing that individual ribosomal proteins are encoded by two or three related genes whose promoters contain a HomolD-box. Interestingly, in each gene family, at least one promoter possesses a tandem repeat ACCCTACCCT or the inverted form (AGGGTAGGGT) upstream of the HomolD-box. This sequence corresponds to the HomolE-box, which is considered a proximal UAS-like sequence for HomolD-box-containing promoters, since the presence of this element strongly increases *in vivo* transcription directed by the HomolD-box. Both promoter elements HomolD- and HomolE-boxes must be in the same orientation to be functional. The distance between the boxes is critical in transcription modulation of RPGs, and it has been described that the smaller the distance between HomolD and HomolE, the higher the transcription activity. This distance ranges from 0 to 32 nucleotides.

Now that the complete genome of *Schizosaccharomyces pombe* is available and searchable, it is known that of the 141 RPG promoters, 140 promoters contain a HomolD-box and 62 contain a HomolE-box upstream of the HomolD-box. In contrast, only 5 of the 59 non-RPG promoters containing a HomolD-box contain a HomolE-box [40].

The activity of the HomolE-box must be related to a transcriptional activator in the same manner as the IFH1 element is recognized by a TF (Ifh1p) in *Saccharomyces cerevisiae* RPGs [51, 52]. This element has been identified upstream of the Rap1 sequence in RPG promoters. It is likely that a member of the same family of Ifh1p recognizes the HomolE-box in *S. pombe*. However, the gene encoding this protein has not been identified yet.

The modular architecture of the *Schizosaccharomyces pombe* RPG promoters where the HomolE-box is always found upstream of HomolD indicates that some of the promoters are under the control of the HomolE-box. This fact suggests that there must be a mechanism that regulates RPG transcription through the HomolE-box under specific growth conditions.

Further investigations must be performed to understand how RPG expression is regulated and which are the mechanisms involved in the coordination between HomolD- and HomolE-boxes during RPG transcription. Despite the fact that there are several factors and mechanisms studied in RNA pol II-directed transcription, most of the promoters studied possess a TATA-box, whereas RPG promoters are TATA-less. Moreover, transcriptional initiation and activation from TATA-less promoters are poorly understood both in metazoan and yeast cells. Thus, the RPG promoters and the arrangement of HomolE-HomolD could provide a model to study transcription in TATA-less promoters using a promoter element such as HomolD that is analogous to the TATA-box.

5. Coordinated regulation of the expression of ribosomal components in *Schizosaccharomyces pombe*

Ribosome biogenesis is one of the most complicated processes in eukaryotic cells, requiring coordinated expression of all ribosome components, which are

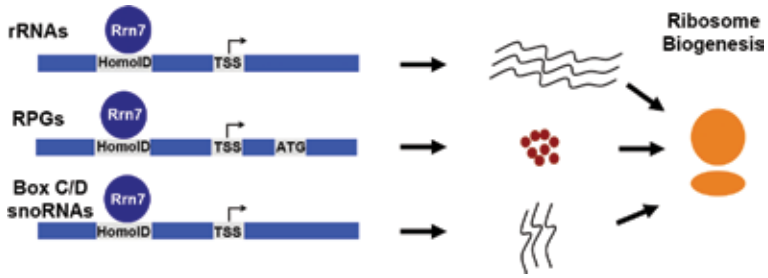

Figure 2.
Role of HomolD-box-containing promoters in ribosome biogenesis. *It is believed that the coordinated expression of genes that encode ribosome components (rRNA, RPG, box C/D snoRNA) is due to a common DNA element (HomolD-box) that is able to bind a protein factor (Rrn7). The presence of the HomolD-box in the promoters of several genes encoding ribosome components indicates a common regulation. Until now, experiments have demonstrated that Rrn7 binds to rRNA and RPG promoters in Schizosaccharomyces pombe. However, binding of Rrn7 to box C/D snoRNA promoters has not been demonstrated yet, although those promoters contain a HomolD-box, which is critical for in vivo transcription.*

essential for accurate translation activity. The coordinated regulation and expression of the RPG with other ribosomal components is still poorly understood. However, in the fission yeast *Schizosaccharomyces pombe*, it is known that rRNAs, RPGs, and box C/D snoRNAs contain in their promoters a HomolD-box [39, 53], which might be able to control the expression of those genes. Moreover, the Rrn7 transcription factor, which is the HomolD-box-binding protein in RPG, was found to be responsible for the control of the gene expression of box C/D snoRNAs and RPGs *in vivo* in *Schizosaccharomyces pombe* cells [53]. Interestingly, the yeast orthologs of the human RECQL/DDB1 complex may recognize the HomolD-box and down-regulate RPG expression [53]. Taking all these results together, we propose a model, in which the HomolD-box is bound by Rrn7 and co-regulates the transcription of RPG, box C/D snoRNA and rRNA genes in the fission yeast. This model is summarized in **Figure 2**.

Unlike *Schizosaccharomyces pombe* in the case of *Saccharomyces cerevisiae*, there is accumulated evidence that show a putative coordinated model to regulate biogenesis of ribosome components. In this model, CK2 protein kinase is part of protein complexes that regulate RPG expression and rRNA synthesis [54] and interact with the protein Fhl1p that is associated to Ifh1p, which binds to the IFH1 element near to the Rap1 sequence. CK2 and Ifh1p are part of the complex CURI associated to rRNA processing and RPG transcription [55]. Also, in *S. cerevisiae,* another protein has been identified and named protein HmoI, which is associated with the transcription regulation of RPGs and rRNAs [56].

In addition to the role of CK2 to modulate Rrn7 function in *Schizosaccharomyces pombe* during HomolD-box directed transcription, there might be another points of regulation related to protein complexes, such as those described in the yeast *Saccharomyces cerevisiae*.

6. The TCT-motif module in metazoan RPG

The analyses of insect and mammalian RPG promoters have shown the presence of a common core promoter element that is part of the poly-pyrimidine initiator (TCT)-motif family, which is a novel core promoter element necessary to initiate transcription in those genes [35, 57]. In these promoters, the transcription start site (TSS) involves the TCT motif and is positioned around −2 to +6 relative to TSS, competing with exactly the same position as the Inr. However, the features of

the TCT-containing promoters are dissimilar to those Inr-containing promoters. The function of a TCT motif cannot be replaced by an Inr, and the TFIID complex cannot bind to the TCT motif [57]. Recently, studies in *Drosophila* RPG promoters, which contain a TCT motif, have shown the dependence on a TBP-related factor 2 (TRF2) but not TBP. Using a TRF2-depleted *Drosophila* whole cell extract, it was shown that human TRF2 [58, 59] and *Drosophila* TRF2 [60] were able to support TCT-dependent transcription. The TATA-binding protein TBP was unable to support TCT-dependent transcription. However, whether or not the TBP factor is required for TCT-dependent transcription remains to be determined. The proteins able to recognize this element are still unknown, because TRF2 is unable to bind directly to the TCT motif. It is possible that TRF2 interacts with other members of the RNA pol II basal transcription machinery and forms a PIC associated with the TCT motif. In addition, TRF2 is able to bind to the vicinity of the TSS of other genes, since it can be crosslinked and immunoprecipitated from that region, but whether or not this factor binds directly to the Inr motif is still unknown [61].

7. Conclusions

The fission yeast *Schizosaccharomyces pombe* provides an excellent biological model to study the coordinated expression of ribosome components. The finding that rDNAs, RPGs, and box C/D snoRNAs genes contain a HomolD-box, which is most likely bound by Rrn7, provides the starting point to investigate this issue. The most important questions to answer are: (i) to determine whether or not box C/D snoRNA genes are transcribed by the same transcription apparatus that transcribes RPG; (ii) to identify the signal that activates transcription of HomolD-box containing genes, and (iii) to identify the HomolE-binding protein. The resolution of all these issues would contribute to understand the regulation of RPG transcription in the fission yeast and most likely could be extrapolated to metazoan organisms.

Acknowledgements

This work has been funded by the Corporación Nacional de Ciencia y Tecnología de Chile (CONICYT), Fondo de Desarrollo de Ciencia y Tecnología (FONDECYT), and FONDEF (ID16I10145). We are very grateful to Mrs. Teresa Donoso for her continued support to F. U. and E. M.

Conflict of interest

The authors do not have any conflict of interest.

Abbreviations

CP	core promoter
CPE	core promoter element
DPE	downstream promoter element
GTF	general transcription factor
Inr	initiator
PIC	pre-initiation complex
RNA pol	RNA polymerase

RPG	ribosomal protein genes
rRNA	ribosomal RNA
snoRNA	small nucleolar RNA
TAF	TBP-associated factor
TBP	TATA-binding protein
TCT	poly-pyrimidine initiator motif
TF	transcription factor
TRF	TBP-related factor
TSS	transcription start site

Author details

Diego A. Rojas[1], Sandra Moreira-Ramos[2], Fabiola Urbina[2] and Edio Maldonado[2]*

1 Programa de Microbiología y Micología, Instituto de Ciencias Biomedicas, Facultad de Medicina, Universidad de Chile, Independencia, Santiago, Chile

2 Programa de Biología Celular y Molecular, Instituto de Ciencias Biomedicas, Facultad de Medicina, Universidad de Chile, Independencia, Santiago, Chile

*Address all correspondence to: emaldona@med.uchile.cl

References

[1] Kandiah E, Trowitzsch S, Gupta K, Haffke M, Berger I. More pieces to the puzzle: Recent structural insights into class II transcription initiation. Current Opinion in Structural Biology. 2014;**24**:91-97. DOI: 10.1016/j.sbi.2013.12.005

[2] Shandilya J, Roberts SG. The transcription cycle in eukaryotes: From productive initiation to RNA polymerase II recycling. Biochimica et Biophysica Acta. 2012;**1819**:391-400. DOI: 10.1016/j.bbagrm.2012.01.010

[3] Wasylyk B, Derbyshire R, Guy A, Molko D, Roget A, Teoule R, et al. Specific in vitro transcription of conalbumin gene is drastically decreased by single-point mutation in T-A-T-A box homology sequence. Proceedings of the National Academy of Sciences of the United States of America. 1980;**77**:7024-7028

[4] Sassone-Corsi P, Corden J, Kedinger C, Chambon P. Promotion of specific in vitro transcription by excised "TATA" box sequences inserted in a foreign nucleotide environment. Nucleic Acids Research. 1981;**9**:3941-3958

[5] Concino MF, Lee RF, Merryweather JP, Weinmann R. The adenovirus major late promoter TATA box and initiation site are both necessary for transcription in vitro. Nucleic Acids Research. 1984;**12**:7423-7433

[6] Orphanides G, Reinberg D. A unified theory of gene expression. Cell. 2002;**108**:439-451. DOI: 10.1016/S0092-8674(02)00655-4

[7] Thomas M, Chiang C. The general transcription machinery and general cofactors. Critical Reviews in Biochemistry and Molecular Biology. 2006;**41**:105-178. DOI: 10.1080/10409230600648736

[8] Sikorski T, Buratowski S. The basal initiation machinery: Beyond the general transcription factors. Current Opinion in Cell Biology. 2009;**21**:344-351. DOI: 10.1016/j.ceb.2009.03.006

[9] Maldonado E, Ha I, Cortes P, Weis L, Reinberg D. Factors involved in specific transcription by mammalian RNA polymerase II: Role of transcription factors IIA, IID, and IIB during formation of a transcription-competent complex. Molecular and Cellular Biology. 1990;**10**:6335-6347. DOI: 10.1128/MCB.10.12.6335

[10] Murakami K, Elmlund H, Kalisman N, Bushnell DA, Adams CM, Azubel M, et al. Architecture of an RNA polymerase II transcription pre-initiation complex. Science. 2013;**342**:1238724. DOI: 10.1126/science.1238724

[11] Lee TI, Young RA. Transcription of eukaryotic protein-coding genes. Annual Review of Genetics. 2000;**34**:77-137. DOI: 10.1146/annurev.genet.34.1.77

[12] Hahn S. Structure and mechanism of the RNA polymerase II transcription machinery. Nature Structural & Molecular Biology. 2004;**11**:394-403. DOI: 10.1038/nsmb763

[13] Esnault C. Mediator-dependent recruitment of TFIIH modules in preinitiation complex. Molecular Cell. 2008;**31**:337-346. DOI: 10.1016/j.molcel.2008.06.021

[14] Jeronimo C, Robert F. The mediator complex: At the nexus of RNA polymerase II transcription. Trends in Cell Biology. 2017;**27**:765-783. DOI: 10.1016/j.tcb.2017.07.001

[15] Baptista T, Grünberg S, Minoungou N, Koster MJE, Timmers HTM, Hahn S, et al. SAGA Is a general cofactor for RNA polymerase II transcription.

Molecular Cell. 2017;**68**:130-143. DOI: 10.1016/j.molcel.2017.08.016

[16] Burley SK, Roeder RG. Biochemistry and structural biology of transcription factor IID (TFIID). Annual Review of Biochemistry. 1996;**65**:769-799. DOI: 10.1146/annurev.bi.65.070196.004005

[17] Donczew R, Hahn S. Mechanistic differences in transcription initiation at TATA-less and TATA-containing promoters. Molecular and Cellular Biology. 2017;**38**:e00448-e00417. DOI: 10.1128/MCB.00448-17

[18] Kim T, Barrera L, Zheng M, Qu C, Singer M, Richmond TA, et al. A high-resolution map of active promoters in the human genome. Nature. 2005;**436**: 876-880. DOI: 10.1038/nature03877

[19] Carninci P, Sandelin A, Lenhard B, Katayama S, Shimokawa K, Ponjavic J, et al. Genome-wide analysis of mammalian promoter architecture and evolution. Nature Genetics. 2006;**38**:626-635. DOI: 10.1038/ng1789

[20] Cooper S, Trinklein N, Anton E, Nguyen L, Myers R. Comprehensive analysis of transcriptional promoter structure and function in 1% of the human genome. Genome Research. 2006;**16**:1-10. DOI: 10.1101/gr.4222606

[21] Smale ST, Baltimore D. The "initiator" as a transcription control element. Cell. 1989;**57**:103-113

[22] O'Shea-Greenfield A, Smale ST. Roles of TATA and initiator elements in determining the start site location and direction of RNA polymerase II transcription. The Journal of Biological Chemistry. 1992;**267**:1391-1402

[23] Verrijzer CP, Chen JL, Yokomori K, Tjian R. Binding of TAFs to core elements directs promoter selectivity by RNA polymerase II. Cell. 1995;**81**:1115-1125. DOI: 10.1016/S0092-8674(05)8 0016-9

[24] Kauffmann J, Ahrens K, Koop R, Smale ST, Müller R. CIF150, a human cofactor for transcription factor IID-dependent initiator function. Molecular and Cellular Biology. 1998;**18**:233-239. DOI: 10.1128/MCB.18.1.233

[25] Cheriyath V, Novina CD, Roy AL. TFII-I regulates Vβ promoter activity through an initiator element. Molecular and Cellular Biology. 1998;**18**:4444-4454. DOI: 10.1128/MCB.18.8.4444

[26] Weis L, Reinberg D. Accurate positioning of RNA polymerase II on a natural TATA-less promoter is independent of TATA-binding-protein-associated factors and initiator-binding proteins. Molecular and Cellular Biology. 1997;**17**:2973-2984. DOI: 10.1128/MCB.17.6.2973

[27] Aso T, Conaway JW, Conaway RC. Role of core promoter structure in assembly of the RNA polymerase II preinitiation complex. A common pathway for formation of preinitiation intermediates at many TATA and TATA-less promoters. The Journal of Biological Chemistry. 1994;**269**:26575-26583

[28] Burke TW, Kadonaga JT. *Drosophila* TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. Genes & Development. 1996;**10**: 711-724. DOI: 10.1101/gad.10.6.711

[29] Burke TW, Kadonaga JT. The downstream core promoter element, DPE, is conserved from *Drosophila* to humans and is recognized by TAFII60 of *Drosophila*. Genes & Development. 1997;**11**:3020-3031

[30] Kutach AK, Kadonaga JT. The downstream promoter element DPE appears to be as widely used as the TATA box in *Drosophila* core promoters. Molecular and Cellular Biology. 2000;**20**:4754-4764. DOI: 10.1128/ MCB.20.13.4754-4764. 2000

[31] Lim CY, Santoso B, Boulay T, Dong E, Ohler U, Kadonaga JT. The MTE, a new core promoter element for transcription by RNA polymerase II. Genes & Development. 2004;**18**: 1606-1617. DOI: 10.1101/gad.1193404

[32] Deng W, Roberts S. A core promoter element downstream of the TATA box that is recognized by TFIIB. Genes & Development. 2005;**19**:2418-2423. DOI: 10.1101/gad.342405

[33] Tokusumi Y, Ma Y, Song X, Jacobson R, Takada S. The new core promoter element XCPE1 (X Core Promoter Element 1) directs activator-, mediator-, and TATA-binding protein-dependent but TFIID-independent RNA polymerase II transcription from TATA-less promoters. Molecular and Cellular Biology. 2007;**27**:1844-1858. DOI: 10.1128/MCB.01363-06

[34] Anish R, Hossain M, Jacobson R, Takada S. Characterization of transcription from TATA-less promoters: Identification of a new core promoter element XCPE2 and analysis of factor requirements. PLoS One. 2009;**4**:e5103. DOI: 10.1371/journal.pone.0005103

[35] Parry TJ, Theisen JW, Hsu JY, Wang YL, Corcoran DL, Eustice M, et al. The TCT motif, a key component of an RNA polymerase II transcription system for the translational machinery. Genes & Development. 2010;**24**:2013-2018. DOI: 10.1101/gad.1951110

[36] Witt I, Straub N, Käufer NF, Gross T. The CAGTCACA box in the fission yeast *Schizosaccharomyces pombe* functions like a TATA element and binds a novel factor. The EMBO Journal. 1993;**12**:1201-1208

[37] Witt I, Kwart M, Gross T, Käufer NF. The tandem repeat AGGGTAGGGT is, in the fission yeast, a proximal activation sequence and activates basal transcription mediated by the sequence TGTGACTG. Nucleic Acids Research. 1995;**23**:4296-4302

[38] Gross T, Käufer NF. Cytoplasmic ribosomal protein genes of the fission yeast *Schizosaccharomyces pombe* display a unique promoter type: A suggestion for nomenclature of cytoplasmic ribosomal proteins in databases. Nucleic Acids Research. 1998;**26**:3319-3322

[39] Rojas DA, Moreira-Ramos S, Zock-Emmenthal S, Urbina F, Contreras-Levicoy J, Käufer NF, et al. Rrn7 protein, an RNA polymerase I transcription factor, is required for RNA polymerase II-dependent transcription directed by core promoters with a HomolD box sequence. The Journal of Biological Chemistry. 2011;**286**: 26480-26486. DOI: 10.1074/jbc.M111.224337

[40] Witt I, Kivinen K, Käufer NF. Core promoters in *S. pombe*: TATA and HomolD boxes. In: Egel R, editor. The Molecular Biology of *Schizosaccharomyces pombe*. Berlin/Heidelberg: Springer-Verlag; 2004. pp. 343-351. DOI: 10.1007/978-3-662-10360-9

[41] Contreras-Levicoy J, Moreira-Ramos S, Rojas DA, Urbina F, Maldonado E. Transcription directed by human core promoters with a HomolD box sequence requires DDB1, RECQL and RNA polymerase II machinery. Gene. 2012;**505**:318-323. DOI: 10.1016/j.gene.2012.05.059

[42] Tanay A, Regev A, Shamir R. Conservation and evolvability in regulatory networks: The evolution of ribosomal regulation in yeast. Proceedings of the National Academy of Sciences of the United States of America. 2005;**102**:7203-7208. DOI: 10.1073/pnas.0502521102

[43] Li B, Nierras CR, Warner JR. Transcriptional elements involved in the repression of ribosomal protein synthesis. Molecular and Cellular

Biology. 1999;**19**:5393-5404. DOI: 10.1128/MCB.19.8.5393

[44] Knutson BA, Hahn S. Yeast Rrn7 and human TAF1B are TFIIB-related RNA polymerase I general transcription factors. Science. 2011;**333**:1637-1640. DOI: 10.1126/science.1207699

[45] Moreira-Ramos S, Rojas DA, Montes M, Urbina F, Miralles VJ, Maldonado E. Casein kinase 2 inhibits HomolD-directed transcription by Rrn7 in *Schizosaccharomyces pombe*. The FEBS Journal. 2015;**282**:491-503. DOI: 10.1111/febs.13157

[46] Montes M, Moreira-Ramos S, Rojas DA, Urbina F, Käufer NF, Maldonado E. RNA polymerase II components and Rrn7 form a preinitiation complex on the HomolD box to promote ribosomal protein gene expression in *Schizosaccharomyces pombe*. The FEBS Journal. 2017;**284**:615-633. DOI: 10.1111/febs.14006

[47] Nischt R, Thüroff E, Küfer NF. Molecular cloning of a ribosomal protein gene from the fission yeast *Schizosaccharomyces pombe*. Current Genetics. 1986;**10**:365-370

[48] Nischt R, Gross T, Gatermann K, Swida U, Käufer NF. Sequence and regulatory responses of a ribosomal protein gene from the fission yeast *Schizosaccharomyces pombe*. Nucleic Acids Research. 1987;**15**:1477-1492

[49] Gross T, Nischt R, Gatermann K, Swida U, Käufer NF. Primary structure of the ribosomal protein gene S6 from *Schizosaccharomyces pombe*. Current Genetics. 1988;**13**:57-63

[50] Liebich I, Köhler G, Witt I, Gross T, Käufer NF. Two genes encoding ribosomal protein L3 of *Schizosaccharomyces pombe* and their proximal promoter regions. Gene. 1994;**142**:119-122. DOI: 10.1016/0378-1119(94)90365-4

[51] Schawalder SB, Kabani M, Howald I, Choudhury U, Werner M, Shore D. Growth-regulated recruitment of the essential yeast ribosomal protein gene activator Ifh1. Nature. 2004;**432**:1058-1061. DOI: 10.1038/nature03200

[52] Wade JT, Hall DB, Struhl K. The transcription factor Ifh1 is a key regulator of yeast ribosomal protein genes. Nature. 2004;**432**:1054-1058. DOI: 10.1038/nature03175

[53] Diao LT, Xiao ZD, Leng XM, Li B, Li JH, Luo YP, et al. Conservation and divergence of transcriptional coregulations between box C/D snoRNA and ribosomal protein genes in Ascomycota. RNA. 2014;**20**:1376-1385. DOI: 10.1261/rna.042309.113

[54] Rudra D, Mallick J, Zhao Y, Warner JR. Potential interface between ribosomal protein production and pre-rRNA processing. Molecular and Cellular Biology. 2007;**27**:4815-4824. DOI: 10.1128/MCB.02062-06

[55] Rudra D, Zhao Y, Warner JR. Central role of Ifhl1p-Fhl1p interaction in the synthesis of yeast ribosomal proteins. The EMBO Journal. 2005;**24**:533-542. DOI: 10.1038/sj.emboj.7600553

[56] Hall DB, Wade JT, Struhl K. An HMG protein, Hmo1, associates with promoters of many ribosomal protein genes and throughout the rRNA gene locus in *Saccharomyces cerevisiae*. Molecular and Cellular Biology. 2006;**26**:3672-3679. DOI: 10.1128/MCB.26.9.3672-3679.2006

[57] Perry RP. The architecture of mammalian ribosomal protein promoters. BMC Evolutionary Biology. 2005;**5**:15. DOI: 10.1186/1471-2148-5-15

[58] Maldonado E. Transcriptional functions of a new mammalian TATA-binding protein-related factor. The Journal of Biological Chemistry. 1999;**274**:12963-12966. DOI: 10.1074/jbc.274.19.12963

[59] Moore PA, Ozer J, Salunek M, Jan G, Zerby D, Campbell S, et al. A human TATA binding protein-related protein with altered DNA binding specificity inhibits transcription from multiple promoters and activators. Molecular and Cellular Biology. 1999;**19**:7610-7620. DOI: 10.1128/MCB.19.11.7610

[60] Wang YL, Duttke SH, Chen K, Johnston J, Kassavetis GA, Zeitlinger J, et al. TRF2, but not TBP, mediates the transcription of ribosomal protein genes. Genes & Development. 2014;**28**: 1550-1555. DOI: 10.1101/gad.245662.114

[61] Kedmi A, Zehavi Y, Glick Y, Orenstein Y, Ideses D, Wachtel C, et al. Drosophila TRF2 is a preferential core promoter regulator. Genes & Development. 2014;**28**:2163-2174. DOI: 10.1101/gad.245670.114

Chapter 9

Repurposing *E. coli* by Engineering Quorum Sensing and Redox Genetic Circuits

Sally Wang, Gregory F. Payne and William E. Bentley

Abstract

Because cells have the extraordinary ability to sense and respond to even subtle environmental changes by intricately regulating their gene expression patterns, their behaviors can be intentionally "tuned" by altering the state of their environments in a prescribed or rational manner. Rational control of both external and internal molecular stimuli provides a basis for many biotechnological applications including the expression of foreign protein products. This is done by coordinately controlling product synthesis while retaining the cell in a productive state. Quorum sensing (QS), a molecular signaling modality that mediates cell-cell communication, autonomously facilitates both inter- and intra-species gene regulation. This process can be rewired to enable autonomously actuated, but molecularly programmed, genetic control. Recently, even electrical signals, which have long been used to control the most sophisticated of man-made devices, are now employed to alter cell signaling processes enabling computer programmed behavior, particularly in cells suitably engineered to accommodate electrical signals. By minimally engineering these genetic circuits, new applications have emerged for the repurposing of *Escherichia coli*, from creating innovative sensor concepts to stimulating the emerging field of electrogenetics.

Keywords: protein expression, quorum sensing, autonomous induction, cell-cell communication, redox signaling, electro-induction, synthetic biology

1. Introduction

Controlling the processes or functioning of biological systems has profound implications in biotechnological and other applications [1]. By controlling gene expression, cell behavior and responses to environmental cues can, in turn, be regulated. Ever since the dawn of biotechnology, scientists have been searching for new and better methods to specifically modulate gene expression. Biological systems, however, possess the ability to sense and respond to internal or environmental changes through tuning their own genetic networks. For example, they are capable of detecting metabolic stress during foreign protein expression, and in response, express genes that brake or facilitate the process. Cells are also able to receive signaling molecules from their 'neighbors', to then begin functioning as a 'collective' or population by activating certain genetic regulons. In this chapter, gene-regulating technologies of prokaryotes are discussed that intentionally alter the intracellular landscape for protein expression

as well as the extracellular microenvironmental state in the vicinity of 'designer' production strains in order to program gene expression and behavior. These techniques incorporate the understanding of cell metabolism and the transcriptome, cell-cell communication (previously reviewed by [2, 3]), and biological redox reactions (previously reviewed by [4]). This chapter will mainly focus on recent advances in how actuation of genes is accomplished in *Escherichia coli* through methods that require only minimal genetic rewiring and the technologies developed on such platforms, for instance those of biosensors and bioelectric devices.

2. Optimizing protein expression: rational control of cell condition

There is no doubt that among the myriads of systems available for heterologous protein expression, the Gram-negative bacterium *Escherichia coli* remains one of the most popular owing to its relative simplicity, its inexpensive and fast high-density cultivation, its well-known genetics, and the large number of cloning vectors and mutant host strains that are commonly available. Though not every gene can be efficiently and fully expressed in this system, much progress has been made to improve the performance and versatility of this workhorse microbe. One of the most sought after outcomes is the overexpression of high quality target proteins, however difficulties such as stimulated protease activity and reduced growth rate, as pointed out decades ago, often arise accompanying overexpression [5–7]. In this section, a brief review of the general background pertaining to *E. coli* protein overexpression is presented.

2.1 Reducing bottlenecks: protease activity

The reduction in growth rate is particularly problematic, not only does it contribute to segregational plasmid instability, but severe growth rate perturbations at the onset of induced foreign protein synthesis have been shown to inhibit further expression of the desired protein [8]. Therefore, high levels of foreign protein expression are often unsustainable. Moreover, increased protease activity upon induction and overexpression of foreign protein generally leads to increased proteolysis, as described elsewhere [9–11]. These protease activities with uncharacterized specificity can be considered detrimental to the stability of the recombinant protein. Inefficient cell metabolism during overexpression, as indicated by acetate secretion of host cells, also results in lower protein expression [12]. These cell responses can greatly diminish the genetically-focused efforts to maximize both the final yield and concentration of recombinant proteins by increasing gene expression. In attempting to overcome these hurdles, cell dynamics during induced expression of chloramphenicol acetyl-transferase (CAT) expression have been examined and mathematically modeled in [13], suggesting that induction with an optimized amount of inducer (IPTG) at the onset of stationary phase can avoid growth rate suppression and achieve high expression. However, stimulated protease activity can be still observed. Intracellular proteases of recombinant *E. coli* have been differentiated by proteolytic activity and molecular weight and further characterized during the time course of protein overexpression [14]. Enhanced protease activity can respond quickly to induction, quicker than even the accumulation of the recombinant protein itself. To elicit and identify the proteases, transcriptional profiles of *E. coli* under stress of overexpression have been be mapped [15, 16]. Molecular chaperones (*groEL, ibpA*), lysis gene *mltB* and other DNA damage/bacteriophage associated genes (*recA, alpA, uvrB*) are all observed to be up-regulated along with proteases like *degP* and *ftsH*. It is also reported that cytoplasmic overexpression results in increased activity and expression of an outer membrane protease

OmpT [17]. With this understanding, "cell-conditioning" by adding dithiothreitol (DTT) to alter the levels of the aforementioned host cell proteins prior to product (e.g., CAT) overexpression is capable of placing the cell in a particularly productive state, the result being a doubling of product level [15]. Other methods such as RNA interference (RNAi), and more recent CRISPR technologies can be exploited to downregulate bottlenecks, such as proteases, while ensuring maximal expression of the desired genes. These methods can be targeted to specific genes or even entire regulons, depending on the applied stress and the desired effects.

2.2 Reducing bottlenecks: transcription factors

Levels of the global heat shock transcription factor, σ^{32}, for example, have been shown to increase rapidly during stress, including the stress associated with heterologous protein overexpression [18–23]. Indeed, a variety of cellular stresses induce the σ^{32}-mediated stress response, including both ethanol and heat shock [19–22, 24]. While σ^{32} accumulation could be mediated by control of transcription and translation, its accumulation following production of recombinant protein is mainly due to an altering of its otherwise chaperone-sequestered state [19, 25]. To facilitate protein expression in recombinant *E. coli*, many have posited that simultaneous downregulation of global regulators (such as σ^{32}) could simultaneously reduce the level of negative bottlenecks, such as the σ^{32}-activated proteases. Noting that σ^{32}-mutation is lethal at elevated temperatures [25, 26], methods such as RNAi were shown to transiently downregulate the σ^{32} stress response *in vivo* and these proved to be immensely advantageous. That is, using plasmids constructed with an antisense fragment of the σ^{32} gene, an early study showed that this successfully downregulated the expression of σ^{32} during the production of organophosphorus hydrolase (OPH), resulting increase specific OPH activity by six-fold compared to non-antisense-producing cultures [27].

2.3 Reducing bottlenecks: perspectives

Indeed, there have been countless studies demonstrating techniques to enhance the production of protein over the past 40+ years since recombinant DNA technology was first introduced. Besides choosing the right amount and type of inducer, optimal fermentation conditions have been developed to alleviate the reduction of growth rate during overexpression and enhance yield. Increasing stability of the protein product can also overcome the increased protease activity, this in addition to downregulation of protease-specific regulators. On top of the examples described above, an excellent review by Makrides [28] and a more recent review by Rosano [29] have discussed the various niches within which one can dig deeper in order to achieve higher yield and activity of the desired recombinant protein product.

We note that the majority of these methodologies have targeted either cell-based genetic regulatory structures, the sequence space and alterations of the protein of interest, or the operating policies of the reactors used to cultivate the overproducing cells. These cells, in turn, have typically been monocultures of an optimized host. Rarely have methodologies appeared in which collectives of cells, either monocultures or controlled co-cultures or consortia, and the exogenous signaling thereof are used to produce products such as recombinant proteins. Particularly useful when the engineering of a particular host overburdens its natural regulatory circuitry, cell consortia or collectives provide an interesting alternative. Co-culture and small consortia concepts have recently emerged. Moreover, new methodologies for orthogonal stimulation of genetic circuits can minimize pleiotropic or off-target effects normally accompanying more common chemical inducers. In the sections

that follow, we describe efforts to minimally alter the native bacterial signaling processes of quorum sensing and oxidative stress to repurpose *E. coli* for application in new platforms.

3. Decipher the bacterial dialog: quorum sensing

Gene expression in bacteria can be regulated by a wide array of intra- and extracellular cues. On top of the common chemical inducers that are most often introduced manually to initiate protein overexpression, bacteria are actually capable of producing their own extracellular signals for intercellular communication. The term "quorum sensing (QS)" was coined by EP Greenberg and colleagues decades ago, to describe the phenomena where the secretion and perception of small signaling molecules are transduced to coordinate behavior of a minimal unit (quorum) of microorganisms. Since then, there's been an explosion in understanding how bacteria communicate with themselves. In this section, well-characterized quorum-sensing systems and types of signals, receptors, mechanisms of signal transduction, and target outputs of each system are introduced. In addition, since quorum sensing in many bacteria is also shown to control gene expression in a global manner, several regulons will be introduced, again with the focus on *E. coli* and their potential application. Lastly, beyond controlling gene expression on a global scale, quorum sensing allows bacteria to communicate within and between species. Common pathways and inducers of interspecies communication will be introduced, and we will highlight some of the many applications built upon this ability to communicate not only between species, but also between kingdoms and non-biological substances. That is, by introducing QS phenomena, we develop its potential for keying protein expression via genetic or other means to cue its signaling processes.

3.1 Quorum sensing and its networks

Quorum sensing bacteria produce and release chemical signal molecules termed autoinducers, whose external concentration increases as a function of increasing cell-population density. Once the bacteria detect that autoinducers have reached a minimal threshold level of stimulatory concentration, they will respond by altering their gene expression and behavior. Autoinducers are the cues by which QS bacteria communicate and synchronize particular behaviors on a population-wide scale, thus gaining the ability to function as a multicellular organism.

3.1.1 LuxIR system of V. fischeri

Quorum sensing mechanisms vary from species to species, and hence here we introduce the first-described QS system of the bioluminescent marine bacterium *Vibrio fischeri* as a paradigm for most systems in Gram-negative bacteria [30]. Relevant differences for each organism will be provided as necessary, yet an excellent review by Waters [31] has described most known systems in detail. *V. fischeri* infects higher order organisms, such as luminescent Hawaiian squid *Euprymna scolopes*, within its light organ is completely occupied by the bacterium. When confined, the bacterial population density can reach up to 10^{11} cells per ml and at that point luminescence genes are expressed through a QS mechanism. The luminescence shed by the bacterial consortium can be used, presumably for counterillumination to mask the squid's shadow so that it avoids predation.

Figure 1 illustrates the QS system of *V. fischeri*. Protein LuxI and LuxR control expression of the luciferase operon (*luxICDABE*) required for luminescence

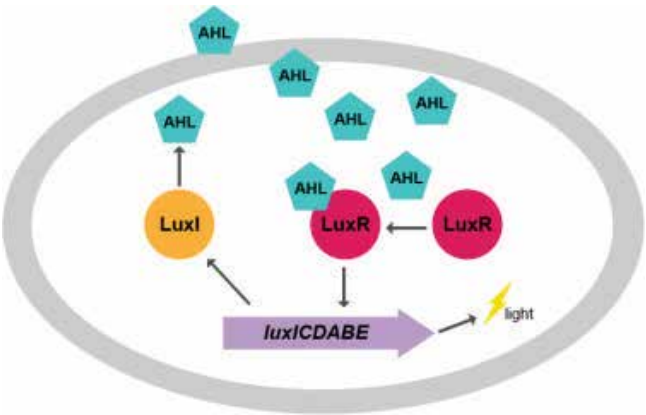

Figure 1.
Quorum sensing in Vibrio fischeri green pentagons denote AHL autoinducer that LuxI produces (3OC6-homoserine lactone). Transcriptional regulator, LuxR, modulates expression of AHL synthase, LuxI, and the lux operon, leading to luciferase-mediated light emission.

production. *LuxI* encodes for an autoinducer synthase that produces the acyl-homoserine lactone (AHL) autoinducer 3OC6-homoserine lactone. Following its production, the AHL will begin to accumulate - its concentration increasing as the cell density increases. Upon reaching a critical level, LuxR the cytoplasmic autoinducer receptor/DNA-binding transcriptional activator, will bind to AHL and this complex will initiate the expression of the luciferase operon. This actuates a positive feedback loop, as LuxI is encoded in the operon, and soon the environment will be flooded with AHL which, in turn, switches all bacteria nearby to the QS active, light-producing mode [32]. The system observed in other *Vibrio* species is more complex, with additional sensing and phosphorylation components in the upstream of *luxR* [33]. In addition, small RNA (sRNA) have been shown to play a vital role in regulating the quorum circuits of *Vibrio harveyi* and *Vibrio cholerae* [34]. These LuxRI-type systems are mostly used for intraspecies communication, as extreme specificity exists between LuxR proteins and their cognate AHL autoinducer ligands.

3.1.2 LuxS/AI-2 system of E. coli

While some of the *Vibrio* QS components are present in *E. coli* (and *Salmonella* strains), the QS system of both species has been found to be distinctively different than that of the *Vibrio*. Several interspecies signaling systems have been identified: those mediated by LuxR homolog SdiA; the LuxS/autoinducer 2 (AI-2) system; an AI-3 system; and a signaling system mediated by indole [35]. Remarkably, the LuxS/AI-2 system possesses the unique feature of endowing cell-population-dependent behavior while interacting with central metabolism through the intracellular activated methyl cycle. LuxS intervenes in central metabolism by functioning in the pathway for metabolism of *S*-adenosylmethionine (SAM), the major cellular methyl donor. Transfer of the methyl moiety to various substrates produces the toxic by-product *S*-adenosylhomocysteine (SAH); while LuxS-containing bacteria have two enzymes (Pfs and LuxS) acting sequentially to convert SAH to adenine, homocysteine, and the signaling molecule DPD [31]. Together, LuxS/AI-2 system has the potential to regulate both gene expression and the cell fitness.

The *luxS* gene, which has a wide range of functions between numerous species, is responsible for AI-2 synthesis in QS networks. However, it was noted that the *luxS* transcriptional profile was reportedly unsynchronized with the accumulation profile of extracellular AI-2 in bacterial supernatants. Confounding its interpretation,

researchers turned toward the signal recognition motif. Thus, another component of the system: the *luxS*-regulated (Lsr) transporter that intakes the extracellular AI-2 was later discovered to be the reason behind the decrease in extracellular AI-2, and not LuxS protein, during stationary phase. As a part of the *lsr* operon, this ATP-binding cassette (ABC) transporter is regulated by both cyclic AMP/cyclic AMP receptor protein and LsrK/LsrR proteins that are transcribed in its own *lsrRK* operon located upstream of *lsr* [36]. The fact that AI-2 intake requires a separate transporter (LsrACDB) is backed up by [37]. Comparing to AI-1 (AHL, 3OC6-homoserine lactone), AI-2 (4,5-dihydroxy-2,3-pentanedione, DPD) is found to be less membrane active and does not intercalate into the bacterial membrane. After modification with carbon chains, products (especially heptyl AI-2) display strong surface activity. These results indicate that AI-2, a more hydrophilic entity, shows less affinity to lipids and thus requires a transportation system. **Figure 2** provides a schematic illustration of the *lsr* circuit comprising of *lsrACDB* (encoding the Lsr transporter), *lsrR* (encodes the transcriptional repressor), *lsrK* (encodes the AI-2 kinase), and *lsrFG* (encodes phosphorylated AI-2 (AI-2P) degradation enzymes) which are all directly regulated by AI-2. A recent mathematical model of this system was provided by Graff and Bentley [38], which helps to discriminate among hypothetical Lsr regulatory mechanisms and points to the importance of repressor LsrR dimer formation and binding on genetic regulation. Desynchronization of Lsr QS system, unlike the LuxIR system where its topology only consists of positive feedback, can display bimodal Lsr signaling and fractional induction. This phenomenon has been both observed in experiments and was also simulated with a mathematical model [39].

3.2 Global quorum sensing regulons

3.2.1 Global genetic regulation of LuxIR and AI-2/Lsr systems

The dawn of genomic profiling has unveiled that quorum sensing, in many bacteria, controls gene expression in a global manner. QS-mutants of *S. pneumoniae*

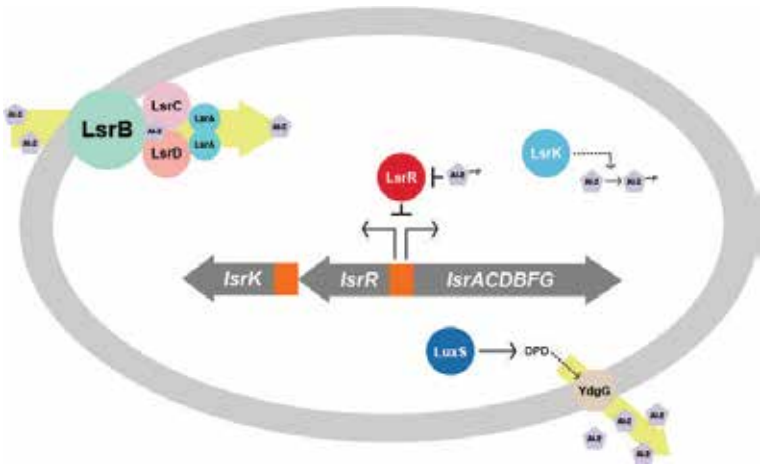

Figure 2.
Regulatory mechanisms of the lsr/AI-2 circuit in E. coli. AI-2 is imported by the Lsr transporter (LsrACDB) and in turn, is processed by LsrK, transforming to its phosphorylated form (AI-2P). As AI-2P binds LsrR, it relieves the repression of LsrR on the Lsr genes and accelerates AI-2 intake. LuxS produces DPD, the precursor of AI-2. The autoinducer is then transported out of the bacteria by YdgG (TqsA), a putative transporter belonging to the exporter superfamily [40].

and related *Streptococci* show defects in multiple pathways, including biofilm formation, acid tolerance, bacteriocin production, and virulence [31]. *E. coli*, too, has been reported to elicit broad QS activities. For example, the quantity and architecture of biofilms are regulated by *lsrR/K* through motility QS regulator (MqsR, B3022), as well as the generation of several small RNAs [36, 41]. Together, these and other reports suggest that QS systems control many aspects of the whole genome rather than just one key gene locus. Further evidence that quorum sensing coordinates the control of a large subset of genes comes from transcriptome analyses of an *E. coli luxS* mutant, which showed that 242 genes (5.6% of the whole genome) exhibited significant transcriptional changes upon a 300-fold AI-2 signaling differential [42–44].

Interestingly, AI-2 synthesis and signaling levels are linked to the accumulation of protein product expressed from plasmid-encoded genes [44]. This suggests that recombinant *E. coli* are able to communicate the burden of overexpressing heterologous protein through AI-2 QS pathways. Most recently, the sugar metabolism of *E. coli* was found to be directly connected to the LuxS/AI-2 QS system. That is, HPr, a phosphocarrier protein central to the sugar phosphotransferase system, was recently reported to copurify with LsrK such that the activity of LsrK was inhibited when bound to HPr [45]. In sum, these finding shed new light on how bacteria respond to changing nutrient levels on a population scale. The intentional manipulation of the QS signaling processes, therefore, has become an interesting target for heterologous gene expression in *E. coli* among many other applications [46].

3.2.2 Applications

De novo engineering of gene circuits inside cells is proven to be difficult, in large part due to connectivity to non-targeted pathways and genes [47, 48]. QS regulons, coupling intraspecies communication and global genome regulation, can serve as excellent platforms for many technologies to be built upon, particularly if one understands the regulatory "reach" of the genetic circuits. Attempting to eliminate the variation in phenotype between cells, You et al. coupled gene expression to cell survival and death using the LuxIR QS system [49]. With the 'population control' gene circuit, they successfully regulated the density of an *E. coli* population autonomously and were able to program the dynamics of an entire population despite behavioral variability between individual cells. Based on the same LuxIR system of *V. fischeri* and the QS system of *Bacillus thuringiensis*, a synchronized genetic clock was engineered [50]. This novel gene network with global intercellular coupling can generate synchronized oscillations in a growing population of cells. In biology, synchronized oscillation holds the same importance as in physics and engineering, where it governs many fundamental physiological processes such as cardiac function and circadian rhythm [51]. These studies have set the stage for future development of using microbes as macroscopic biosensors with oscillatory output, as the colony-level synchronized oscillation could diminish single-cell variability in most synthetic gene networks and increase the sensitivity and robustness of response to external signals.

On top of employing the LuxIR system as a platform for innovative genetic and population regulators, intentional rewiring of *E. coli*'s native QS networks can also benefit biotechnological applications. For example, in [52], autonomous induction of recombinant proteins is realized through minimal rewiring of the AI-2/Lsr system. Since the QS network is capable of 'reporting' the metabolic state of a bacterial population and the metabolic burden is self-indicated by this network [44], Tsao et al. made it possible to achieve metabolically-balanced coordination of the entire culture for a user-specified purpose through minimal rewiring of the QS network

and signal amplification by the T7 RNA polymerase [53, 54]. This study demonstrated one cell population was able to guide protein synthesis process of another by guiding intraspecies communication. Moreover, it was reported in [55] that by simply adding conditioned medium (containing a high amount of AI-2) during recombinant protein induction, one can double the yield of active product. Also, by altering the coincident *luxS* expression to control the AI-2 concentration while also inducing heterologous protein expression, they found an optimal condition where protein yield is dramatically increased. The authors further elucidated the mechanism behind this phenomenon: chaperone GroEL was shown to be coincidentally upregulated post-transcriptionally by AI-2. Because of its native role as a stabilizer of heterologous protein and its role in folding, the upregulation of GroEL might be the reason behind the higher product yield.

More endeavors have been made [56] to increase protein yield in this autonomous system through a different approach. With the same intention in mind [49, 50], a new study showed that reduced heterogeneity between independent cells could be achieved by inserting an enhanced feedback loop to the *E. coli*'s native AI-2 QS system. Upon activation of the engineered system, not only does the foreign pET plasmid concurrently express more sfGFP signal, but it also transcribed more LsrACDB and LsrK than the native *lsr* operon [57]. This overexpression resulted in increased uptake of AI-2, leading to amplified system response and minimized heterogeneity. Heterogeneity, on the other hand, could also be leveraged. In [58], quantized *E. coli* quorums were intentionally assembled through independent engineering of the AI-2 transduction cascade increasing the sensitivity of detector cells. Upon encountering a particular AI-2 level, a discretized sub-population of cells emerge with the desired phenotype. This sensitive, robust detection process could pave the way for future cell-based biosensors for AI-2 and subsequent programmed cell function.

That is, in [59] and as shown in **Figure 3**, *E. coli* were modified to enable programmed motility, sensing and actuation based on the density of user-selected features on nearby surfaces. These 'smart' bacteria can then express marker proteins to indicate phenotypic response based on calculated feature density displayed on the surfaces of nearby eukaryotic cells. Specifically, the AI-2/Lsr signaling pathway was rewired an introduced onto the eukaryotic cells as a 'nanofactory' to direct *E. coli* to swim toward a cancer cell line (SCCHN), where they then initiated synthesis of a drug surrogate based on a threshold density of epidermal growth factor receptor (EGFR). This novel technology represented a new type of targeted drug synthesis and delivery and a new area-based switch that could serve multiple purposes within in the field of synthetic biology.

3.3 Interspecies communication

3.3.1 Universal autoinducer AI-2

Beyond controlling genetic expression on a global scale, quorum sensing allows bacteria to communicate within and between species. This notion arose with the study and discovery of the aforementioned autoinducer AI-2. Derived from SAM as a part of bacterial 1-carbon metabolism, AI-2 is a general term for a family of cyclic furanones utilized in interspecies communication [60]. In LuxS-containing bacteria, SAM is converted into SAH and then broken down by enzymes Pfs and LuxS sequentially into signaling molecule DPD and other byproducts. Due to the high reactivity of DPD, many distinct but related products could be recognized by different bacterial species as AI-2. Though it is postulated that small molecules of similar structure as AI-2 could serve as potential antagonists that halt the bacterial conversation, only a handful are found (compared to a large number of AI-1 inhibitors). In [61], C-1

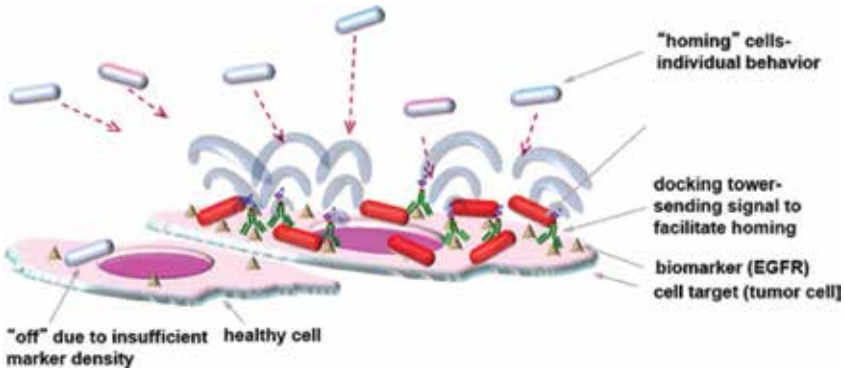

Figure 3.
Biological nanofactories that synthesize AI-2 are targeted to EGFR on the surface of SCCHN cells. AI-2 is emitted from the cell surface and recognized by reprogrammed bacteria, which swim to the site of signal generation and decide, based on AI-2 level (proportional to the EGFR surface density), whether to initiate gene expression (adapted from Wu et al. [59]).

alkyl analogs of AI-2 that quench QS responses in multiple bacterial species simultaneously were developed and synthesized. Interestingly, addition of a single carbon to the C1-alkyl chain of the analog plays a critical role in determining the effect on quenching the QS response. This analog, isobutyl-DPD, was later used to inhibit maturation of *E. coli* biofilms [62]. An expanded and diverse array of AI-2 analogs, including aromatic and cyclic C1-alkyl analogs are synthesized in [63]. Some were identified as species-specific QS disruptors for *E. coli* and *Salmonella typhimurium*, and so were QS quenchers for *Pseudomonas aeruginosa*. Remarkably, these synthetic analogs selectively antagonized quorum sensing among individual bacterial strains within a physiologically relevant polymicrobial culture.

AI-2 is also one of the several signals used by marine bacteria *V. harveyi*. Specifically, AI-2 encoding *luxS* has been found in roughly half of all sequenced bacterial genomes. AI-2 production has been verified in over 80 species, and AI-2 controls gene expression in a variety of bacteria. By using Local Modular Network Alignment Similarity Tool (LMNAST) to study gene order and generate homologous loci, the AI-2/Lsr system was reported to be phylogenetically more dispersed than the well-studied *lac* operon, while its distribution remained densest among gammaproteobacteria [64]. These findings together reinforced the hypothesis that bacteria use AI-2 to communicate between species [31, 65].

Interkingdom communication was also shown to be mediated by AI-2. In [66], transcriptomic effects of bacterial secretions from two nonpathogenic *E. coli* strains (BL21 and W3110) on the human colonic cell line HCT-8 were explored using RNA-Seq. Expression of inflammatory cytokine interleukin 8 (IL-8) in HCT-8 cells was found to respond to AI-2 with a pattern of rapid upregulation followed by a subsequent downregulation after 24 h. This discovery helps provide a deeper understanding of the relationship of microbiome and the host, which is of significant importance in maintaining human health.

3.3.2 Applications

This discovery suggests that AI-2 QS manipulation might find application in guiding human physiology and that 'smart' bacteria, those making heterologous proteins such as drugs or essential nutrients and that otherwise serve as decision makers, might find application in a variety of other fields. As an extension, Lentini et al. [67] engineered minimal 'artificial' cells capable of expressing AI-2 synthesizing fusion protein His_6-LuxS-Pfs-Tyr_5 (HLPT) [68] wherein newly synthesized

AI-2 was proven to induce luminescence in nearby cells, particularly an AI-2 reporter strain of *V. harveyi*. This not only demonstrates QS-mediate communication between cells and non-biological, artificial cell mimics, but presents a new technique to alter the complex networks of natural cells without tampering with the original genetic makeup.

Developing, silencing, or intervening with the communication between cells has revolutionized the way we control gene expression. In [69], communication between cells is developed further by modifying the biological 'nanofactories' proposed by LeDuc et al. [70] to trigger QS responses in the absence of autoinducers. They are self-assembled and comprised of four functional modules: a targeting module (an antibody), a material sensing module, an assembly module, and a synthesis module (fusion protein His_6-Protein G-LuxS-Pfs-Tyr_5, HGLPT, (**Figure 4**). Protein G (assembly module) allows the chimeric enzyme to attach to a targeting antibody *ex vivo*, and LuxS and Pfs together convert raw material, SAH, into autoinducer AI-2. The targeting antibody is proven to successfully attach onto targeted *S. typhimurium* in a mixed culture that also includes *E. coli*. Remarkably, this study built up interspecies 'conversation' between cells that do not usually communicate with each other. After *E. coli*-targeted nanofactories were added to non-QS *E. coli* to 'unmute' the null *E. coli*, the activated *E. coli* are co-cultured with reporter *luxS* null *S. typhimurium*. As the levels of activated *E. coli* increased, *S. typhimurium* begin to 'respond' as they received the AI-2 produced by activated *E. coli* and initiate the expression of their own reporter gene. Interkingdom communication between *E. coli* and human intestine epithelial (Caco-2) cells was also developed using this technique [71]. This tool may be very useful for interrogating and interpreting signaling events in human GI tract.

Perhaps next generation antimicrobials can be created by intercepting bacterial communication and creating 'smart' bacteria. Instead of targeting the viability of pathogenic strains, interruption of their communication is proposed, as it is hypothesized that there will be less selective pressure to develop resistance if instead one targets the mechanisms keyed to pathogenicity [72]. As a global autoinducer, inhibition of the signal AI-2 could possibly lead to decreased virulence in a variety of bacterial species. Many parts of the AI-2/LuxS system, from signal generators (Pfs and LuxS) to signal receptors are all likely targets for inhibition, especially as there are many synthesized AI-2 analogs that are available for quorum quenching [61–63, 73]. In another case [74], probiotic *E. coli* were themselves, engineered to eliminate and prevent *P. aeruginosa* gut infection by reducing biofilm formation. However, it was the *P. aeruginosa*-secreted, species-specific autoinducer AHL ($3OC_{12}HSL$) secreted detected by the probiotic *E. coli* and served as the trigger the for the expression of an anti-biofilm enzyme dispersin B (DspB) and a *P. aeruginosa* toxin, pyocin.

In addition to potential for therapeutic synthesis and delivery, *E. coli* cells can be rewired to serve in networks that provide molecular information about their surroundings or as cell sensors or 'sentinels'. For example in [75], engineered *E. coli* sentinels are made to recognize and move toward hydrogen peroxide, a non-native chemoattractant and potential toxin. Similarly, commensal gastrointestinal strain *E. coli* Nissle 1917 are engineered to recognize gastrointestinal dysfunction biomarker nitric oxide (NO) [76]. These 'smart' bacterial sensors can generate strong fluorescent response upon NO recognition and may serve as simple diagnostic tool for diseases like Crohn's Disease and ulcerative colitis. In [77], nano-guided cell networks that serve as conveyors of molecular communication are developed (**Figure 5**). This system interprets molecular information by intercepting diverse molecular inputs, processes them autonomously through independent cell units within the system and refines output to include positive responders that are viewed via orthogonal, simple optical means. That is,

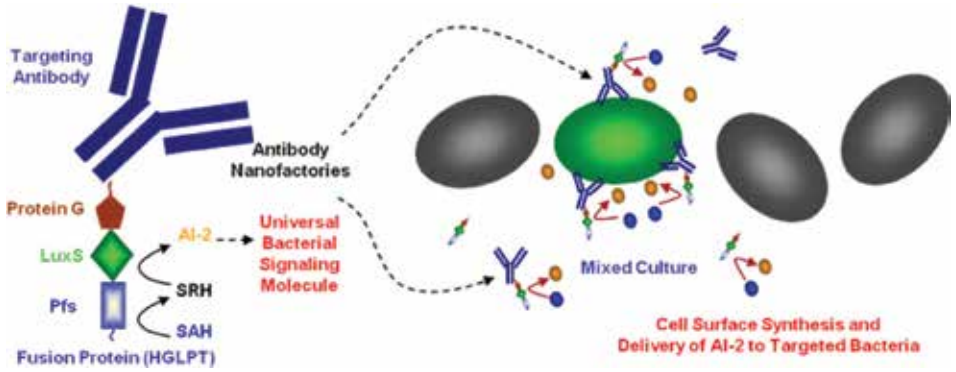

Figure 4.
Biological nanofactory induced interspecies communication. SAH (blue circle) is converted into AI-2 (yellow circle) by the nanofactory fusion protein anchored onto E. coli. AI-2 thus activated QS gene expression in reporter cells (adapted from Fernandes et al. [69]).

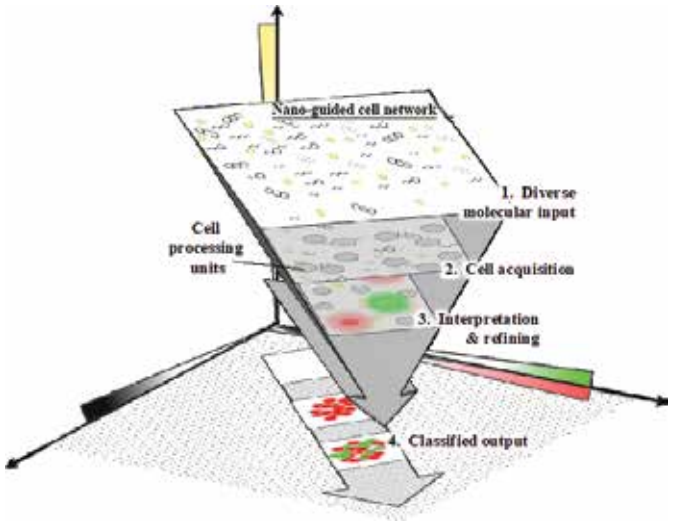

Figure 5.
Schematic of a cell population and nanomaterial-based network. This conceptual system describes cells and magnetic nanoparticle networks that intercept diverse molecular inputs, process them autonomously through independent cell units, and refines output to include positive responders that are viewed via visual classification (red or red and green, adapted from Terrell et al. [77]).

in the preceding sections we have described how engineering cells and the signaling processes that guide their behavior can be used to enhance the overall expression of proteins, but also that when coupled with more advanced functions, cells can serve as their own autonomous factories or surveyors of various microenvironments. A key to performing these functions in an optimal manner is the control of the signaling process, the signal itself, its positioning, its strength or frequency, and its recognition by the engineered cell.

4. Bridging the bio-electro interface: Redox signaling and electrogenetic systems

In addition to quorum sensing, bacteria use numerous other small chemical molecules to build up conversations between themselves and with the environment.

It is well known that redox reactions and redox based signaling pervade living cells and are extremely crucial to both anabolic and catabolic metabolism. Redox-based molecular systems, however, are also leveraged by bacteria for communication. Cells must detect a variety of oxidative stressors and quickly respond so as to avoid oxidative damage and maintain redox balance in order to survive. In this section, several redox signaling pathways will be introduced, yet emphasis will be on how redox signaling and electrochemistry help connect communication and information transfer between biological systems and electronic devices. In this way, redox molecules can serve as exogenous and electronically-programmed controllers of biological function.

4.1 Redox signaling in biological systems

In response to redox imbalance, new metabolic pathways are initiated, the repair or bypassing of damaged cellular components is coordinated and systems that protect the cell from further damage are induced. Throughout the years, many studies have revealed a vast repertoire of elegant solutions that have evolved to allow bacteria to sense and respond to different redox signals [78]. Below, two oxidative stress sensors, SoxR and OxyR, and their corresponding signaling pathway will be introduced. These systems are later shown to enable electrical control of gene expression in *E. coli*.

4.1.1 SoxR: [Fe-S]-cluster based, superoxide/nitric oxide stress sensor

The *E. coli* SoxRS system enhances the production of ~45 proteins in response to superoxide exposure, including those in detoxification (*sodA*, manganese superoxide dismutase), DNA repair (*nfo*, endonuclease IV), maintaining cellular reducing power (*zwf*, glucose-6-phosphate dehydrogenase) and central metabolism (*fumC*, superoxide-stable fumarase C and *acnA*, aconitase A). The *E. coli* SoxR protein exists as a homodimer that contains one [2Fe–2S] cluster per subunit. During aerobic growth, up to 95% of SoxR are held in the reduced ($[2Fe–2S]^{1+}$) state. Upon sensing conditions that promote the production of superoxide, SoxR is oxidized to ($[2Fe–2S]^{2+}$) clusters and it leaves the *SoxR/S* promoter region (*psoxRS*) to activate the expression of transcription factor SoxS. SoxS, unlike SoxR, when bound to *psoxRS* initiates the expression of the proteins listed above located downstream of the promoter [78, 79].

4.1.2 OxyR: thiol-based, peroxide stress sensor

The *E. coli* transcriptional activator, OxyR, is a member of the LysR family of transcriptional regulators. Although it is often cited as the model for bacterial redox sensors, the precise mechanism of thiol modification and the consequences for OxyR activity are the subject of ongoing controversy [80]. Like SoxR, OxyR acts as a repressor of *oxyS* RNA transcription in *E. coli*. Oxidation of cysteine residues in OxyR results in a dramatic secondary structure rearrangement, which leads to a change in the DNA-binding specificity of OxyR, recruitment of RNA polymerase to *OxyR/S* promoters, and the subsequent transcriptional activation of downstream genes such as *oxyS*. OxyS RNA, in turn, is a global oxidative stress regulator mediating the activation or repression of over 40 genes, including several detoxifying enzymes such as hydroperoxidase I (*katG*) and alkylhydroperoxide reductase (*ahpCF*) [75, 78]. Responses of *katG* and *ahpCF*, along with many genes in *SoxR/S* regulon (*sodA, zwf, fumC* and *acnA*) upon paraquat (superoxide ion regenerating redox reagent) insult have been revealed in [81].

4.2 Redox capacitor and bio-electrode interface communication

To probe bio-related redox reactions/signaling simply and readily, recently developed redox-capacitor films can serve as a bio-electrode interface. These are well-described and have been reviewed [82]. In brief, these electrochemical tools are capable of accepting, storing and donating electrons from mediators commonly used in electrochemistry and also in biology. Biofabricated from catechol and the polysaccharide chitosan, the former can be readily (and reversibly) oxidized. When catechol is oxidized, quinone is formed and it can be covalently grafted onto chitosan. In addition, chitosan can be easily 'electro-assembled' onto electrodes owing to its pH-responsive properties. That is, when a voltage is applied to an electrode submersed in an aqueous solution containing chitosan, the pH near the electrode can be controlled. When basic (above the pKa of chitosan, ~6.5), chitosan will form a hydrophobic network and assemble onto the electrode as a film or hydrogel, depending on the application of the electronic charge. When the catechol/quinone redox couple is integrated into the film, it can serve as a source or sink of electrons. Diffusible redox mediators can be added as they can exchange electrons ('charge/discharge') with the redox-active films. Common biology-related mediators include molecules such as ascorbate and NADH, which can charge and discharge the film. Pyocyanin, a toxin secreted by *P. aeruginosa*, is also found to be able to donate electrons to catechol-chitosan film (charging). This metabolite is noted because it, like many other mediators, can also undergo redox-cycling in the film to amplify outputs and facilitate detection of is host cell. It can similarly carry electrons from electrodes directly to proteins or cells near the electrode where such transfer of 'information' can control biological processes.

4.3 Electrical process modulation and gene induction

Many researchers have endeavored finding new ways to control cell processes. The use of optical means to regulate gene expression has garnered significant attention and resulted in an entire field of optogenetics [83]. Genetic switches that operate on optical signals (even small changes in wavelength or color) have been shown to be powerful exogenous controllers of cell function [84, 85]. More recently, researchers have turned to electronic devices to directly control biochemical reactions. In [86], a transistor-like device is engineered to control glucose metabolism of yeast (*S. cerevisiae*). Changes in gating voltage of the device are reported to bring about acceleration or deceleration of the depletion rate of glucose, and in turn the production rate of end-products (ATP and ethanol). Biofabrication and cell-based communication can also be enhanced through electrical control. In a nano-biosystem [87], electrical signals were used to assemble and tune an enzymatic pathway. The assembly comprised of electrodeposited chitosan film on top of a gold electrode, followed by the enzymatic and covalent grafting of a model enzyme HLPT [68] onto the chitosan scaffold. Through different electrical signals and with the help of diffusible redox mediators (pyocyanin), not only the amount of assembled enzymes but their activity was found to be tunable.

Even more recently, a synthetic, mammalian electro-genetic transcription circuit was created [88]. This was done by linking the electrochemical oxidation of ethanol to acetaldehyde, triggering an acetaldehyde-inducible gene expression circuit. While an indirect outcome of the applied voltage, this was the first study whereby specifically intended gene expression was induced by electronic means. A more direct methodology recently appeared [89] in which the engineered genetic circuit responds directly to the electrode-oxidized signal molecule, opening an entirely new modality for bioelectronic control (**Figure 6**). Again, pyocyanin was

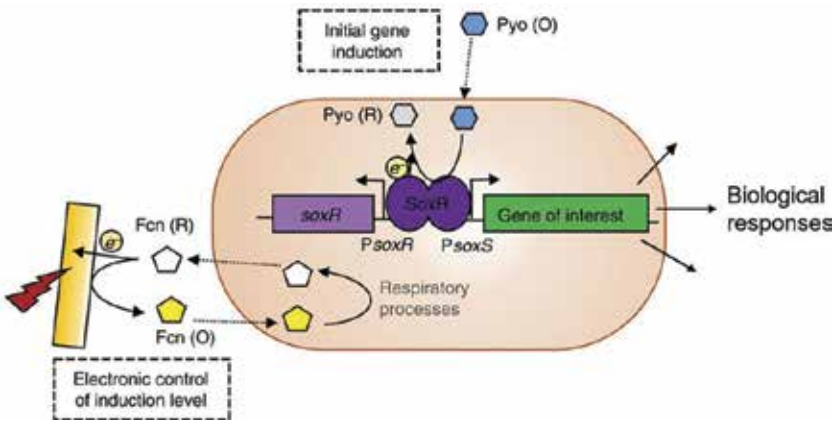

Figure 6.
Electrogenetic induction system scheme. Pyo (O) initiates gene induction and Fcn (R/O), through interactions with respiratory machinery, allows electronic control of induction level. Fcn (R/O), ferro/ferricyanide; Pyo, pyocyanin. Encircled 'e⁻'and arrows indicate electron movement (adapted from Tschirhart et al. [89]).

used in their system, it is responsible for translating electrical signals into a biochemical redox signal that, in turn, can be sensed by SoxR and in sequence, initiate the expression from *psoxS* promoter. Strikingly, gene expression controlled by this device is functionally reversible on relatively short time scales (30–45 min). It was also found to be quite robust, as oscillatory behaviors were shown over many cycles. Accordingly, both optogenetic and electrogenetic systems will require that an entirely new 'suite' of genetic elements be developed that respond to and coordinate these environmental cues. In the recent study, the expression of AHL-synthesizing enzyme LuxI was electronically actuated, resulting in electronic control of QS behavior of nearby cells. Analogously, motility regulator CheZ was also electronically stimulated demonstrating the electronic initiation of cell motility. This study is the first in which electronic signals guided engineered cells and those, in turn, guided others. While this chapter has focused on gene expression in *E. coli*, it also attempts to show how the simultaneous coordination of gene expression and of the host cells can result in interesting and new application areas.

5. Conclusion

Researchers in biotechnology are constantly seeking novel platforms or techniques from which to address problems: those that in a broad sense, have enhanced efficacy, while maintaining or intensifying specificity. In this chapter, innovative means that focus on controlling environmental cues to regulate gene expression are introduced. To optimize heterologous protein expression, methods seeking to repress stress responses and retain cells in a 'productive' state are carried out by carefully engineering host cells to respond to various cues that are either introduced exogenously or endogenously. QS systems have appeared that provide targets for controlling bacterial behavior. They are also shown to report on the prevailing metabolic state of a product-producing cell. Early methodologies such as RNAi, genetic mutation, product protein-directed evolution, all successful means to enhance yield, can be reexamined based on new understanding of how cells communicate with one another. That is, QS systems enable the rewiring of endogenous metabolism for the coordinated control of entire populations of cells. This ushers in a new way of viewing protein or product-producing cells as a cell 'collective' rather than as individual cells each identical to one another, responding to cues or inducers such as

IPTG for the controlled overexpression of heterologous proteins. QS systems enable autonomous global gene regulation based on cell density. That is, instead of direct interrogation and control of genetic circuits, QS-based cell-cell communication allows indirect gene regulation through self-secretion and uptake of small signaling molecules. Further, exogenous and orthogonal signals, such as those provided by optical and electrical means can be interfaced with cells, providing exquisite control of gene expression. Importantly, in host cells were synthetic components contribute minimal perturbation to native systems, exogenously signaled protein expression can be coupled with exogenously controlled cell behavior (e.g., swimming or decision making). Electrochemistry, along with the invention of redox capacitors, thusly opens a new niche for genetic induction. That is, by leveraging the ability of mediators to translate electrical signals into chemical cues, researchers can cue changes in environmental electrical state that, in turn, are capable of inducing gene expression. These innovative methods will no doubt continue to generate impactful applications in fields such as biotechnology and biosensing.

Acknowledgements

This work was supported by DTRA (HDTRA1-13-0037) and NSF (DMREF #1435957) and by the National Institutes of Health (R21EB024102).

Conflict of interest

We declare that we have no conflicts of interest associated with the submitted work.

Author details

Sally Wang[1,2], Gregory F. Payne[2] and William E. Bentley[1,2]*

1 Fischell Department of Bioengineering, University of Maryland, College Park, MD, USA

2 Institute for Bioscience and Biotechnology Research, University of Maryland, College Park, MD, USA

*Address all correspondence to: bentley@umd.edu

References

[1] Kitano H. Systems biology: A brief overview. Science. 2002;**295**(5560):1662-1664

[2] Fuqua C, Parsek MR, Greenberg EP. Regulation of gene expression by cell-to-cell communication: Acyl-homoserine lactone quorum sensing. Annual Review of Genetics. 2001;**35**:439-468

[3] Bassler BL. How bacteria talk to each other: Regulation of gene expression by quorum sensing. Current Opinion in Microbiology. 1999;**2**(6):582-587

[4] Dalton TP, Shertzer HG, Puga A. Regulation of gene expression by reactive oxygen. Annual Review of Pharmacology and Toxicology. 1999;**39**:67-101

[5] Betenbaugh MJ, Beaty C, Dhurjati P. Effects of plasmid amplification and recombinant gene expression on the growth kinetics of recombinant *E. coli*. Biotechnology and Bioengineering. 1989;**33**(11):1425-1436

[6] Goff SA, Goldberg AL. An increased content of protease La, the lon gene product, increases protein degradation and blocks growth in *Escherichia coli*. The Journal of Biological Chemistry. 1987;**262**(10):4508-4515

[7] Okita B, Arcuri E, Turner K, Sharr D, Del Tito B, Swanson J, et al. Effect of induction temperature on the production of malaria antigens in recombinant *E. coli*. Biotechnology and Bioengineering. 1989;**34**(6):854-862

[8] Bentley W. Effects of plasmid-mediated activity on bacterial metabolism and culture stability [thesis]. Boulder, CO: University of Colorado; 1989

[9] Gottesman S, Maurizi MR. Regulation by proteolysis: Energy-dependent proteases and their targets. Microbiological Reviews. 1992;**56**(4):592-621

[10] Maurizi MR. Proteases and protein degradation in *Escherichia coli*. Experientia. 1992;**48**(2):178-201

[11] Grossman AD, Taylor WE, Burton ZF, Burgess RR, Gross CA. Stringent response in *Escherichia coli* induces expression of heat shock proteins. Journal of Molecular Biology. 1985;**186**(2):357-365

[12] Ko YF, Bentley WE, Weigand WA. The effect of cellular energetics on foreign protein production. Applied Biochemistry and Biotechnology. 1995;**50**(2):145-159

[13] Bentley WE, Davis RH, Kompala DS. Dynamics of induced CAT expression in *E. coli*. Biotechnology and Bioengineering. 1991;**38**(7):749-760

[14] Harcum SW, Bentley WE. Response dynamics of 26-, 34-, 39-, 54-, and 80-kDa proteases in induced cultures of recombinant *Escherichia coli*. Biotechnology and Bioengineering. 1993;**42**(6):675-685

[15] Gill RT, DeLisa MP, Valdes JJ, Bentley WE. Genomic analysis of high-cell-density recombinant *Escherichia coli* fermentation and "cell conditioning" for improved recombinant protein yield. Biotechnology and Bioengineering. 2001;**72**(1):85-95

[16] Gill RT, Valdes JJ, Bentley WE. A comparative study of global stress gene regulation in response to overexpression of recombinant proteins in *Escherichia coli*. Metabolic Engineering. 2000;**2**(3):178-189

[17] Gill RT, DeLisa MP, Shiloach M, Holoman TR, Bentley WE. OmpT expression and activity increase in response to recombinant

chloramphenicol acetyltransferase overexpression and heat shock in *E. coli*. Journal of Molecular Microbiology and Biotechnology. 2000;**2**(3):283-289

[18] Blaszczak A, Zylicz M, Georgopoulos C, Liberek K. Both ambient temperature and the DnaK chaperone machine modulate the heat shock response in *Escherichia coli* by regulating the switch between sigma 70 and sigma 32 factors assembled with RNA polymerase. The EMBO Journal. 1995;**14**(20):5085-5093

[19] Kanemori M, Mori H, Yura T. Induction of heat shock proteins by abnormal proteins results from stabilization and not increased synthesis of sigma 32 in *Escherichia coli*. Journal of Bacteriology. 1994;**176**(18):5648-5653

[20] Lesley SA, Thompson NE, Burgess RR. Studies of the role of the *Escherichia coli* heat shock regulatory protein sigma 32 by the use of monoclonal antibodies. The Journal of Biological Chemistry. 1987;**262**(11):5404-5407

[21] Straus DB, Walter WA, Gross CA. The heat shock response of *E. coli* is regulated by changes in the concentration of sigma 32. Nature. 1987;**329**(6137):348-351

[22] Straus DB, Walter WA, Gross CA. The activity of sigma 32 is reduced under conditions of excess heat shock protein production in *Escherichia coli*. Genes & Development. 1989;**3**(12A):2003-2010

[23] Tilly K, Spence J, Georgopoulos C. Modulation of stability of the *Escherichia coli* heat shock regulatory factor sigma. Journal of Bacteriology. 1989;**171**(3):1585-1589

[24] Parsell DA, Lindquist S. The function of heat-shock proteins in stress tolerance: Degradation and reactivation of damaged proteins. Annual Review of Genetics. 1993;**27**:437-496

[25] Neidhardt FC, Curtiss R. *Escherichia coli* and Salmonella: Cellular and Molecular Biology. Vol. 2. 2nd ed. Washington, DC: ASM Press; 1996. (xx, 2822 p.) p

[26] Zhou YN, Kusukawa N, Erickson JW, Gross CA, Yura T. Isolation and characterization of *Escherichia coli* mutants that lack the heat shock sigma factor sigma 32. Journal of Bacteriology. 1988;**170**(8):3640-3649

[27] Srivastava R, Cha HJ, Peterson MS, Bentley WE. Antisense downregulation of sigma(32) as a transient metabolic controller in *Escherichia coli*: Effects on yield of active organophosphorus hydrolase. Applied and Environmental Microbiology. 2000;**66**(10):4366-4371

[28] Makrides SC. Strategies for achieving high-level expression of genes in *Escherichia coli*. Microbiological Reviews. 1996;**60**(3):512-538

[29] Rosano GL, Ceccarelli EA. Recombinant protein expression in *Escherichia coli*: Advances and challenges. Frontiers in Microbiology. 2014;**5**:172

[30] Nealson KH, Hastings JW. Bacterial bioluminescence: its control and ecological significance. Microbiological Reviews. 1979;**43**(4):496-518

[31] Waters CM, Bassler BL. Quorum sensing: Cell-to-cell communication in bacteria. Annual Review of Cell and Developmental Biology. 2005;**21**:319-346

[32] Miller MB, Bassler BL. Quorum sensing in bacteria. Annual Review of Microbiology. 2001;**55**:165-199

[33] March JC, Bentley WE. Quorum sensing and bacterial cross-talk in biotechnology. Current Opinion in Biotechnology. 2004;**15**(5):495-502

[34] Lenz DH, Mok KC, Lilley BN, Kulkarni RV, Wingreen NS, Bassler BL. The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae*. Cell. 2004;**118**(1):69-82

[35] Lee J, Zhang XS, Hegde M, Bentley WE, Jayaraman A, Wood TK. Indole cell signaling occurs primarily at low temperatures in *Escherichia coli*. The ISME Journal. 2008;**2**(10):1007-1023

[36] Li J, Attila C, Wang L, Wood TK, Valdes JJ, Bentley WE. Quorum sensing in *Escherichia coli* is signaled by AI-2/LsrR: Effects on small RNA and biofilm architecture. Journal of Bacteriology. 2007;**189**(16):6011-6020

[37] Kamaraju K, Smith J, Wang J, Roy V, Sintim HO, Bentley WE, et al. Effects on membrane lateral pressure suggest permeation mechanisms for bacterial quorum signaling molecules. Biochemistry. 2011;**50**(32):6983-6993

[38] Graff SM, Bentley WE. Mathematical model of LsrR-binding and derepression in *Escherichia coli* K12. Journal of Bioinformatics and Computational Biology. 2017;**15**(1):1650039

[39] Quan DN, Tsao CY, Wu HC, Bentley WE. Quorum sensing desynchronization leads to bimodality and patterned behaviors. PLoS Computational Biology. 2016;**12**(4):e1004781

[40] Herzberg M, Kaye IK, Peti W, Wood TK. Ydg G (Tqs A) controls biofilm formation in *Escherichia coli* K-12 through autoinducer 2 transport. Journal of Bacteriology. 2006;**188**(2):587-598

[41] Gonzalez Barrios AF, Zuo R, Hashimoto Y, Yang L, Bentley WE, Wood TK. Autoinducer 2 controls biofilm formation in *Escherichia coli* through a novel motility quorum-sensing regulator (Mqs R, B3022). Journal of Bacteriology. 2006;**188**(1):305-316

[42] DeLisa MP, Wu CF, Wang L, Valdes JJ, Bentley WE. DNA microarray-based identification of genes controlled by autoinducer 2-stimulated quorum sensing in *Escherichia coli*. Journal of Bacteriology. 2001;**183**(18):5239-5247

[43] Wang L, Li J, March JC, Valdes JJ, Bentley WE. luxS-dependent gene regulation in *Escherichia coli* K-12 revealed by genomic expression profiling. Journal of Bacteriology. 2005;**187**(24):8350-8360

[44] DeLisa MP, Valdes JJ, Bentley WE. Quorum signaling via AI-2 communicates the "metabolic burden" associated with heterologous protein production in *Escherichia coli*. Biotechnology and Bioengineering. 2001;**75**(4):439-450

[45] Ha JH, Hauk P, Cho K, Eo Y, Ma X, Stephens K, et al. Evidence of link between quorum sensing and sugar metabolism in *Escherichia coli* revealed via cocrystal structures of LsrK and HPr. Science Advances. 2018;**4**(6):eaar7063

[46] DeLisa MP, Bentley WE. Bacterial autoinduction: Looking outside the cell for new metabolic engineering targets. Microbial Cell Factories. 2002;**1**(1):5

[47] Yokobayashi Y, Weiss R, Arnold FH. Directed evolution of a genetic circuit. Proceedings of the National Academy of Sciences of the United States of America. 2002;**99**(26):16587-16591

[48] Chen W, Kallio PT, Bailey JE. Construction and characterization of a novel cross-regulation system for regulating cloned gene expression in *Escherichia coli*. Gene. 1993;**130**(1):15-22

[49] You L, Cox RS 3rd, Weiss R, Arnold FH. Programmed population control by cell-cell communication and regulated killing. Nature. 2004;**428**(6985):868-871

[50] Danino T, Mondragon-Palomino O, Tsimring L, Hasty J. A synchronized quorum of genetic clocks. Nature. 2010;**463**(7279):326-330

[51] Glass L. Synchronization and rhythmic processes in physiology. Nature. 2001;**410**(6825):277-284

[52] Tsao CY, Hooshangi S, Wu HC, Valdes JJ, Bentley WE. Autonomous induction of recombinant proteins by minimally rewiring native quorum sensing regulon of *E. coli*. Metabolic Engineering. 2010;**12**(3):291-297

[53] Studier FW. Protein production by auto-induction in high density shaking cultures. Protein Expression and Purification. 2005;**41**(1):207-234

[54] Studier FW, Moffatt BA. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. Journal of Molecular Biology. 1986;**189**(1):113-130

[55] Tsao CY, Wang L, Hashimoto Y, Yi H, March JC, DeLisa MP, et al. LuxS coexpression enhances yields of recombinant proteins in *Escherichia coli* in part through posttranscriptional control of GroEL. Applied and Environmental Microbiology. 2011;**77**(6):2141-2152

[56] Zargar A, Quan DN, Bentley WE. Enhancing intercellular coordination: Rewiring quorum sensing networks for increased protein expression through autonomous induction. ACS Synthetic Biology. 2016;**5**(9):923-928

[57] Zargar A, Quan DN, Emamian M, Tsao CY, Wu HC, Virgile CR, et al. Rational design of 'controller cells' to manipulate protein and phenotype expression. Metabolic Engineering. 2015;**30**:61-68

[58] Servinsky MD, Terrell JL, Tsao CY, Wu HC, Quan DN, Zargar A, et al. Directed assembly of a bacterial quorum. The ISME Journal. 2016;**10**(1):158-169

[59] Wu HC, Tsao CY, Quan DN, Cheng Y, Servinsky MD, Carter KK, et al. Autonomous bacterial localization and gene expression based on nearby cell receptor density. Molecular Systems Biology. 2013;**9**:636

[60] Singh V, Lee JE, Nunez S, Howell PL, Schramm VL. Transition state structure of 5′-methylthioadenosine/S-adenosylhomocysteine nucleosidase from *Escherichia coli* and its similarity to transition state analogues. Biochemistry. 2005;**44**(35):11647-11659

[61] Roy V, Smith JA, Wang J, Stewart JE, Bentley WE, Sintim HO. Synthetic analogs tailor native AI-2 signaling across bacterial species. Journal of the American Chemical Society. 2010;**132**(32):11141-11150

[62] Roy V, Meyer MT, Smith JA, Gamby S, Sintim HO, Ghodssi R, et al. AI-2 analogs and antibiotics: A synergistic approach to reduce bacterial biofilms. Applied Microbiology and Biotechnology. 2013;**97**(6):2627-2638

[63] Gamby S, Roy V, Guo M, Smith JA, Wang J, Stewart JE, et al. Altering the communication networks of multispecies microbial systems using a diverse toolbox of AI-2 analogues. ACS Chemical Biology. 2012;**7**(6):1023-1030

[64] Quan DN, Bentley WE. Gene network homology in prokaryotes using a similarity search approach: Queries of quorum sensing signal transduction. PLoS Computational Biology. 2012;**8**(8):e1002637

[65] Xavier KB, Bassler BL. LuxS quorum sensing: More than just a numbers

game. Current Opinion in Microbiology. 2003;**6**(2):191-197

[66] Zargar A, Quan DN, Carter KK, Guo M, Sintim HO, Payne GF, et al. Bacterial secretions of nonpathogenic *Escherichia coli* elicit inflammatory pathways: A closer investigation of interkingdom signaling. MBio Journal. 2015;**6**(2):e00025

[67] Lentini R, Martin NY, Forlin M, Belmonte L, Fontana J, Cornella M, et al. Two-way chemical communication between artificial and natural cells. ACS Central Science. 2017;**3**(2):117-123

[68] Fernandes R, Bentley WE. AI-2 biosynthesis module in a magnetic nanofactory alters bacterial response via localized synthesis and delivery. Biotechnology and Bioengineering. 2009;**102**(2):390-399

[69] Fernandes R, Roy V, Wu HC, Bentley WE. Engineered biological nanofactories trigger quorum sensing response in targeted bacteria. Nature Nanotechnology. 2010;**5**(3):213-217

[70] Leduc PR, Wong MS, Ferreira PM, Groff RE, Haslinger K, Koonce MP, et al. Towards an in vivo biologically inspired nanofactory. Nature Nanotechnology. 2007;**2**(1):3-7

[71] Hebert CG, Gupta A, Fernandes R, Tsao CY, Valdes JJ, Bentley WE. Biological nanofactories target and activate epithelial cell surfaces for modulating bacterial quorum sensing and interspecies signaling. ACS Nano. 2010;**4**(11):6923-6931

[72] Hung DT, Shakhnovich EA, Pierson E, Mekalanos JJ. Small-molecule inhibitor of *Vibrio cholerae* virulence and intestinal colonization. Science. 2005;**310**(5748):670-674

[73] Roy V, Adams BL, Bentley WE. Developing next generation antimicrobials by intercepting AI-2 mediated quorum sensing. Enzyme and Microbial Technology. 2011;**49**(2):113-123

[74] Hwang IY, Koh E, Wong A, March JC, Bentley WE, Lee YS, et al. Engineered probiotic *Escherichia coli* can eliminate and prevent *Pseudomonas aeruginosa* gut infection in animal models. Nature Communications. 2017;**8**:15028

[75] Virgile C, Hauk P, Wu HC, Shang W, Tsao CY, Payne GF, et al. Engineering bacterial motility towards hydrogen-peroxide. PLoS One. 2018;**13**(5):e0196999

[76] McKay R, Hauk P, Quan D, Bentley WE. Development of cell-based sentinels for nitric oxide: Ensuring marker expression and unimodality. ACS Synthetic Biology. 2018;**7**(7):1694-1701

[77] Terrell JL, Wu HC, Tsao CY, Barber NB, Servinsky MD, Payne GF, et al. Nano-guided cell networks as conveyors of molecular communication. Nature Communications. 2015;**6**:8500

[78] Green J, Paget MS. Bacterial redox sensors. Nature Reviews. Microbiology. 2004;**2**(12):954-966

[79] Demple B. Redox signaling and gene control in the *Escherichia coli* soxRS oxidative stress regulon—A review. Gene. 1996;**179**(1):53-57

[80] Zheng M, Wang X, Templeton LJ, Smulski DR, LaRossa RA, Storz G. DNA microarray-mediated transcriptional profiling of the *Escherichia coli* response to hydrogen peroxide. Journal of Bacteriology. 2001;**183**(15):4562-4570

[81] Lu C, Albano CR, Bentley WE, Rao G. Quantitative and kinetic study of oxidative stress regulons using green fluorescent protein. Biotechnology and Bioengineering. 2005;**89**(5):574-587

[82] Kim E, Leverage WT, Liu Y, White IM, Bentley WE, Payne GF. Redox-capacitor to connect electrochemistry to redox-biology. The Analyst. 2014;**139**(1):32-43

[83] Fenno L, Yizhar O, Deisseroth K. The development and application of optogenetics. Annual Review of Neuroscience. 2011;**34**:389-412

[84] Levskaya A, Chevalier AA, Tabor JJ, Simpson ZB, Lavery LA, Levy M, et al. Synthetic biology: Engineering *Escherichia coli* to see light. Nature. 2005;**438**(7067):441-442

[85] Schmidl SR, Sheth RU, Wu A, Tabor JJ. Refactoring and optimization of light-switchable *Escherichia coli* two-component systems. ACS Synthetic Biology. 2014;**3**(11):820-831

[86] Song Y, Wang J, Yau ST. Controlled glucose consumption in yeast using a transistor-like device. Scientific Reports. 2014;**4**:5429

[87] Gordonov T, Kim E, Cheng Y, Ben-Yoav H, Ghodssi R, Rubloff G, et al. Electronic modulation of biochemical signal generation. Nature Nanotechnology. 2014;**9**(8):605-610

[88] Weber W, Luzi S, Karlsson M, Sanchez-Bustamante CD, Frey U, Hierlemann A, et al. A synthetic mammalian electro-genetic transcription circuit. Nucleic Acids Research. 2009;**37**(4):e33

[89] Tschirhart T, Kim E, McKay R, Ueda H, Wu HC, Pottash AE, et al. Electronic control of gene expression and cell behaviour in *Escherichia coli* through redox signalling. Nature Communications. 2017;**8**:14030

Edited by Fumiaki Uchiumi

Transcription is the most fundamental nuclear event, by which the information of nucleotide sequences on DNA is transcribed into RNA by multiple proteins, including RNA polymerases. Transcription determines the functions of proteins and the behaviour of cells, appropriately responding to environmental changes.This book is intended for scientists, especially those who are interested in the future prospect of gene expression and control in medicine and industry. This book consists of 9 chapters, divided into four parts. Each chapter is written by experts both in the basic and applied scientific field. A collection of articles presented by active and laboratory-based investigators provides evidence from the research, giving us a rigid platform to discuss "Gene Expression and Control."

Published in London, UK

IntechOpen

ISBN 978-1-83962-113-0