The HAR1F gene: a Darwinian paradox

Peer Terborg and Royal Truman

Unexpected scientific outliers are always interesting, as they offer the opportunity to either identify unsuspected causal factors or to discredit cherished theories. With this in mind, some human genes show such marked dissimilarities to those of chimpanzees that they invite careful reflection. Perhaps the data conforms well with a designed cause. Alternatively, various evolutionary explanations may be invoked. Might one interpretative framework be more plausible than the other?

According to a recent report in *Nature*, *non-random or ordered mutations* can be accepted as part of an evolutionary framework. This sounds suspiciously like *post facto* rationalization and is remarkable since *streng verboten* teleological implications quickly come to mind. The article speculates on why human brains are so distinct from the brains of chimpanzees. Scientists at the Center for Biomolecular Science and Engineering at the University of California, Santa Cruz, believe they may have found a key gene, *HAR1F*, which helped the human brain evolve from that of putative chimp-like ancestors.[1] The Associated Press offered[2] a tantalizing overview:

'Human brains are triple the size of chimp brains.

'Looking at 49 areas that have changed the most between the human and chimpanzee genomes, Haussler zeroed in on an area with "a very dramatic change in a relatively short period of time".

'That one gene [known as *HAR1F*] didn't exist until 300 million years ago and is present only in mammals and birds, not fish or animals without backbones.

'But then it didn't change much at all. There are only two differences in that one gene between a chimp and a chicken, Haussler said.

'But there are 18 differences in that one gene between human and chimp and they all seemed to occur in the development of man, he said.

'Andrew Clark, a Cornell University professor molecular biology who was not part of Haussler's team, said that if true, the change in genes would be the fastest and most dramatic in humans and would be "terrifically exciting". However, the gene changed so fast that Clark said that he has a hard time believing it unless something unusual happened in a mutation. It's not part of normal evolution, he said. Haussler attributed the dramatic change to the stress of man getting out of trees and walking on two feet.

'And it's not just that this gene changed a lot. There is also its involvement with the cerebral cortex, which is responsible for some of the more complex brain functions, including language and information processing.

'"It looks like in fact it is important in the development of brain [*sic*]," said co-author Sofie Salama, a research biologist at Santa Cruz who led the efforts to identify where the gene is active in the body.

'The scientists still don't know specifically what the gene does. But they know that this same gene turns on in human fetuses at seven weeks after conception and then shuts down at 19 weeks, Haussler said.'

Human accelerated regions

The *HAR1F* gene was first identified after comparing the human genome with the genomes of the chimpanzee, the mouse and the rat. DNA segments present only in the human genes and showing dramatic sequence differences are believed to be fast-evolving areas (table 1) and have been dubbed *human accelerated regions*—abbreviated HARs.

'… HARs are often associated with regions that undergo a high rate of recombination—the process by which an offspring obtains a blend of parental genes. Recombination, and its associated process, biased gene conversion, are thought to favour the inclusion of G and C nucleotides over the other two possible nucleotides, A and T … . As all of the nucleotide substitutions observed in *HAR1* are of this type, high (and biased) mutation rates might explain part of the rapid evolution of *HAR1*. Nevertheless, this process cannot explain the authors' other observations, such as the pairs of substitutions that together further stabilize the structure of *HAR1* RNA.'[3]

The *HAR1F* gene is part of a newly discovered RNA gene expressed by a particular type of brain cells (called Cajal-Retzius cells), and it regulates how the six layers of the cortex are laid down during the development of the human embryo.

Analysis

The existence of the *HAR1F* gene in chimps and chickens places constraints on evolutionary interpretations. Independent convergence can hardly be expected to produce almost identical nucleotide sequences. This leaves as the best alternative a common ancestor which preceded by eons the origin of all birds, perhaps 200 million years ago. The claim that *HAR1F* is 'present only in mammals and birds'[2] starting about 300 million years ago is remarkable, and poses the obvious question as to why it appears to be absent in reptiles. Recall the popular story that birds evolved from a dinosaur line.

Highly mutable or highly conserved?

Now, for 200 million years two distinct lineages would be able to accumulate mutations. The fact that chimps and chickens share but two differences can only mean that

Table 1. Comparison of part of the *HAR1* RNA gene beween humans and some primates shows significant differences.[9] Part of the 2,794 nucleotide long *HAR1* gene[10] shows differences too large to be accounted for by random mutations. Note that the chimpanzee, gorilla, orangutan and macaque genomes in this region are 100% identical, with one exception (macaque, fifth column).(From Pollard *et al.*[9])

										10										20										30						
human	t	G	t	c	a	G	c	t	g	a	a	a	t	G	a	t	G	g	g	C	g	t	a	g	a	c	G	c	a	C	g	t	c	a	g	c
chimp	t	A	t	c	a	A	c	t	g	a	a	a	t	T	a	t	A	g	g	T	g	t	a	g	a	c	A	c	a	T	g	t	c	a	g	c
gorilla	t	A	t	c	a	A	c	t	g	a	a	a	t	T	a	t	A	g	g	T	g	t	a	g	a	c	A	c	a	T	g	t	c	a	g	c
orang	t	A	t	c	a	A	c	t	g	a	a	a	t	T	a	t	A	g	g	T	g	t	a	g	a	c	A	c	a	T	g	t	c	a	g	c
macaque	t	A	t	c	a	G	c	t	g	a	a	a	t	T	a	t	A	g	g	T	g	t	a	g	a	c	A	c	a	T	g	t	c	a	g	c

			40								50										60									70					
human	G	g	C	g	g	a	a	a	t	G	g	t	t	t	c	t	a	t	c	a	a	a	a	t	G	a	a	a	g	t	G	t	t	t	a	g
chimp	A	g	T	g	g	a	a	a	t	A	g	t	t	t	c	t	a	t	c	a	a	a	a	t	T	a	a	a	g	t	A	t	t	t	a	g
gorilla	A	g	T	g	g	a	a	a	t	A	g	t	t	t	c	t	a	t	c	a	a	a	a	t	T	a	a	a	g	t	A	t	t	t	a	g
orang	A	g	T	g	g	a	a	a	t	A	g	t	t	t	c	t	a	t	c	a	a	a	a	t	T	a	a	a	g	t	A	t	t	t	a	g
macaque	A	g	T	g	g	a	a	a	t	A	g	t	t	t	c	t	a	t	c	a	a	a	a	t	T	a	a	a	g	t	A	t	t	t	a	g

				80										90								100								
human	a	g	a	t	t	t	t	c	c	t	c	a	a	G	t	t	t	c	a	a	a	t	G	a	G	g	c	g	a	a	t	c	c	c
chimp	a	g	a	t	t	t	t	c	c	t	c	a	a	A	t	t	t	c	a	a	a	t	T	a	T	g	c	g	a	a	t	c	c	c
gorilla	a	g	a	t	t	t	t	c	c	t	c	a	a	A	t	t	t	c	a	a	a	t	T	a	T	g	c	g	a	a	t	c	c	c
orang	a	g	a	t	t	t	t	c	c	t	c	a	a	A	t	t	t	c	a	a	a	t	T	a	T	g	c	g	a	a	t	c	c	c
macaque	a	g	a	t	t	t	t	c	c	t	c	a	a	A	t	t	t	c	a	a	a	t	T	a	T	g	c	g	a	a	t	c	c	c

this very valuable gene is extremely resistant to mutation. Evolutionists postulate that humans and chimps diverged from a common ancestor some seven million years ago. This provides only 4% as much time for mutations to occur (7 out of 200 million years). Furthermore, considering the much longer human and chimp generation times involved suggests far less than 1% as many generations, and thereby mutational opportunities, from the proposed bird/chimp ancestor to chimp. How many differences in the *HAR1F* gene of chimps and humans would one expect to find? The evolutionary model predicts no fixed mutations (perhaps one or two at most) for chimps during this period. And why should a chimp variant leading eventually to humans be any different?

However, in humans we find *18* different nucleotides compared to chimpanzees. This is very remarkable and by no means consistent with evolutionary thinking, which still relies heavily on a 'molecular clock' model whereby mutations accumulate at a fairly constant rate.

How could the *HAR1F* gene suddenly take on 18 mutations in the last few million years? It is implausible to argue that this region of the genome contains a huge number of hidden mutational opportunities waiting to be revealed by any one of a large number of candidate mutational combinations. If so many recombinations could produce this effect, then why did numerous other organisms not 'discover' some of these during 200 million years of trial-and-error, using much greater population sizes than a chimpanzee ancestral population could have had?

Recombination occurs between identical or very similar sequences. The unique *HAR1F* sequences were not already present elsewhere on the genome, so apparently one is invoking a large number of recombinations, one after the other, using sexual mates having similar regions on their chromosomes but with enough differences to produce countless new combinations. But evidence for this would be easy to find, and the data does not support the notion of a huge variety of *HAR1F* sequences scrambled by wild recombinations. Quite the opposite, extreme intolerance to change must be assumed by evolutionists if some 200 million years of opportunities generated only two differences between chickens and chimps.

Climbing Mount Improbable

We read above that pairs of substitutions together further stabilize the structure of *HAR1* RNA. This kind of fine-tuning would require several base pair (bp) mutations. But these are very rare, in the order of 5×10^{-11} per bp each generation.[4] An average effective population size of about 10,000 individuals is commonly assumed by evolutionists.[5] Assuming the differences are distributed throughout all the reported 2,795 bp[6] positions of the gene, it would take about 1,000 generations just to obtain the first random mutation in one individual. The great majority of random mutations have no beneficial value, and even those which do have very poor chances of becoming fixed in a population due to genetic drift. The chance of obtaining several bp mutations simultaneously on the same gene for such small populations is negligible, and not a serious option.

On average, 1,000 generation intervals are needed to produce another random mutation on this gene somewhere among the population members. One such mutation must not only be favourable, but must occur in at least one of the offsprings from the first favoured mutant for this scenario to work. Moreover, this a scenario failed during 200 million year's of attempts using vastly greater population sizes. For statistical reasons, being constrained to less than 10^6

generations, it cannot be made to work even when huge populations and directional selection are invoked.[7]

If against all odds, one wishes to believe this is how the differences between human and chimp *HAR1F* gene arose, then one must accept some sobering evolutionary side-effects:

- that natural selection has only been focusing on a single trait, favouring improved *HAR1F* versions. This means that the rest of the genome would be subject to weak purifying selection and would therefore degrade. This is because only so many offspring are available to be sacrificed in weeding out bad mutations while still avoiding extinction. As can be deduced by Remine's insights into the meaning of 'cost' in natural selection,[8] a gene version among mammals can only be selected for at the expense of eliminating fewer deleterious mutations. Only a limited number of human offspring are produced, and the sum of causes of death, including the permission for positive selection of a new *HAR1F*, must be less than the number of survivors required to ensure a stable population size.

- as natural selection is focused on spreading the *HAR1F* gene throughout the human population, it cannot simultaneously favour other mutations. Therefore, natural selection cannot also account for all the other differences between the human and chimpanzee genome, including 48 *other* non *HAR1F* areas with huge differences, mentioned in the quote above.[1]

Factors such as 'the stress of man getting out of trees and walking on two legs'[2] have no causal influence in generating the necessary useful mutations, and the best way to avoid such 'stress' would simply be to return to the trees where the other primates have remained, and walk on all fours!

Non-random mutations or special creation?

There is no plausible evolutionary answer to Pollard's findings, unless materialists wish to argue that: (a) the mutations required to stabilize the structure of the mature RNA molecule are of a non-random character and were induced by unknown external factors—a speculative tool which could always be invoked *ad hoc*; or (b) a series of events so statistically unlikely occurred that these would be indistinguishable from a miracle.

On the other hand, the *HAR1F* gene can readily be interpreted as being the result of special creation. It 'turns on in human fetuses at seven weeks after conception and then shuts down at 19 weeks'.[2] Such precise regulation during extraordinarily complex developmental processes clearly suggests design. Special signals in the form of nucleotide sequences and the associated transcription factors need to be exactly synchronized *a priori* for this perfect timing to work. The evolutionary scenario requires more precise mutations, fixed throughout the human population in a relatively short time span, than the theory can accommodate.

Instead of an unreasonable series of unfathomable accidents, the design model can offer some positive research suggestions. The research community is urged to examine the structure of the molecular partners associated with *HAR1F* (proteins, RNA), as well as the nature of their regulatory pathways. A recent human-chimpanzee

common ancestry predicts that very few useful differences in this ensemble of molecules could have been produced by mutations and natural selection. Instead, we believe that even more differences associated with *HAR1F's* functions will be discovered in the future. These differences will be be too subtle to attract much attention, and may appear to evolutionists to merely be the result of uninteresting random mutations. We believe these differences are going to reveal some delicate fine-tuning, which will be too complex to be simply explained by random mutation. As we examine more carefully the supposedly small differences between some chimpanzee and human genes, we will be also able to discern that the differences between these two species are considerably more significant than implied in the evolutionary literature.

References

1. Pollard, K.S., *et al.* An RNA gene expressed during cortical development evolved rapidly in humans, *Nature* **443**(7108):167–72, 2006.

2. Associated Press, Scientists link gene to human brain evolution, 17 August 2006, <www.ctv.ca/servlet/ArticleNews/story/CTVNews/20060817/brain_evolution_gene_060817/20060817?hub=SciTech>.

3. Ponting, C.P. and Lunter G., Evolutionary biology: human brain gene wins genome race, *Nature* **443**(7108):149–150, 2006.

4. Drake, J.W., Charlesworth, B., Charleswroth, D. and Crow, J.F., Rates of spontaneous mutation, *Genetics* **148**:1667–1686, 1998.

5. Hartl D.L. and Clark, A.G., *Principles of Population Genetics*, 3rd edition,. Sinauer Associates, Inc., Sunderland, MA, 1997.

6. <www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?db=Nucleotide&dopt=GenBank&al=DQ860409>

7. Pollard, K.S., *et al.*, Forces shaping the fastest evolving regions in the human genome, *PLoS Genet.* **2**(10):1599–1611, Published Online 13 Oct. 2006.

8. ReMine, W.J., Cost theory and the cost of substitution—a clarification, *J. Creation* **19**(1):113–125, 2005.

9. Pollard *et al.*, ref. 7, supporting information, figure S8, <genetics.plosjournals.org/archive/1553-7404/2/10/supinfo/10.1371_journal.pgen.0020168.sg008.pdf>.

10. Gene length for the human version according to GenBank: <www.ncbi.nlm.nih.gov/BLAST/> accession DQ860409. Note: the gene HAR1A is also called HAR1F.

Peer Terborg has an M.Sc. in Biology (Hons. biochemistry and molecular genetics) and a Ph.D. in Medical Sciences from the University of Groningen, The Netherlands. He is currently working on the cellular and molecular aspects of pulmonary diseases, such as asthma and COPD, and is an expert on the molecular biology of signal transduction and gene expression.

Royal Truman has bachelor's degrees in chemistry and in computer science from SUNY Buffalo, an M.B.A from the University of Michigan, a Ph.D. in organic chemistry from Michigan State University and post-graduate studies in bioinformatics from the universities of Heidelberg and Mannheim. He works for a large multinational in Europe.