

**Shared mutations:  
Common descent or common mechanism?**

**Peter Borger**

Corresponding address: The Independent Research Institute on Origins  
c/o Peter Borger, PhD  
Schwarzwaldstrasse 21  
79589 Binzen  
Germany  
Phone: +49 7621 161 0366  
e-mail: [peterborger@hotmail.com](mailto:peterborger@hotmail.com)

## ***Abstract***

Mutations are a fact of life. Darwin gave mutations, which he called natural variation between individuals, a key role to explain the origin of species. The origin and nature of mutations is one of the most fundamental questions of biology, and are a hot topic in origin debates. If mutations are merely a matter of chance, then the alignment of mutations in distinct species that do not reproduce together qualifies as independent molecular evidence of common descent. We know now, however, that mutations are not utterly chance driven phenomena as the DNA context may determine to a considerable extent where mutations occur. If mutations are modulated because of biophysical mechanisms the question is not whether rules and laws determine where mutations are introduced, but rather – do non-random mutations affect phylogenetic analysis? The DNA analysis of the 1G5 gene in *Drosophila melanogaster* demonstrates that over 70 percent of the mutations that are shared between subpopulations of species that do not interbreed are independent of common descent. Likewise, over 50 percent of the mutations in the *GULO* pseudogene that are shared between humans and the great apes are mutational hot spots also found in guinea pigs – they exactly match the mutations that set humans and primates apart from the rat and line up independent of common ancestry. This paper advances a new hypothesis to understand alignment of mutations in homologous DNA sequences of separated species as the result of a common mechanism operating in similar genomes, and provides the first biological evidence that the location where a mutation will occur and the type of mutation (transition or transversion) are largely predetermined. The consequence is that we may not be able to discriminate between common descent and this common mechanism.

## ***Introduction***

In *The Origin of Species* Darwin considered natural variation not to be necessarily random. He commented, however, that it is reasonable to treat them as random because their cause and origin were unknown [Darwin, 1859]. In his latter days, Darwin started to believe that... “the birth both of the species and of the individual are equally parts of that grand sequence of events, which our minds refuse to accept as the result of blind chance” [Darwin, 1871]. Variation in populations of organisms is merely an observation, however, and it is logically and scientifically incorrect to refer to an observation of which cause and origin are unknown as ‘random’ or ‘blind chance’. We do not know the cause and origin of gravity, for instance, but we know that gravity does not act at random. We do also not know the cause and origin of mutations, but the current consensus is that they are the result of an entirely random process. The current consensus is based, I am afraid, on a misinterpretation of the Luria-Delbrück experiments.

## ***The current consensus***

In the 1940s, two microbiologists, Salvador Luria and Max Delbrück, conducted a series of experiments to prove that mutations are not related to the environment [Luria & Delbrück, 1943]. To this end they cultured strains of *Escherichia coli*, which they exposed to a lethal selective pressure – the bacterial virus T1. As this virus immediately kills non-resistant cells, the only bacteria to survive are those with a pre-existing specific mutation that renders immunity. Luria and Delbrück analyzed the number and distribution of surviving bacteria and concluded that the relevant mutations had occurred at random and had been present before they applied the selective pressure. The Luria-

Delbrück experiments proved that a particular form of variation – that is the mutations that induced resistance to the T1 virus – was not induced by the environment. Rather, the resistance was present from the start otherwise none of bacteria would have survived the lethal exposure. It is important to realize that the experiments did not reveal the nature of the pre-existing mutations. Where exactly had mutations been introduced? In what genes had changes occurred? And, how had they been introduced there? The experiments also did not address whether the resistance was the result of point-mutations in genes, or whether they resulted from gene inversions, gene duplications and losses, or simply reflected the reshuffling of the bacterial genome. Nevertheless, the Luria-Delbrück experiments were interpreted as evidence that *all* mutations are random. All variation found in organisms is the result of random mutations in DNA sequences; mutations are unpredictable, not associated with behaviour or food; neither with social or environmental conditions. Mutations just happen and afterwards they are inherited. The current consensus is thus that mutations in a DNA sequence are introduced at random and occur only once.

Mutations that spontaneously and randomly appear in the germ line can be inherited from the one generation to the next and can be followed in time. The alignment of a particular mutation in DNA sequences obtained from several reproducing species can therefore be taken as evidence of common descent; the mutation appeared in an ancestor and is now present in many descendants due to the principle of inheritance. Because of the current consensus, any alignment of mutations in DNA sequences, also between species that do not reproduce together, is considered molecular evidence of common descent. Mutations

that do not line up in phylogenetic analyses are novel mutations that were introduced after the organism split into distinct species and are evidence of the random character of mutations. What if mutations are not so random after all, as Caporale has argued; what if mutations are modulated [Caporale 1999; Caporale 2000; Caporale 2003]? It may turn out that the alignment of mutations may not be evidence of common descent at all, but rather reflect common biophysical properties of organisms. The external environment of the organism may not be directing mutations, as the Luria-Delbrück experiments indicated, but there could be another environment that determines where a mutations is introduced – the DNA environment. If the position where mutations occur is predetermined due to physiochemical properties of the DNA sequence this will have far reaching consequences for phylogenetic analysis and molecular evidence that supports evolutionary theory.

### ***Non-random mutations in Drosophila***

The idea that mutations might probably not be an entirely random phenomenon first occurred to me when I read a paper of Schmid and Tautz that discussed the 1G5 gene of *Drosophila melanogaster* and *D. simulans* [Schmid & Tautz, 1997]. The gene is found in both species as a unique single copy gene; it is of unknown function but not a pseudogene. The 1G5 gene had the authors' particular interest because it was the fastest changing gene of their study. The sequence of the 1G5 gene counts 1'081 base pairs, which includes only one small intron of 61 base pairs. The major part of the individual genes is identical and not interesting for analysis. As shown in figure 1, Schmid and Tautz located 75 polymorphic sites in an 864 base pair segment, which included the

intron, in thirteen populations of *D. melanogaster* and four of *D. simulans*. Most of them are point mutations, but there are also indel mutations in *D. simulans*. Analysis of the mutations showed that the fraction of variable sites in the intron and the exon is approximately the same. The authors concluded that almost none of the amino acid positions may be under strong selective constraint, because the fraction of polymorphic sites in the intron is comparable to the fraction of polymorphic sites in the coding region, and that a comparison between fixed and polymorphic sites between the two species shows also no significant deviation from the assumption of a neutral evolution in this region [7]. Surprisingly, the authors completely missed the intriguing observation of non-random mutations in the 1G5 gene.

#### ***A reanalysis of the 1G5 gene***

As shown in figure 1, the introns of the individual genes vary considerably *between* the two species (13/61 nucleotides are different: 21%), whereas introns *within* subpopulations of *D. melanogaster* show almost no variation (1/61 nucleotides vary in only 3 out of 13 subpopulations: 1.6%). Likewise, introns in the 1G5 gene found in the subpopulations of *D. simulans* do not demonstrate variation at all. Comparing the introns of the two species reveals ten polymorphic sites immediately adjacent to each other (figure 1, position 153-162). The chance that ten point mutations occur at random in the intron equals  $1.4 \times 10^{-18}$ . By way of contrast, the chance that ten adjacent mutations occur in the intron equals  $2.2 \times 10^{-14}$ . Because the authors demonstrate no significant deviation from the assumption of a neutral evolution in this region, natural selection cannot explain this cluster of mutations. It is therefore reasonable to assume that the cluster of mutations

observed in the introns was not the result of a random accumulation of (point) mutations. Another remarkable observation is that the 1G5 genes in subpopulations of *D. melanogaster*, as far apart as Australia (*D. mel-1*), Asia (*D. mel-13*) or Canada (*D. mel-4*) are completely identical. Likewise, the populations present in Europe (*D. mel-7*), South-America (*D. mel-9*), and North America (*D. mel-12*) share exactly the same five mutations – one present in the intron and four in exon 2; the polymorphisms at position 835 are observed in Japan (*D. mel-8*) and Peru (*D. mel-9*), while that at position 861 is present in Australia (*D. mel-3*) and the USA (*D. mel-10*). Analysis of the polymorphic sites in *D. melanogaster* shows that 8 out of 11 are shared and line up as the result of mutational hot spots – that is over 70 percent.

### ***The illusion of common descent***

The alignment of mutations in the 1G5 genes are not explained by Darwin's selection hypothesis – the genes behave neutrally and if natural selection shaped these genes we must introduce neutral selection. Besides, the sequences were obtained from species inhabiting separated continents and are therefore reproductively isolated. The logical consequence is that the shared mutations in the 1G5 genes are due to a biological or physical mechanism. In other words, the mutations in the 1G5 gene are *non-random mutations* that would produce an alignment of mutations in separated species that do not reproduce together. The alignment is not due to common descent, however, but rather through a common mechanism. If there was no additional information for the 1G5 genes, we would most probably be inclined to argue that *D. mel-1*, *-4*, *-5*, *-6*, *-11* and *-13* are very closely related, since they all have exactly the same gene and thus have a very recent

common ancestor. Likewise, *D. mel-7*, *-9* and *-12* must have a recent common ancestor as they have several shared point mutation; for the same reason *D. mel-3* and *-10* must share a recent common ancestor. This may seem a logically sound conclusion because of the consensus that mutations are merely introduced at random and shared mutations are evidence of common descent. But, the 1G5 data show that an alignment of mutations can also be the result of a common mechanism active in the genomes of the organisms. The 1G5 genes demonstrate that the shared mutations in separated, reproductively isolated organisms could as well be the result of a non-random mechanism – a mechanism that produces the *illusion of common descent*. An important question to be addressed is whether such non-random mutations are the rule rather than the exception? If they are common, it may be impossible to discriminate between common descent and a common mechanism. This is a very extraordinary claim – extraordinary claims require extraordinary evidence.

### ***The extraordinary evidence: the GULO gene***

I have always had a fascination for the *GULO* gene. The *GULO* gene specifies the protein gulono-lactono-oxidase that catalyses the final step in the biosynthesis of ascorbic acid – vitamin C. With a few exceptions, all animal species are able to synthesize vitamin C. Those who do not produce it require a regular dietary intake, otherwise they will develop scurvy. Humans do not synthesize vitamin C. In fact, all the families tested from one suborder of primates – the *Platyrrhini* – also lack the ability to make their own vitamin C, whereas those from the other suborder – the *Catarrhini* – make the vitamin in the liver. Humans and the great apes are not the only organisms lacking the ability to synthesize

vitamin C, however. All tested species of bats (*Chiroptera*) are deficient in producing vitamin C [Birney *et al*, 1976]. Another example is the guinea pig, which appears to be the only non-producer among the rodents. Degenerative loss of vitamin C biosynthesis has evidently occurred quite frequently<sup>1</sup>. The study of *GULO* pseudogene sequences may be informative to reveal whether mutations are a non-random phenomenon, since – due to lack of purifying selection – mutations easily accumulate in pseudogenes.

In 1999, Japanese scientists at Wakayama Medical University determined the exact DNA sequences of the *GULO* pseudogenes of humans, the great apes and the macaque, and compared their data to the active gene found in the rat [Ohta & Nishikimi, 1999]. The sequences they reported show that the pseudogenes of the great apes are almost identical, but differ on several positions to that of the rat's active gene. As shown in figure 2, comparing the sequences of human, chimpanzee, gorilla and orangutan reveals a single nucleotide deletion at position 97 in the coding region of exon X in all primates. If the genes had been active genes, the deletion would have caused a frame shift and inactivated the gene. It is this shared mutation that is often presented as molecular evidence that proves common descent of the *Platyrrhini*. The mutations in the *GULO* pseudogene could indeed provide such evidence, but then *non-random mutations that create the illusion of common descent must be excluded*.

In 2003, the same Japanese group published the complete sequence of the guinea pig *GULO* pseudogene and compared it to that of humans [Inai *et al*, 2003]. They reported a

---

<sup>1</sup> An intriguing question is – why would an inactivated *GULO* pseudogene be fixed in a common ancestor? Darwinians must propose the inactivation of the gene provided a selective advantage.

large number of shared mutations (substitutions) present in both organisms. Under the assumption of an equal chance of substitution throughout the sequence, the probability of the same substitutions in both humans and guinea pigs occurring at the observed number of positions and more was calculated to be  $1.84 \times 10^{-12}$ . This extreme small probability indicates the presence of many mutational hot spots in the sequences [Inai *et al*, 2003]. Remarkably, the mutational hot spots found in guinea pigs and humans exactly match the mutations that set humans and primates apart from the rat. The *GULO* sequences presented in [Ohta & Nishikimi, 1999] and [Inai *et al*, 2003] are combined in figure 2. It shows that 11 out of 21 positions of exon X are shared and line up as the result of mutational hot spots – that is more than 50 percent. Even the single nucleotide deletion at position 97 is a non-random position. If non-random mutations make up 50-70 percent of the mutations in a DNA sequence they would have dramatic consequences for the molecular evidence of common descent; 50 percent is approximately the part of mutations shared between humans and chimpanzees sequences that are usually interpreted as the result of common descent<sup>2</sup>.

### ***Conclusions***

The presented examples challenge the idea that alignment of mutations is compelling evidence of common descent at the molecular level. Rather, shared mutations may be the result of common mechanisms. Up to 50 percent of all mutations of homologous DNA sequences of distinct species may line up due to such mechanisms and create a genetic mirage – the illusion of common descent. How these mechanisms operate is not yet

---

<sup>2</sup> Of course, there are also randomly introduced mutations, but these do not line up in genetic comparisons.

understood, but the elucidation poses an excellent challenge for the scientific community concerned with alignment of mutations. In the meantime, the principle of Occam's razor – also known as the *principle of parsimony* – dictates that scientific explanations must be simple; they should not contain unnecessary assumptions. The simplest explanation for shared mutations between humans and the great apes is not common descent, but rather a common mechanism that introduced the mutations on the same spot in the DNA sequence. Non-random mutations may not only explain alignment of mutations, they could also be the basis to understand linkage disequilibrium, convergent evolution and the recent observation that selection appears not to act on genes with a high mutation rate [Wyckoff *et al*, 2005].

## *References*

Birney, E. C., R. Jenness and A. M. Kayaz [1976], Inability of bats to synthesise L-ascorbic acid. *Nature* **260**: 626-628.

Caporale, L. H. [1999], Chance favors the prepared genome. *Ann. N. Y. Acad. Sci.* **870**:1-21.

Caporale, L. H. [2000], Mutation is modulated: implications for evolution. *Bioessays* **22**: 388-395.

Caporale, L. H. [2003], Natural selection and the emergence of a mutation phenotype: an update of the evolutionary synthesis considering mechanisms that affect genome variation. *Annu. Rev. Microbiol.* **57**: 467-485.

Darwin, C. [1859], *The Origin of Species by means of Natural Selection or The Preservation of Favoured Races in the Struggle for Life*. John Murray, London.

Cited from: Penguin Classics, 1985.

Darwin, C. [1871], *The Descent of Man*. John Murray, London. Cited from: <http://www.literature.org/authors/darwin-charles/the-descent-of-man/chapter-21.html>

Inai, Y., Y. Ohta and M. Nishikimi. [2003], The whole structure of the human nonfunctional L-gulono-gamma-lactone oxidase gene--the gene responsible for scurvy--and the evolution of repetitive sequences thereon. *J. Nutr. Sci. Vitaminol. (Tokyo)* **49**: 315-319.

Luria S. E. and M. Delbrück [1943], Mutations of bacteria from virus sensitive to virus resistance. *Genetics* **28**: 491-511.

Ohta, Y. and M. Nishikimi [1999], Random nucleotide substitutions in primate nonfunctional gene for L-gulono-gamma-lactone oxidase, the missing enzyme in L-ascorbic acid biosynthesis. *Biochim Biophys Acta* **1472**: 408-411.

Schmid, J. K. and D. Tautz [1997], A screen for fast evolving genes from *Drosophila*. *Proc Natl Acad Sci USA* **94**: 9746-9750.

Wyckoff G, C. M. Malcom, E. J. Vallender and B. T. Lahn [2005], A highly unexpected strong correlation between fixation probability of nonsynonymous mutations and mutation rate. *Trends Genet* **21**: 381-5.

### *Legends to figures*

Figure 1.

Non-random mutations in the 1G5 gene of 13 isolated populations of *Drosophila melanogaster* and 4 populations of *D. simulans*. The numbers indicated at top (read downward 141, 153, 154, etcetera) refer to the exact positions of the polymorphic sites in the 1G5 gene. The exon and intron are indicated at bottom. Non-random mutations ('hot spots') are indicated in bold font and make up over 70 percent (8/11) of all polymorphic sites found in *D. melanogaster*.

Figure 2.

Overview of the mutations in exon X of the *GULO* pseudogene of human, chimpanzee, orangutan, and guinea pig and the active *GULO* gene of the rat. The illusion of common descent is created because of non-random mutations ('hot spots') and indicated by bold font. Over 50 percent (11/21) of the polymorphic sites shared between humans, the great apes and guinea pig line up independent of common descent. Other alignments of mutations observed in the sequences of the great apes and human, e.g. positions 55 and 131, may also be the result of non-random mutations. Arrows indicate the nucleotide positions; the deletion mutation on position 97 is indicated with a hyphen (-).

```

111111111111112223333344444444555555555556666666666777777777788888888999
455555556666714723448134688999122355677789011333444890122344557811366889002
134567890124814154364842069568667725302950546578238878202149174349515490472
AGAAATGTATTGATTAGTGGGAAGAAAGGCTCACAGGCACACAGAGCGGGGGTACAATCTGCGAGAAGGTT
D.Mel-1 (Australia I) .....C.....
D.Mel-2 (Australia II) .....A.....T.....
D.Mel-3 (Australia III) .....A.....T.....
D.Mel-4 (Canada) .....
D.Mel-5 (Cyprus) .....
D.Mel-6 (Iraq) .....
D.Mel-7 (Italy) .....A..GG..A.A.....
D.Mel-8 (Japan) .....T.....A.....T..
D.Mel-9 (Peru) .....A..GG..A.A.....C.....A.....
D.Mel-10 (USA I) .....T.....
D.Mel-11 (USA II) .....
D.Mel-12 (USA III) .....A..GG..A.A.....C.....
D.Mel-13 (USSR) .....

D.sim-1 (Colombia) GCCTGGCAATGGAGGCAATGACCCGAAAAATCGTTACATTG-AAAGGGCCAAACACTCACCCAG-AGTACGGC
D.sim-2 (Mexico I) .....T.....GC.....A..-GG...G..T..T.....A...-A.....
D.sim-3 (Mexico II) .....T.....GC.....A..-GG...G..T..T.....A...-A.....
D.sim-4 (Zimbabwe) .....G.A.....A.GGC.AT.....A..GGG.TT.G.....-A.G...
{--intron--}{-----exon2-----}

```

**FIGURE 1**

	01	10	20	30	40	50	60
	↓	↓	↓	↓	↓	↓	↓
Rat	<b>GGAGAAGACCAAGGAGGCCCTACTGGAGCTAAAGGCCATGCTGGAGGCCACCCCAAAGT</b>						
Human	<b>AA</b> .....C..... <b>G</b> ..... <b>G</b> .....G..... <b>TG</b> . <b>G</b> ..						
Chimpanzee	<b>AA</b> .....C..... <b>G</b> ..... <b>G</b> ..........G. <b>G</b> ..						
Orangutan	<b>AA</b> .....C..... <b>G</b> ..... <b>G</b> .......... <b>TG</b> . <b>G</b> ..						
Guinea pig	<b>A</b> ..... <b>G</b> ..... <b>G</b> ..AG.....A..... <b>G</b> ..						
	61	70	80	90	100	110	120
	↓	↓	↓	↓	↓	↓	↓
Rat	<b>GGTAGCCCACTACCCCGTAGAGGTGCGCTTCACCCGAGGCGATGACATTCTGCTGAGCCC</b>						
Human	...GT..... <b>TG</b> .. <b>G</b> . <b>G</b> ...A.....T.-A..... <b>C</b> ..A.....						
Chimpanzee	...GT..... <b>TG</b> .. <b>G</b> . <b>G</b> .C.A.....T.-A..... <b>C</b> ..A.....						
Orangutan	...GT..... <b>G</b> .. <b>G</b> . <b>G</b> .....A-A.....G. <b>C</b> ..A.....						
Guinea pig	..C..... <b>T</b> .. <b>G</b> . <b>G</b> ..... <b>G</b> ...C..... <b>C</b> .....						
	121	130	140	150	160		
	↓	↓	↓	↓	↓		
Rat	<b>CTGCTTCCAGAGGGACAGCTGCTACATGAACATCATTATGTACAG</b>						
Human	..... <b>T</b> .....C.....C..... <b>ACC</b> .....						
Chimpanzee	..... <b>C</b> .....C.....C..... <b>ACC</b> .....						
Orangutan	..... <b>CA</b> .....C.....TC..... <b>ACC</b> .....						
Guinea pig	..C..... <b>TGC</b> ..A.....						

**FIGURE 2**