The *Working* Genome is RNA - Evolution of a Conceptual Challenge

Klaus Scherrer, Paris (klaus.scherrer31@gmail.com)

ABSTRACT

The genome is encoded in DNA but, at least in eukaryotic organisms, the *Working Genome* is RNA. The reason is simple chemistry which is *the basis of life:* encoding, transmitting and expressing information. DNA is incredibly stable; genomes a million years old have been partially recovered and sequenced. In contrast RNA is labile and thus easily modified by chemical reactions. Nevertheless, RNA can be stabilised and saved for years. Maternal mRNA in human oocytes is stored for at least 30 years, and some high Mr RNAs are even carried over in germ-line cells from generation to generation. During cell differentiation, partially processed pre-mRNAs are stabilised in time and stored in the nucleus along the *Cascade of Regulation*. The first indication of this was the discovery in 1961 (DOI 10.1016/0006-291X(62)90341-8) of *high Mr RNA and RNA-processing*, as well as later on, of *pervasive transcription* in animal cells. Conceptually, the general acceptance and integration of these facts took decades, and even today there is debate about some of the basic principles. We will try to retrace here the evolution of these paradigms of gene expression which were challenged from the beginning. - Before about 1961 the largest cellular RNA known was the 28S ribosomal RNA (rRNA) in animal cells. The high Mr RNA we observed sedimenting at about 45S turned out to be the G-C rich metabolic precursor of rRNA, pre-rRNA, which looses half of its sequence while being processed into 28S and 18S rRNA. This was the *first observation of RNA processing*, a mechanism, as discovered later, operating on most eukaryotic primary gene transcripts. - But there was also some much larger RNA sedimenting faster than the 45S pre-rRNA and this turned out to be A-U rich RNA, hence "DNA-like". While the existence and role of pre-rRNA seemed clear from the beginning, the A-U rich "giant" RNA was considered to be an artefact. Confirmation came in 1973 when we developed an original method of RNA spreading and obtained EM-pictures which revealed the real size of 28S and 18S rRNAs, the 45S pre-rRNA as well as demonstrating the existence of much larger molecules in the A-U-rich RNA fraction. Thereafter the existence of the giant RNA molecules became accepted and were shown later to be *pre-messenger RNA* (pre-mRNA) prior to being processed into mRNA. - Another surprising observation of the 1960s was *pervasive transcription*: red blood cells in which globin represents about 90% of protein synthesis, were found to transcribe thousands of genes. Indeed, according to latest data, the *whole genome is transcribed* in the lifetime of an organism. Thus, DNA serves just as the archive of the genome, and its expression and regulation operate essentially through transcription, RNA-processing and mRNA translation (c.f. 2018 review in DOI 10.1016/j.yexcr.2018.09.011). Thus the *Working Genome* acts at the RNA-level. Note that during differentiation, the 3D-organisation of DNA can be modified, prior to transcription, leading to new kinds of pre-mRNAs. - If the early data were dismissed, even worse was the rejection at the conceptual level of ideas based on early biochemical data. Full domain transcription into primary high Mr RNA has only recently been accepted as the basis of expression of protein-coding genes thanks to the recent method of "amplification-free long range sequencing" of RNA molecules. - In the sixties, only 3 laboratories interpreted correctly the emerging data showing genomic transcription into high Mr primary transcripts and its subsequent processing into mRNA; 99% of labs rejected the data and the new ideas. Confirmation of the few facing a large majority almost looks like a miracle !