Preface

This book takes readers on a captivating journey to the core of ribonucleic acid (RNA), a vital molecule for life. The book focuses on one critical aspect: RNA extraction. RNA extraction is a crucial step in molecular biology laboratories that aids in understanding how genes function. It has contributed to significant breakthroughs in medicine and agriculture. The aim of this book is to simplify this complex process and make it comprehensible to those who are new to it.

RNA extraction is a precise and intricate process necessary for isolating RNA from various biological samples. This book provides a detailed and visual exploration of the technique, complete with realistic illustrations of the laboratory equipment used in the process. The goal is to provide a clear and practical understanding of the RNA extraction method so that anyone, even without laboratory experience, can comprehend the steps involved.

Mastering the process of RNA extraction puts individuals at the forefront of a crucial scientific process, which has implications in medical research, genetic mechanisms, and therapy development. This book serves as a gateway to gaining in-depth, practical knowledge for those interested in entering the field of molecular biology or understanding what RNA is and how it is extracted.

The idea for writing this book came from a training session I conducted at the University of Ibadan in Nigeria. The training was sponsored by Alexander von Humboldt Foundation through the research group linkage program between Dirk Albach and Mubo A. Sonibare. It was about RNA extraction and data analysis. During the training, I realized that it would be helpful to have a book that makes it easy for students, even for people without prior knowledge in molecular biology, to extract RNA. As far as I know, there is no book available on RNA extraction with real lab device images that allow anyone, especially people in the tropics, to perform the extraction. That is why I thought a book with real lab images would be useful.

I wrote this book during my postdoctoral position at the Prinzessin Therese von Bayern Chair of Systematics, Biodiversity, and Evolution of Plants at the Faculty of Biology of the Ludwig-Maximilians Universität

München. Thanks to this position, I had access to the RNA extraction equipment presented in this book. Additionally, I became more familiar with the RNA extraction technique through various projects that I worked on, courses that I taught, and students that I supervised. I would like to express my gratitude to Gudrun Kadereit, the head of the Chair. This book is dedicated to my parents, Marie Kameni and Damase Wemo, who despite economic difficulties, ensured my education.

Christian Siadjeu
Menzinger Str. 67, Munich

1 Introduction

1.1 RNA – the central workhorse of life

Ribonucleic Acid (RNA) is often regarded as something like the small brother of Deoxyribonucleic Acid (DNA), which is the central molecule for the inheritance of genetic information. However, it is neglected that RNA is the workforce of the cell, ribose is the precursor of deoxyribose, and RNA preceded DNA in the origin of life.

RNA is a nucleic acid. Nucleic acids are biopolymers composed of a backbone of 5-carbon sugars (Fig. 1), the ribose in RNA (Fig. 1a), connected via phosphate groups on the 3′ and 5′ carbon and to a nitrogenous base on the 1′ carbon (Fig. 2a-e). Whereas sugar (ribose in RNA, deoxyribose in DNA) (Fig. 1a-b) and phosphate are homogenous in nucleic acids, the nitrogenous bases vary, thus, storing the information of the nucleic acid. DNA can have adenine (A), cytosine (C), guanine (G), or thymine (T) as bases. Still, the latter is replaced by uracil in RNA (Fig. 3a). Adenine and guanine are purines. In contrast, cytosine, thymine and uracil are pyrimidine derivatives. Pyrimidines are cyclic organic compounds with four carbons and two nitrogen in the cycle (Fig. 2a-c). Purines are derived from pyrimidines but with a second cycle attached to two carbons with another two nitrogens and one carbon atom (Fig. 2d-e). In DNA, two sugar-phosphate strings are attached to each other by hydrogen bonds between one purine and one pyrimidine, either adenine and thymine or cytosine and guanine (Fig. 3b). In RNA, these organic bases are free to attach to other compounds or to other bases of the same strand, thus forming the center of the catalytic function of RNA (Fig. 3a).

It has been discussed for a long time whether DNA or RNA came first in the origin of life. Since deoxyribonucleotides, the single parts of DNA consisting of one sugar, one phosphate group, and one base, are produced from ribonucleotides, the latter are not just considered the initial step in biosynthesis but also evolutionarily. This is the origin of the idea that once a world existed without DNA but RNA (RNA world hypothesis)[1]. Robertson & Joyce[2] discussed intensively how RNA could have formed from

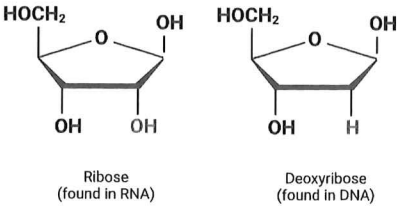

Fig. 1 Sugar bases of DNA and RNA. (a): ribose, (b): deoxyribose (generated using BioRender)

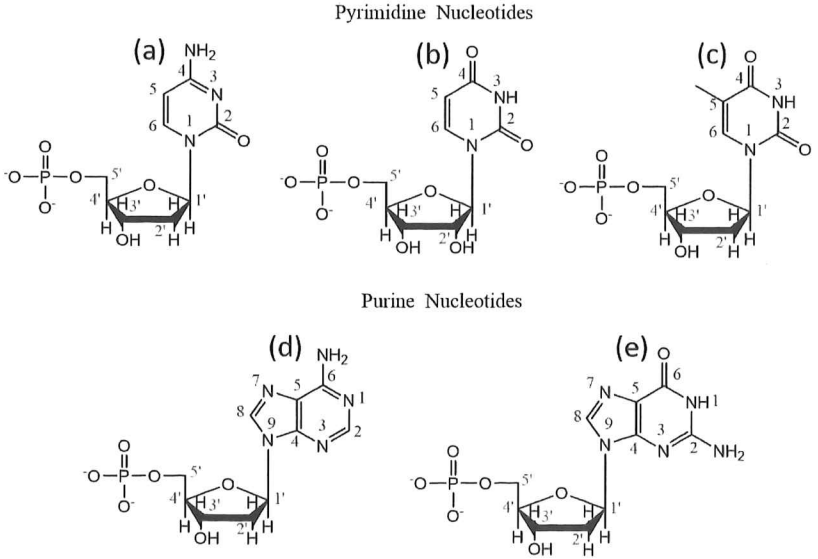

Fig. 2 Nucleotides forming RNA and DNA. (a): cytidine, (b): uridine, (c): thymidine, (d): adenosine, (e): guanosine (generated using ChemDraw v21.0.0)

inorganic and primitive organic compounds. Ribonucleotides, therefore, fulfill three vital criteria for early life. They can be formed from simple compounds occurring in nature, they can store information, and they can catalyze chemical reactions necessary to reproduce new ribonucleotides[3].

In the evolution of the genetic system, DNA has the advantage of higher stability based on the hydrogen bonds between the two nucleotide strands (Fig. 3b). Human chromosomes can have more than 250 million nucleotides in length[4] and they are not the longest. In contrast, RNA rarely has

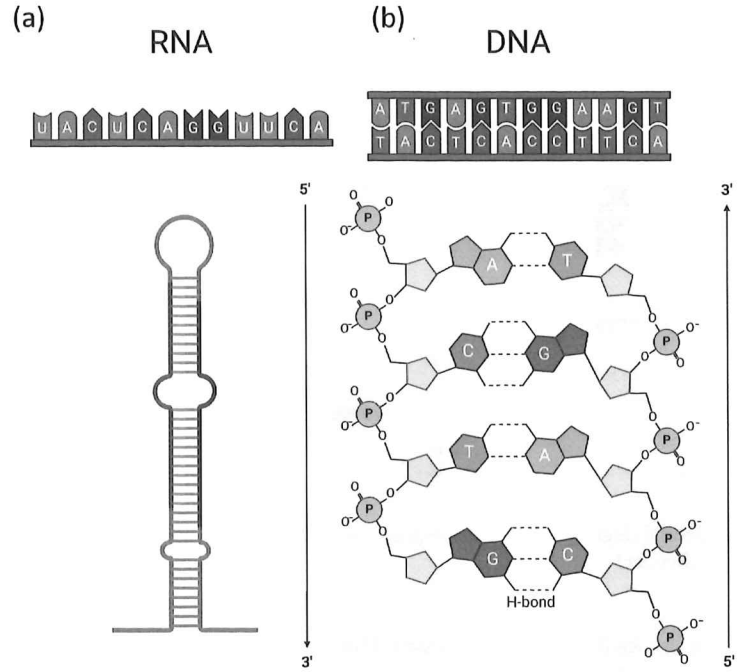

Fig. 3 DNA and RNA molecules. (a): single-stranded RNA forming secondary structure, (b): double-stranded DNA with hydrogen bonds between the two nucleotide strands (generated using BioRender)

more than 3000 nucleotides in length, which partly binds to each other to form a secondary structure (Fig. 3a). Especially higher temperatures and extreme pH can denature RNA faster than DNA. There is also a difference between bacterial mRNA, which is usually degraded within minutes, making their study difficult, and mRNA of eukaryotes, which can be stable for more than 24 hours[5]. However, there is a large variation in stability depending on RNA sequence, RNA secondary structure, temperature, and acidity. This has important implications for the study of differential gene expression, for example when comparing transcriptomes of organisms at different temperatures. Besides these inherent factors limiting the life of RNA molecules, RNA is also actively degraded depending on multiple regulatory elements[6,7].

Nowadays, RNA is the workhorse of all living organisms. Whereas DNA is stored in the cell nucleus with all the information to organize the

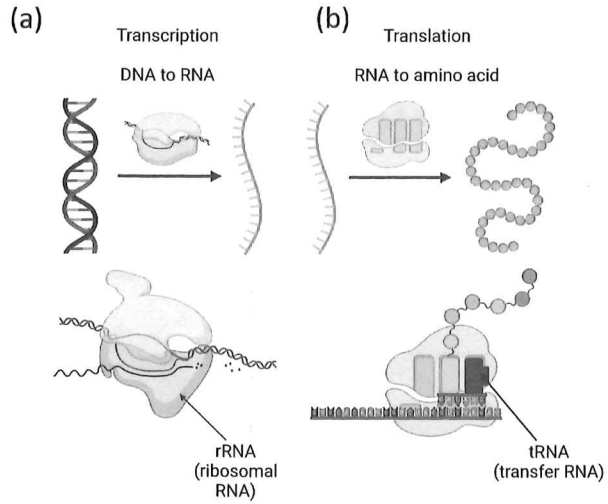

Fig. 4 Central dogma of molecular biology. (a): transcription, (b): translation (generated using BioRender)

cell, much like the honeybee queen, RNA is the link between the DNA and the different functions of the cell. For this purpose, different RNA types have different functions. There are messenger RNA (mRNA), transfer RNA (tRNA), which function in the translation of mRNA to amino acid chains, ribosomal RNA (rRNA), which form the ribosome doing the transcription (Fig. 4), and many other types of RNA involved in RNA modification or gene regulation. However, most studies of RNA have only mRNA in mind. This compound is the result of transcription from DNA and the basis for the translation, which results in proteins. Thus, it contains the information of DNA but is at the same time involved in the actual reaction process of the organism. Only recently, short RNA (sRNA) have reached more of the focus of research as central regulators of gene activity regulation.

Given the higher stability of DNA, it makes better long-term information storage. RNA, however, is a better catalyst. A catalyst facilitates and enhances a chemical reaction without changing itself. These reactions can be aminoacylation during translation or splicing of DNA in the case of snRNA[8]. These reactions do not need to be in the cell in which the RNA has been produced since RNA can be transported long distance within a plant via the phloem[9]. They can even be transported between species,

e.g., between host and microbe[10], or even in exceptional cases, transmitted vertically between generations[11].

This position at the crossroads of vertical transmission of information between generations and the horizontal transmission of information between the DNA and the environment makes RNA a fascinating compound for understanding how an organism reacts to its environment and, at the same time, how species evolve in reaction to natural selection. This is why the field of transcriptomics, the study of mRNA variation, is a central part of plant genetics and evolutionary biology.

1.2 Advantages of studying RNA

Studying the transcriptome has several advantages over studying the genome when interested in natural selection. For example, in metabarcoding of microbial life (bacteria, archaea, fungi, etc.) RNA demonstrates the activity of an organism rather than just the presence of an organism[12]. Studying mRNA provides information on the temporal and spatial patterns of genetic activity[13,14]. It also offers the opportunity to study genetic differences between genotypes since no sequence variation is necessary in the coding region of the respective gene if the expression is regulated by the promoter region only or epigenetically. As a third possibility, variation in the transcripts in the absence of sequence variation is caused by alternative splicing[15].

The other types of RNA are also used in evolutionary biology but mostly as phylogenetic markers only (but see below). The reason for their use as phylogenetic markers is their ubiquity as central actor in the functioning of the cell. For example, rRNA and tRNA are essential for transcription. Any mutation changing their functioning is likely to be detrimental. Thus, their evolutionary change is slow, and they can be used to infer phylogenetic relationships across divergent organisms[16]. Also, tRNA are used in phylogenetic analyses, especially those present in the chloroplast genome.

In contrast to those RNA types involved in transcription, those involved in the post-transcriptional modification and gene regulation are much less studied. However, the importance of these has gained more and more acceptance. Among these, microRNA and small interfering RNA are involved in gene regulation of most eukaryotes, especially in some immune systems against viruses and self-replicating DNA and RNA such

as transposons. MicroRNA were first detected in 1993[17] and only gained interest in the new millennium. Today, more than 38.000 microRNAs are known (www.mirbase.org). They are 21–24 nucleotides long and have a hairpin structure. Their biogenesis and action have been described by Kidner et al.[18]. They regulate transcription factors and are primarily involved in small scale-limitation of gene action by spatial or temporal limitation of RNA distribution. In vertebrates and plants, more than 20.000 different microRNAs are known, and their number is suggested to be correlated with organismic complexity[19]. Taylor et al.[20] indicated that bursts in the diversification of microRNA are correlated with the phenotypic evolution of plants. Given the importance of microRNA in the stress response of plants[21], the study of microRNA is certainly gaining increased attention in the future of plant breeding and plant evolution.

1.3 Regulation of DNA transcription

From the DNA in the genome to the functional protein, eight steps are required and can be regulated (Fig. 5). These steps are 1) remodeling the chromatin, the structure of the chromosome, 2) control of transcription, regulating that RNA is formed at all, 3) control of processing, the process by which mRNA is formed from the initially transcribed RNA, 4) control of transport, regulating which RNA leaves the nucleus, 5) control by the stability of RNA, the lifetime of RNA, 6) control of translation of RNA to protein, 7) post-translational changes of enzymes, and 8) protein degradation. Only the first five steps affect the transcriptome, the amount and diversity of mRNA in a cell or tissue. Nevertheless, the other steps are also important for the formation of the phenotype, because these steps lead to the fact that the amount of an enzyme does not always correlate with the amount of RNA, which is often ignored but essential to understand transcriptomics. Translational repression of protein formation from mRNA may, for example, mislead to inferences of activity of biosynthetic pathways that are, however, down-regulated at the protein level. Thus, in case of conflict between transcriptomic and phenotypic data, proteomic approaches may be necessary for conclusive answers.

With regards to transcriptomics, especially important are the transcription factors influencing whether or not a gene is transcribed and how much is transcribed. There are enhancers and repressors. These

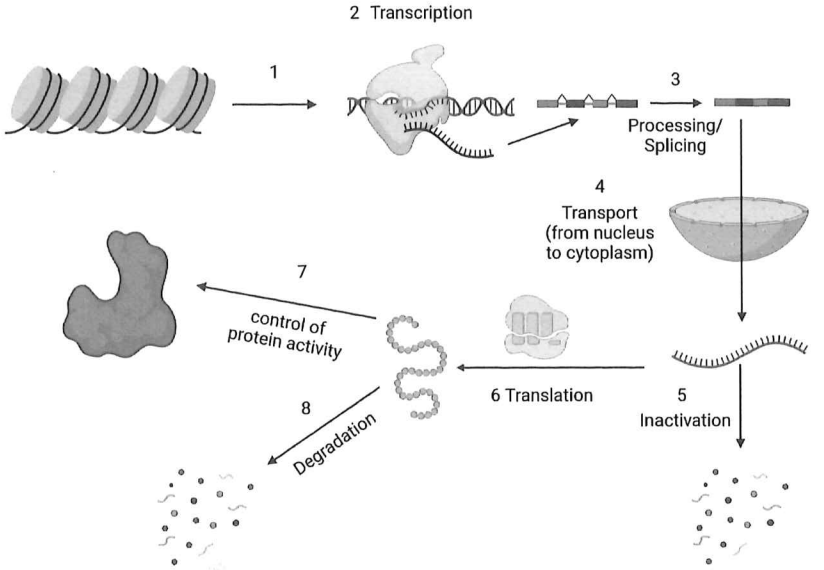

Fig. 5 Steps of transcriptional regulation (generated using BioRender)

transcription factors can be expressed constitutively but also organ-specific, controlled by environmental factors, or in a developmental stage- or cell cycle-specific manner[22]. Environmental stimuli can be light, hypoxia, water, temperature, salinity, or biotic interactions (e.g., bacteria or fungi invading the plant or bites by herbivores). These environmental stimuli are perceived mostly as oxidative stress by an increased amount of ROS (reactive oxygen species) or RNS (reactive nitrogen species) or ionic stress by changes in minerals such as calcium ions. However, there are also enzymes able to perceive external stimuli, for example phytochromes for light and heat shock proteins for heat[23]. These signal perceptions lead to a signaling pathway, which involves hormones (e.g., gibberellins, auxins, abscisic acid, jasmonates, etc.) and sRNA. Hormones are essential factors regulating transcription via master regulators, whereas sRNA regulate gene activity via post-transcriptional mechanisms. However, they can also regulate gene activity at the transcriptional level indirectly by regulating hormone production post-transcriptionally.

The result is an environment-dependent but tissue-specific amount of mRNA for any gene. Since some of these processes occur on a short time

scale, sometimes within minutes, and often depend on prior environmental stimuli, transcriptomic studies can significantly vary in their results. Even when comparing clones or similar tissues within one individual, excluding (mostly) genetic factors, epigenetic variation may still cause variation. This blurring of significant signals in the data by uncontrollable variation may cause frustration among scientists but is part of the wonder of diversity in life. Living organisms remain different from totally predictable machines.