

PublisherInfo		
PublisherName	:	BioMed Central
PublisherLocation	:	London
PublisherImprintName	:	BioMed Central

Further resolution of gene transfer issues

ArticleInfo		
ArticleID	:	3934
ArticleDOI	:	10.1186/gb-2001-2-8-reports0027
ArticleCitationID	:	reports0027
ArticleSequenceNumber	:	25
ArticleCategory	:	Paper report
ArticleFirstPage	:	1
ArticleLastPage	:	4
ArticleHistory	:	RegistrationDate : 2001-6-27 Received : 2001-6-27 OnlineDate : 2001-8-8
ArticleCopyright	:	BioMed Central Ltd2001
ArticleGrants	:	

Reiner Veitia

Abstract

Genomic and phylogenetic analyses have suggested that previously reported cases of putative lateral gene transfer from bacteria to vertebrates could result from gene losses and sampling artifacts.

Significance and context

Horizontal gene transfer (HGT) is the transmission of genetic material among organisms that are not directly related by descent. It has been very well documented in bacteria, in which it can occur by a variety of mechanisms including transformation. The existence of homologous sequences in very distantly related taxa but not in phylogenetic intermediates is suggestive of HGT. For instance, analysis of the draft sequence of the human genome showed 113 genes that were similar to bacterial sequences but apparently lacked homologs in the non-vertebrate species considered (including *Caenorhabditis elegans* and *Arabidopsis thaliana*). These genes were considered as putative cases of HGT from bacteria to vertebrates. As pointed out by Salzberg *et al.*, such an event implies the transfer of bacterial DNA into the germ line of the vertebrate ancestor and stable incorporation of this genetic material into the host genome. Such an event is conceivable, but to spread throughout the population, the transferred genes should confer a selective advantage to the recipient or show selfish properties such as the ability to multiply or transpose. This amplification step is harder to imagine. It therefore became essential to assess the likelihood of HGT over such a wide evolutionary distance as bacteria to vertebrates. Salzberg *et al.* first compared the proteomes of the Ensembl and Celera databases against all publicly available complete bacterial genomic sequences. This provided a list of proteins homologous between humans and bacteria. These sets of proteins were then compared with non-vertebrate proteomes. The criterion for automatically ruling out HGT was the detection of non-vertebrate homologs.

Key results

In the comparison of bacterial-plus-human homologs against the non-vertebrate databases the authors noticed, as expected, that inclusion of additional, even unfinished, proteomes in the comparison rapidly reduced the number of putative transferred genes (to about 50-100, depending on the database used). The authors expect the final list to continue to shrink as further genomic sequence is taken into account in the analysis. They propose that in many instances the absence of non-vertebrate homologs could be due to gene loss, amplified by a sampling effect. Let us consider their scholarly yet elegant example.

Assume the last common vertebrate ancestor contained 10,000 genes and that each lineage can lose 30% of the genes. When comparing four non-vertebrate organisms, the probability that the same gene is lost in all four is $(0.3)^4$ (or nearly 81 genes for independent events). If one further considers that 20% of the proteome is essential and cannot be lost, 30% loss means 65 genes lost. This figure is of the same order of magnitude as the number of putative HGT events. The authors therefore consider that losses of genes inherited by vertical transmission can account for a large proportion of the putative HGT.

The human genome draft analysis presented by the Human Genome Program (HGP) consortium used the parameters of the BLAST report as suggestive (or not) of putative HGT events. In the present work, the authors 'play' with the BLAST statistical output to get a more accurate result. For instance, if in the BLAST search they use $E = 10^{-7}$ (lower stringency, including fewer similar sequences) instead of 10^{-10} (more stringent) as a cutoff to exclude homologs in non-vertebrates, they are able to cut down the putative HGT of the Ensembl dataset from 114 to 74 cases. That is, they would have missed homologs in non-vertebrates when using the stringent cutoff (as the HGP did). A phylogenetic analysis was also performed in the cases where sufficient numbers of homologous sequences were available. The results seem to rule out HGT. At the end of the genomic and evolutionary analysis the paper leaves us with some 40-50 putative cases of HGT. The authors prefer the hypothesis of gene loss to that of HGT, however.

Links

A complementary paper from another team of researchers that comes to a similar conclusion is reported in [Genome Biology 2\(8\):reports0026](#).

Reporter's comments

This paper shows comprehensibly a path of reasoning with which to test a biological genomic hypothesis. But the results of the authors' phylogenetic analysis should have been included (perhaps as supplementary material). The HGT 'candidacy' of the cases left at the end of the analysis must be re-evaluated in the near future when more non-vertebrate genomic sequences are available. As an indication of how fast this field moves, the formiminotransferase cyclodeamidase gene retained in the final list of putative HGT cases by Salzberg *et al.* will soon be removed because a homolog has been found in the slime mold *Dictyostelium*.

Table of links

[Science](#)

References

1. Salzberg SL, White O, Peterson S, Eisen J: Microbial genes in the human genome: lateral transfer or gene loss?. *Science*. 2001, 292: 1903-1906. 0036-8075