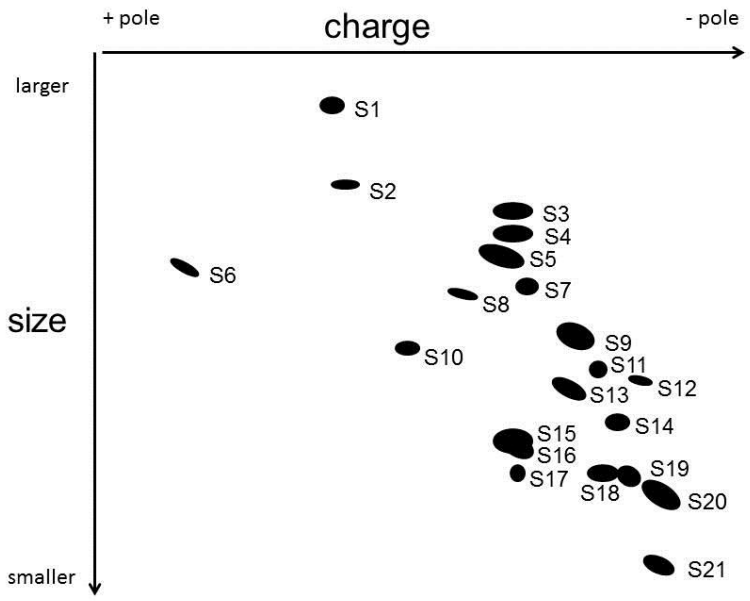
Here lies one whose name is writ in water

Or so says the headstone on the grave of John Keats in a cemetery in Rome. These words bring to mind the fleeting image of a name, somehow created on the water's surface only to disappear as Brownian motion and diffusion conspire to cause the name to dissipate. Imagine if you will, this name disappearing and then reappearing as another. The second law of thermodynamics would tend to argue against this possibility - systems spontaneously move to maximize entropy (disorder) - to reverse this process and create order from chaos, energy much be added to a system.

With this in mind, let us turn to this month's *Journal Club*. The chaos in question is a nomenclature for ribosomal proteins that has been cobbled together over a 40+-year span creating considerable confusion and frustration for individuals working in the field. The energy to reverse this trend and make order from chaos comes from a group of ribosomologists that spend their days puzzling over the three-dimensional structures of these enormously complex particles. Their studies began with ribosomes from bacteria found in exotic locations, like thermal vents in the sea, because these ribosomes were more stable to bombardment by the X-rays used to solve their complex structures. They then turned their gaze to ribosomes from more garden variety microbes like *Escherichia coli* and have more recently turned their attention to ribosomes from organisms like us. The problem for the structural biochemists is that the names used for ribosomal proteins in the different systems described above can differ even though the ribosomal proteins from one system to the next can be related; having common structures and functions. The structural biochemists have decided that nomenclature for ribosomal proteins that has arisen over the years is untenable and must be changed to a more uniform system allowing ready comparisons of ribosomal proteins from one system to the next. Most others working in the ribosome field agree to this change, although it is generally realized that this transition will not be without a certain amount of discomfort and anguish from interested parties.

One such interested party is the DBA community, particularly those patients with mutations in the eleven ribosomal proteins genes that have so far been shown to cause DBA. **The names for each of these genes will change. In some cases this change will be rather subtle. In others, the change will be more dramatic.** Before getting into these specific changes, I thought I would first provide a little background as to how we got to where we are today and the need for this change.

The first detailed studies on the structure and composition of ribosomes came in the early days of molecular biology back in the 1960's. Ribosomes were shown to be ribonucleoprotein complexes composed of two subunits of unequal size. The small ribosomal subunit was shown to be composed of a single RNA and 21 proteins. These proteins could be separated by two-dimensional gel electrophoresis with the first dimension separating according to net charge (from right to left in the image below) and the second dimension separating according to size (decreasing size from top to bottom). Separation using this technique gave a pattern similar to that shown in the figure below.



Proteins were then named for how they separated on these two-dimensional gels starting from the largest protein at the top of the gel and ending with the smallest protein at the lower right. If two proteins had the same size, they were named based on their separation on the basis of charge going from left to right; such as can be seen for S17, S18 and S19 in the image above. The ribosomal proteins were given the designation S because they came from the small ribosomal subunit. A similar nomenclature was developed for ribosomal proteins of the large subunit, which were given the designation L.

This system seemed fine and during the 70's was applied to new studies being performed on mammalian ribosomes, again naming ribosomal proteins based on their migration in two-dimensional gels. Mammalian ribosomes were larger and had more ribosomal proteins than those from bacteria and so the numbering system extended from S21 to S31. About this time researchers were determining the primary structures of ribosomal proteins and the genes that encode them. Proteins are polymers of amino acids and the primary structure of a protein is the order of amino acids along its chain length. The order of amino acids in a protein can be determined either directly from the protein or deduced from the sequence of DNA which encodes it. One of the mantas I use in teaching elements of protein structure to medical students is when discussing proteins "structure gives rise to function and common structures give rise to common functions."

What we began to learn from these sequences is that some ribosomal proteins in organisms as distinct as *E. coli* and humans had common structures and therefore common functions, most likely because they arose from some common ancestral ribosomal protein at some point in evolution. However, as we began to know more about the primary structures of the different ribosomal proteins and their relatedness, the system of nomenclature used to name them began to crumble. For example, S19 in the bacterial system does not have a counterpart in humans. Thus,

the protein termed S19 in humans doesn't look anything like S19 in bacteria. Apparently, human S19 was a relatively late addition to ribosomes arising after bacterial cells took their own evolutionary path distinct from the path taken by primitive cells with nuclei (eukaryotes) came along and began working together to give, well, us. The upshot of all this is that it is very difficult to keep track of who's who when two totally unrelated proteins are given the exact same name. This confusion is resolved rather easily in the new nomenclature where the family of ribosomal proteins related to bacterial S19 is given the designation S19; whereas the human S19 protein family is given the designation S19e, with the "e" standing for eukaryote-only. You would recognize this protein with the more proper and complete name RPS19e. Fortunately, most of the known DBA ribosomal protein genes are classified as being eukaryotic-specific and so will only change by adding an "e" at the end of their current name.

Unfortunately, this simply solution isn't going to work for all known DBA genes. Take for example the protein, L5 (A.K.A. RPL5). Human L5 is actually structurally related to L18 in bacteria (so they would be merged to a single protein family given the new designation L18), and the icing on the cake is that human RPL11, another DBA protein, is actually related to L5 in bacteria. Thus, the newly proposed nomenclature would change human RPL5 to RPL18 and RPL11 to RPL5, the latter being exceptionally confusing as both names are used now for DBA genes. Sigh! One final quirk in the new nomenclature is that some of the human ribosomal proteins given a number and a letter, for example the DBA gene RPL35A. The letter A will no longer be used and they will instead be named after their yeast (another eukaryote) counterparts. So RPL35A will now be named after its eukaryote-specific yeast family member, RPL33e.

Should we stand up and fight against the tyranny of the structural biochemists? Alas, I think not. The new nomenclature does in fact make a great deal of sense and will greatly simplify the lives of individuals trying to identify new DBA genes. When we identify a new candidate DBA gene based on a polymorphism in a ribosomal protein gene, we often have to draw upon structural studies on bacterial counterparts to determine whether a mutation is pathogenic or not. With the willy-nilly nomenclature we currently have, when we cross boundaries between humans and bacteria it often takes a while for us to gather or wits and gain our bearings to make sure that the two proteins we are comparing are actually counterparts of one another; thereby making the comparison valid.

So how will this all play out, particularly in the transitional period before we all get used to the new system of nomenclature? **First of all let me point out, that for those of you where the name change will be most dramatic, for example the human RPL11 gene going to RPL5…..nothing has changed, except the name of the gene. Your child's gene and mutation therein hasn't changed, it has just been given a different name. This may take a little adjustment for you and your physician, but in order for you to be able to keep up on the scientific literature on DBA which should be using the new nomenclature, this adjustment will be necessary.** To ease the transition, it has been proposed that those of us publishing in the field use both the new and the old name of the proteins we are describing (a solution raised to the ribosome community and supported by none other than our very own Dr. Jeffrey Lipton). For example, RPL11(RPL5), new(old). This will be clumsy, but it will hopefully only be in place for a few years before we completely shift to using the new nomenclature.

As another aide during this transitional period I have created a cheat sheet below providing readers with the current name of known DBA genes and what their new names will be using the new nomenclature.

Human Name	Bacteria Name	New Name
RPS7		RPS7e
RPS10		RPS10e
RPS17		RPS17e
RPS19		RPS19e
RPS24		RPS24e
RPS26		RPS26e
RPL5	L18	RPL18
RPL11	L5	RPL5
RPL15		RPL15e
RPL26	L24	RPL24
RPL35A		RPL33e

The structural biochemists rationale for this nomenclature change and a complete listing of all ribosomal proteins under the new system of nomenclature can be found at the link below.

http://www.elsevierblogs.com/currentcomments/?p=686