Coffee with Ribohipster

MANUEL ARES Jr.
Molecular, Cell and Developmental Biology, University of California, Santa Cruz, California 95064, USA

Back in '95, things were different, not like they are now. To help me recall the last two decades, I went out for coffee with a local Santa Cruz eccentric known as Ribohipster. He has a decided preference for artisanal science in an era of corporatized big government projects. Not one to consort with consortia, he resists the idea of trading a solid finding brought forth with one's own hands for a wheelbarrow of government-purchased, industrially manufactured datinos (10^3 datinos = one piece of data). He is wary of giving up too much of the art. I can see that his nostalgic fondness for vintage RNA research is going to color these reflections.

Back in '95 my lab had several things on its mind. Rhonda Perriman rigged our inside-out group I intron that made RNA circles to make an infinite open reading frame mRNA. *Escherichia coli* ribosomes circumnavigated this RNA to make very long (>800 kDa) proteins. We had dreams of spinning out new strong and long proteins. Circles were a curiosity back then, just like now. Several other groups also made RNA circles (Michael Been, Kevin Jarrell, Mariano Garcia-Blanco) in vitro and in vivo. Peter Zaphiropoulos convincingly showed that natural circles are formed by the spliceosome. Julia Salzman's and others analysis of RNAseq data to look for circles transcriptome-wide will teach us the extent to which circles play important functional roles. Ribohipster became a little agitated and went on a bit of a rant. He likes Julia's paper but is generally irritated by the bandwagon hype when a new technique finds things that have already been discovered, and is turned off by glam journal-induced false priority claims. Nobody reads the old papers he wails. But these new methods are more comprehensive, I argue, they cover the whole genome. It's mass produced and descriptive he replies, as he orders another coffee.

Back in '95, we didn't have any complete eukaryotic genomes. Once in a while, a complete yeast chromosome would come out. I had accidentally discovered yeast U2 while a postdoc with Alan Weiner at Yale, found it was ~1.2 kb (too big for a "small" nuclear RNA), sequenced the gene and showed it was essential. Stephanie Ruby, rest her soul, had told me that U2 was linked to *PRP5*, known to be on chromosome II, but when the chromosome II paper came out it made no mention of U2. The sequence was there, but I guess RNA wasn't as hip in those days. I remark that nowadays we fall over ourselves to annotate every long noncoding RNA we can imagine. Ribohipster finds this hilarious and snorts coffee out his nose laughing. I move on, explaining how I heard from Micha Sammeth in mid-2013, while he was struggling to map RNAseq reads for human U2 transcripts. He must have found my name on a GenBank file with the human U2 gene sequence. I told him that the true U2 genes had so far not appeared in any versions of the "complete" human genome. Instead, scattered pseudogenes for U2 were labeled as genes. Finally, by the end of 2013, we got a human genome (GRCh38/hg38) that displays the true U2 genes (RNU2-1) correctly. Ribohipster notes that it only took 13 years to get that right. He wonders what else is wrong with the genome.

Back in '95 we were sequencing a lot, with ^{32}P and gels. If you found a gene in those days it was for sure you had to sequence it. We did a genetic screen for cold-sensitive U2 suppressors or *CUS* genes. Megan Neville cloned *CUS1* and *CUS2*, and Haller Igel became the sequencing expert, doing BAL31 resections and primer-walking into the rescuing clones. Ribohipster laughs, wondering aloud if anyone knows what those terms mean any more. We'd pair up to read the autorads into the MacIntosh IIsi (with a color cathode ray tube monitor!), one person reading up the gel and the other typing, and emailing it to NCBI's blast server. The database was growing so fast, you had to do this every day to make sure you didn't miss anything. While sequencing *CUS1* we got a very strong tblastx hit to *E. coli* RNase III. It turned out the RNaseIII gene (*RNT1*) was next door to *CUS1*. Postdoc Sherif Abou Elela did the honors following that up. Robin Reed helped us figure out that Cus1 was yeast SF3B2, then known as SAP145 or SF3b145. As we did our daily blast emails, yeast SF3B4 appeared (*HSH49*, characterized by Haller Igel) and then yeast SF3B1 (*HSH155*, taken on by Michelle Pauling). Rhonda Perriman decided to examine splicing extracts lacking Cus2. The extracts worked fine, but she couldn't seem to deplete them of ATP; they just went right ahead without it. After many controls and some demanding chase experiments, she showed that Cus2 is responsible for enforcing the ATP-dependent step in U2 binding to

Corresponding author: ares@ucsc.edu
Article and publication date are at http://www.rnajournal.org/cgi/doi/10.1261/rna.050468.115. Freely available online through the *RNA* Open Access option.

the pre-mRNA. With help from Gia Voeltz in the lab and Imre Barta from Abelson's group, Rhonda found that Cus2 antagonizes Prp5's function with U2. She produced a parade of papers culminating in the discovery of the branchpoint interacting stem loop (BSL) in U2. Recently SF3B1 has attracted new attention as both a target of spliceostatin and for its stereotypic mutations during progression of certain leukemias. Ribohipster grunts at this and smugly points out that fundamental research initiated for no other reason than to know how RNA works is rarely unimportant.

Back in '95 as the yeast sequence was accumulating we got interested in knowing where all the introns were going to be in the yeast genome. Leslie Grate searched sequence for intron patterns and Carrie Davis looked by RT-PCR and sequencing. By 1998 we had a set of about 250, but we knew there must be more. Tyson Clark's rotation project was to compare wild type RNA debranching mutant (*dbr1*) RNA, which Jef Boeke's lab had shown was loaded with shocking amounts of undegraded intron. Back then one could buy nylon membranes spotted with Maynard Olson's ordered yeast genomic fragments in lambda clones. Tyson hybridized these "macroarrays" looking for spots that lit up more with the *dbr1* probe than wild type. He got signal over clones without known introns, but we never guessed right with our primers. Tyson was hooked, but he wasn't going to stand for the primitive approach and wanted to make splicing sensitive microarrays. We had the right design, using junction and exon oligos to resolve signals from alternatively spliced RNA from the same gene, but we still had no microarray printer and no method to stick oligonucleotides to slides. Tyson and Chuck Sugnet teamed up to build our printer using Joe DeRisi's instructions from the Internet. Ribohipster smiles at the idea of home made microarrays. Lily Shiue's experience brought critical knowledge on printing methods. Yeast cDNA libraries were nonexistent, but from Carrie's validations we knew where the splice junctions were. Lily, Chuck, and Tyson printed our first test arrays, employing oligos for the two-intron yeast *SLC1* gene. We made *SLC1* splice site mutant strains that produced mixtures of alternatively spliced RNAs. The arrays worked like a shot. Tyson and Chuck's first paper came out in 2002 and showed the surprising result that deletion of different conserved splicing factors produces very different splicing phenotypes, helping explain why different characteristic spliceosome mutations track with retinitis pigmentosum and leukemias. We gave away many yeast arrays and helped folks like Kathy Gould, Scott Stevens, Grant Hartzog (who helped us with printing), T.-H. Chang, Stefan Jentsch, David Horowitz, and Tracy Johnson. Jeff Pleiss from the Guthrie lab wanted to make his own so he came to learn and went home with printing plates of our oligos. We made smaller arrays for human and mouse designed with help from Doug Black, R.-J. Lin, and Xiang-Dong Fu. These first arrays were very powerful because there was no easier way to capture the responses of hundreds of alternative splicing events in a single experiment. Ribohipster says he senses that I am happy we made these ourselves. I reply that we did ok with the yeast genome but to fully capture mammalian splicing we needed to go to a company. He frowns and pulls at the waxed ends of his mustache, but I continue with the story.

I was on David Kulp's thesis committee back when he was working with David Haussler on eukaryotic gene-finding programs, trying to predict introns from raw sequence. Kulp left to help start Neomorphic, which was later bought by Affymetrix, where he ended up. He invited Tyson, Chuck, Lily and me to help design Affy's first genome-wide splicing array for mouse called the "A-chip." Melissa Cline joined us and became critical to nearly everyone using the A-chip, including Bob Darnell who got early access to these arrays and made good use of them. Our first experiment with the A-chip was to dissect mice and make RNA to identify tissue-specific exons. Grad student Shalu Srinivasan knew her way around a mouse and ran arrays on every tissue she could reasonably dissect. Chuck (by now a graduate student with Haussler) and Melissa developed analysis methods for finding splicing differences. Chuck identified a muscle splicing-associated motif we later showed regulates splicing by *quaking* protein. Julie Ni used the A-chip to detect alternatively spliced NMD targets and found widespread autogenous control of RNA binding protein mRNA levels mediated by ultraconserved sequences. Hongqing Du in collaboration with Maury Swanson's and Charles Thornton's labs revealed that the transcriptome-wide impact of CUG repeat RNA mimicked loss of MBNL1 splicing activity in muscles of mouse models of myotonic dystrophy. And Stephanie Huelga in Gene Yeo's lab put together an awesome set of studies on human hnRNP proteins. Ribohipster grudgingly admits that it would have been hard to do those experiments with homemade arrays.

By 2012, RNAseq methods pushed microarrays to zombie platform status, primarily because only known events can be captured on an array. I explain to Ribohipster that we no longer make or use microarrays. Lisa Munding used RNAseq to uncover the pre-mRNA competition phenomenon in yeast. Shutting down ribosomal protein gene transcription increases splicing efficiency for other pre-mRNAs. She also found previously unrecognized introns in strange places, whose splicing is out-competed in growing cells. We would never have seen this with arrays. He complains in his grumpy voice that RNAseq is the height of corporatized science with its kits and ridiculously expensive machines. How can the artisan survive? I laugh and tell him it is ok because ENCODE will do all the experiments from now on, and we will only need to think about their data. He doesn't appreciate the joke. I explain that what matters more is the idea. Ribohipster wants to see what people do with RNAseq data once the obvious things like "discovering" exons that aren't in RefSeq are over.

Ribohipster still worries that distributing big money to poorly conceived moon-shot projects and so-called centers of excellence is going to stifle "small farmer" scientists. He thinks that great ideas get you through times of no money

better than big money gets you through times of no ideas. As we bus our table and leave the coffee shop I ask him if there is anything new he thinks is good. CRISPRs, he says, it's cheap to get into, works pretty well in a variety of systems, and the results will illuminate mechanism. I ask if he is sure it's not just another trendy bandwagon, and he rolls his eyes. We part vowing to stay authentic about our processes and to call out things that are poorly done. We'll meet in 2035 to see how it's going. He hops on his unicycle and pedals off down the street.

Acknowledgments

Work in the Ares lab over the last 20 years has been supported by grants GM040478 and GM084317.