

# Duke University

DURHAM  
NORTH CAROLINA  
27708-0338

DEPARTMENT OF BIOLOGY  
BOX 90338

TELEPHONE (919) 660-7372  
FAX (919) 660-7293

919-660-7342  
[dmshea@duke.edu](mailto:dmshea@duke.edu)

1 November 2005

Dr. Jeremy Bennett  
Faculty Search Coordinator  
Department of Biology  
Indiana University  
1001 East 3rd Street  
Jordan Hall 127  
Bloomington, IN 47405-3700

Dear Dr. Bennett:

I'm writing in support of Dr. Jeff Marcus's application for a position in your department. I knew Jeff when he was a graduate student here at Duke, but we interacted mainly on a professional level, on account of our joint interest in the complexity of organisms.

The conclusion of this letter is that Jeff is one of the most impressive young scientists I know. But to explain, I first need to explain my understanding of an old problem in molecular biology, the C-value paradox, dating back I believe to the 1960s. The very reasonable hypothesis was raised that number of genes in an organism should be correlated with the complexity of the organism, in particular its phenotypic complexity or complexity of form. The thinking was that a more complex organism ought to require more instructions for its development, and therefore more genes. This was before any robust estimates could be made for numbers of genes in any organism. But we could measure total amount of DNA, the "C value," and this was construed as a kind of first-order estimate of gene number. The difficulty remained was that there was no objective measure of complexity of form, but the *scala natura*, or Great Chain of Being, was deemed by some a suitable proxy. Humans were assumed to be the most complex morphologically, followed by other mammals, then reptiles, amphibians, fish, and so on. The results, however, did not support the hypothesis. The correlation between amount of DNA and Great-Chain complexity was poor. Notably, certain "lower" forms, including some amphibians, have huge amounts of DNA compared to "higher" forms like us. The hypothesis could have been simply rejected, but my guess is that its underlying logic was considered so compelling that it remained on the table, so to speak, and the poor correlation was deemed anomalous. Hence, the

C-value paradox.

In recent years, of course, we have been able to produce some very robust estimates of gene number, based on numbers of open-reading frames, rather than raw amounts of DNA. And some of the anomalous results have been explained. (My understanding is that the high-C-value amphibians, for example, have many DNA copies but not extraordinary numbers of genes.) But a real barrier remains to a serious test of the hypothesis relating gene number and complexity of form. And it is (still) that we have had no objective measures of complexity of form. My own work has focused on developing such measures, both in principle and in practice, and I have devised a number of them, mostly applicable to multicellular organisms. In effect, I've devised ways to estimate numbers of parts in these organisms.

While Jeff was still here at Duke, he devised a way to apply the general part-counting principles I had developed to bacteria and protists. When he first approached me about this project I was fairly skeptical. It wasn't clear to me that the morphology of most bacteria was complex enough – at least insofar as it could be discerned with modern microscopy – to make counting parts possible with my methods. We talked about it at some length, and I agreed to help him where I could, but I was still fairly skeptical. It's now clear that he has done it. And by that I mean not just that he has come up with quantitative morphological complexity estimates for bacteria, but that he has demonstrated the biological meaningfulness of those estimates by showing a statistically significant correlation with gene number. In effect, Marcus has solved the decades-old C-value paradox. He was solved it in principle but showing that direct, objective measures of morphological complexity can be devised for bacteria. And he has begun to solve it in fact, showing that the predicted correlation in fact exists in the prokaryotes and protists for which we have reliable gene counts.

I have told this story at such length because I think it reveals so much about Jeff. First, his scientific style is bold. The obvious near-intractability of the problem seemed not to have daunted him in the least. He was willing – make that eager – to take it on, and to stick with it, despite the obvious high risk of coming up completely empty. More impressive yet, he did this at a point in his career when most shy away from risky research. This is a gutsy scientist. Second, he showed enormous creativity as well. The application of my parts-counting principles to bacteria was not straightforward, and indeed was something I would have bet a modest sum could not be done. But he did it, and on his own.

I will leave direct comment on the rest of Jeff's work to others, adding here just that it seems to me to show the same extraordinary high level of competence and creativity he showed in the C-value project. He is, in my view, one of the top two or three students to graduate from Duke in the past 5 years or more. As a scientist, he is simply first-rate.

Sincerely,

Daniel W. McShea  
Associate Professor