Comment

Count me out

Gregory A Petsko

Address: Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, MA 02454-9110, USA.
E-mail: petsko@brandeis.edu

Published: 8 December 2000

Genome ***Biology*** 2000, **1(6)**:comment1006.1–1006.2

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2000/1/6/comment/1006

The following is a transcript of testimony before the Supreme Court of the State of Florida, United States of America. Eugene Finder, lawyer on behalf of the Human Genome Organization, is arguing his case before the assembled justices, who are questioning him.

Gene Finder: ...and so, Your Honors, I hope that the evidence will convince you that if the decision of the Secretary of State to certify the count is upheld, the number of genes in the human genome will be smaller than anyone ever imagined.

Justice Skeptical: Just how much smaller?

Gene Finder: We don't know yet. That's why we're asking you to delay certifying the official count until we're sure what the correct number is.

Justice Ambivalent: What do you mean you don't know yet? The draft human genome sequence has been available for months. The automated counting of genes has been going on since then. How hard is it to count something properly?

Gene Finder: Uh, well, Your Honors, a large part of the problem is that we don't exactly know what constitutes a legitimate gene. We can identify sequences of DNA that have long open reading frames and things that look like start and stop codons, but we don't know if they're actually transcribed. Some sequences have several apparent start codons embedded in the first couple of exons; we don't know which, if any, is the right one.

Justice Credulous: Can't you just go by the distance to the promoter region? If there's a nearby promoter, doesn't that signify intent on the part of the sequence to be a transcribed gene?

Gene Finder: It might, Your Honor, if we knew what a good promoter looked like in a higher eukaryote, but there are arguments about that. Does a gene need a complete promoter to be a functional gene, or would an incomplete promoter – what we're starting to call 'dimpled promoters' – be enough to signify intent? We also don't know how far away from the start signal a promoter can be in a higher eukaryote and still work. Which brings me to the intron problem.

Justice Credulous: Intron problem? Does that have something to do with Palm Beach County?

Gene Finder: Uh, no, Your Honor; that's a different case. Introns are intervening sequences. We used to think introns didn't code for anything. But some genes have things that look like good start signals in introns, and we now know of several genes that are intron-encoded. We propose, by the way, that such a cryptic start signal should be called a 'chad'.

Justice Ambivalent: You mean genes within genes? That would complicate the machine-based counting of genes, wouldn't it?

Gene Finder: Yes, Your Honor. That is why my client, who has bet a lot of money on there being more than 40,000 genes in the human genome, wants a manual recount, including such disputed genes. We believe that until a universally-agreed automatic method is developed, careful scrutiny of each potential gene by a human being is the right way to go.

Justice Credulous: Wait a minute, can't we just use homology to expressed genes from lower organisms as a criterion?

Gene Finder: We could use it, and we are doing so, but we don't know what weight to give it in drawing conclusions. Just because there is a homolog in a lower organism doesn't mean the gene is functional in humans. It may be a pseudogene. We anticipate that many apparent genes will have to be excluded from the final count because of irregularities such as non-functional promoters, cryptic stop signals, and so on.

Justice Credulous: So if you did that, could we at least count - sorry for the pun - on being able to come up with a total?

Gene Finder: Not in time. We have to delay the certification. If we certified now, on the basis of the little we know about what it takes to make a human gene, the number of genes in the human genome would be less than 40,000. Not only would my client lose her bet, but the figure would mean that we as human beings are only three times more complicated than a fruitfly, in genetic terms. I think that would be very hard for people to accept.

Justice Ambivalent: Well, besides including genes with dimpled promoters and genes where the intent of the sequence is unclear, is there any other way to ensure that all genes have been fairly counted?

Gene Finder: Yes, Your Honor. My client believes that we must also learn how to take into account alternative splicing.

Justice Credulous: Is that like the situation in a ballot, where people punch two candidates instead of voting for only one?

Gene Finder: Very much like it, Your Honor. But alternative splicing can produce many more than two gene products from a single gene. For example, work by Fettiplace at Wisconsin and Fuchs at Johns Hopkins, among others, has shown that mechanosensory hair cells of the vertebrate inner ear contribute to acoustic tuning through feedback processes involving voltage-gated ion channels in the basolateral membrane and mechanotransduction channels in the apical hair bundle. The specific number and kinetics of calcium-activated (BK) potassium channels determine the resonant frequency of electrically tuned hair cells. Kinetic variation among BK channels appears to arise through alternative splicing of mRNA for the *slo* gene and combination with modulatory beta subunits. It's hard to know how many splice variants of these channels may be involved, but it is possible that a single gene may produce dozens or even hundreds of slightly different proteins in the auditory system. And who knows how often that sort of thing happens in other tissues? We know that many genes, especially genes encoding receptors, are alternatively spliced. Do they count as one gene or two?

Justice Skeptical: I don't think you should count your chickens before ...

Gene Finder: Exactly, Your Honor, thank you. Some of these variants were first detected in chickens. But it gets worse, Your Honors. We don't even know how to predict from the sequence whether a gene will be alternatively spliced, or if so, how many different splice variants it will produce. Not enough people are working on that problem. And if we could do that, when we have a gene that produces, say, six alternative transcripts, do we count it as one gene or many? We need a ruling from Your Honors on this matter before we can go forward and certify a final gene count.

Justices: Don't count on it.