Updated August 11th, 2015

CONTRIBUTIONS TO MUTATION STRATEGIES AND OTHER MOLECULAR EVOLUTION TOPICS

JACQUES NINIO

(included in the web site http://www.lps.ens.fr/~ninio)

TOPICS DISCUSSED HERE:

1. *Overview*

2. *Mutation strategies*
2.1 - The need for complex mutational events
2.2 - Transient mutators
2.3 - The connectedness of errors
2.4 - Multiple mutations in higher organisms (the herezygosity-induced mutations hypothesis, or “gene conversion as a focusing device”)
2.5 - Why DNA repair must be error-prone
2.6 – The blind spots of population genetics

3. *Molecular approaches to evolution*

TOPICS NOT DISCUSSED HERE:
- Prebiotic replication and catalysis (see web chapter on the origins of life)
- The evolution of tRNA 3d structure (see web chapter on the origins of life)

1. OVERVIEW

Up until now, my best known contribution to molecular evolution is the « transient mutator » concept, developed in 1991, see [1] and Section 2.2 . A bolder hypothesis developed in 1996 [2] on the generation of complex mutations in higher organisms may turn out to be even more important. This second contribution, predicted that regions of heterozygosity in diploid organisms should trigger mutations in their vicinity, and this prediction is gaining experimental support through the work of William (Bill) Amos [3-5] on HI-induced mutations, and the parent-progeny genome sequencing work of Sihai Yang et al. [6] , see also Arbeithuber et al. [6a] on gene conversion induced mutations.

Concerning transient mutators, using simple semi-quantitative arguments, I showed that when an *E. coli* population is grown in a non-limiting medium, genetic mutators (bacteria, that produce mutations at a much higher rate than the standard rate) produce a rather small fraction of the total mutation output of the population. On the other hand, there must exist, within the population, bacteria that display a mutator *phenotype* for just one or two generations, then switch back to the standard phenotype. These « transient » or « phenotypic » mutators would be responsible for most of the observed double mutation events. Due to their existence, the frequency of double mutations should be far higher (say, 50 times higher) than expected from the single mutation frequency. Such a property would be extremely useful to explain how compensatory mutations may arise in evolution, or how codon usage becomes established.

Concerning mutations triggered by gene conversion, I published a sequel to the transient mutator article in 1996 [2] that is less known, but that is potentially far more important, on how multiple mutations may arise in higher organisms. In this case, they would not arise simultaneously, but one after the other at a recombination hot spot created by local heterozygosity. The title of the 1996 article was “gene conversion as a focusing mechanism for correlated mutations”.

The two articles are contributions to the broader field of « mutation strategies » or « adaptation strategies ». I had written on this topic earlier (e.g. [7, 8]). J. Mark Baldwin who invented the concept of phenocopy was one of the main precursors in this field ([9], review in [10]). The field has recently gained respectability and popularity (e.g., reviews in [11-13], and more references in [14]).

While in my early molecular biology work I had been mostly

concerned with the accuracy of molecular processes (for, instance, how the cell manages to translate messenger RNA into protein with high accuracy) I was now exploring the other side of the coin : How the various forms of variability, both phenotypic and genotypic were controlled and perhaps optimized. The concept of sequence space is at the interface between the two domains. Mutations are often described as the successive steps of a walk in the sequence space. On the other side, the existence of transcription and translation inaccuracies results in the fact that in any given organism, many sequences that are in the neighbourhood of the genetically encoded sequence are present, and this presence may have important functional consequences [15, 16].

In large measure, my book « Molecular approaches to evolution » [17] was an attempt to discuss all that was important in molecular evolution from the viewpoint of someone who believes that a gene produces, beyond its canonical product, a small amount of closely related products. A new gene product, resulting from a gene mutation was there, prior to the mutation, as an erroneous protein synthesis product. From there, subtle properties of molecular evolution can be deduced.

The book was also rooted on earlier interests in the origin of the genetic code, on tRNA 3d structure and its evolution, and on sequence comparisons. The French version was written before the discovery of the non-universality of the genetic code – but I was prepared to that, and before the discovery of RNA splicing (a big surprise for me) and before the discovery of RNA catalysis (less of a surprise, due to my past work on « non-enzymatic replication » -see the web chapter on the origins of life).

2. MUTATION STRATEGIES

2.1 THE NEED FOR COMPLEX MUTATIONAL EVENTS

In a bacterial population of reasonable size ALL mutations are produced ALL the time. There is perhaps some exaggeration in this statement, but consider one gram of *E. coli* cells (10^{12} bacteria). Accepting the standard mutation rate of $3x10^{-3}$ per whole genome replication [18, 19] and a genome size of 5 million base pairs, when this gram of bacteria is duplicated, about $3x10^{9}$ mutations are produced at

constant population size, thus on average, 600 mutations per base pair. Similar back-of-the-envelope estimates, applied to yeast populations, indicate that when one gram of yeast (about 10^{11} cells) are duplicated, there are about 2 mutations per base pair of the $1.4x10^{7}$ nucleotides genome.

These numbers invalidate, in the case of *E. coli* or yeast, the « infinite site » approximation, used by Mooto Kimura [20] to derive his famous equations of the « neutral theory of molecular evolution ». The infinite site approximation states that there occurs no more than a single mutation at any given locus in the genome, in the whole history of the population.

If mutations are so readily available at the population level, many genes can be optimized, by mutation and selection, with respect to single nucleotide changes. Then, in order to improve the situation, larger evolutionary steps are needed, and I will consider here multiple mutations.

There are situations in which the usefulness of double mutational events is intuitive enough. For instance, if an RNA molecule has a functionally important self-complementary secondary structure, any mutation that disrupts a Watson-Crick base-pair would be rather detrimental to the function, so double mutations that change a Watson-Crick base pair into another Watson-Crick base pair may be preferable. Similarly a protein structure may hold through pairs of amino acids that are remote in the primary sequence, yet are in close contact in the 3d structure. Here again, double mutations that change simultaneously the two interacting amino acids could be useful. Quantifying the proportion of « compensatory mutations » in sequence evolution is a rather difficult task (see, e.g. [21-24a]). It is also known that there are strong constraints on codon usage (e.g., [25-27]), that are not entirely explained by mutation pressure. Some of the codon preferences must reflect selective pressures. However, there is practically no in vitro or in vivo evidence of codon replacements that lead to a selective advantage. If single synonymous codon changes are nearly neutral, it might nevertheless be the case that simultaneous multiple codon changes may produce substantial selective advantages or disadvantages. Consider also the problem of how to increase the accuracy of any given process. Errors in one process are due to the noisiness of all the components contributing to the process. A gain in accuracy can be obtained by reducing the noisiness of the most noisy component. At the end, all components would make roughly equal contributions to the global error-rate, so any improvement in the accuracy of a single component will

result in minor overall improvement. Here again, it would seem that mutations, to be beneficial, must hit several components at the same time. A similar argument can be made for a metabolic chain, in which a change in the efficiency of a single enzyme in the chain has very little effect on the efficiency of the complete process [28a, 28b].

2. 2 TRANSIENT MUTATORS

In 1988, Cairns, Overbaugh and Miller published in Nature an article that attracted much attention and generated much controversy [29]. It was soon followed by articles by Barry Hall, pointing in the same direction [30, 31]. The kinetics of appearance of mutants indicated, according to John Cairns that revertants appeared far more frequently than expected on the basis of the known mutation frequency (if one believed in the validity of the Luria-Delbruck test for measuring mutation frequencies).

Since Cairn's argument was based mainly on a quantitative discrepancy between the known standard mutation frequency, and that derived in his case from the Luria-Delbruck fluctuation assay, I tried to figure out whether or not the mutants detected by Cairns might have been produced by a subpopulation of the *E. coli* cells, namely, the mutator population. Conceivably, within the original population, there were genetic mutators – bacteria that produced mutations at a much higher rate than the standard rate (say 1000 times the standard rate). Could it be that most of the revertants originated from this mutator subpopulation ? I made back-of-the-envelope calculations that readily convinced me that this could not be the case. I estimated that in a typical *E. coli* population growing in a non-selective medium, the mutator subpopulation would contribute about 0.3% of the total number of mutations. This would not be sufficient to create a Cairns effect.

I then had the idea to evaluate the contribution to the total number of mutations produced by bacteria that were producing, due to translation errors, faulty DNA polymerases that would be error-prone, and thus responsible for increased levels of mutations, as long as they were present in the cell and were effectively used. A cell using such a DNA polymerase for its replication would behave as a mutator cell, but this would not be a genetic character, and the mutator phenotype would be lost after one or two rounds of genome replication. I considered other possible sources of noise, mainly transcription errors, and an insufficient supply of a limiting mismatch repair protein. Lumping the three

contributions, I estimated that about 1 cell in ten thousand would behave phenotypically as a mutator cell, and that there would be about one in ten thousand such « transient mutators » within the *E. coli* population. These transient mutators would be responsible for more than 10% of the total mutations produced by the population.

Pursuing the back-of-the-envelope calculations, something striking emerged. If you multiply a mutation rate by n = 1000, the rate of double mutations is multiplied by n^2 = a million. So, even though there would be merely one in ten thousand transient mutators, these would be responsible for most of the double mutational events. To put it differently, double mutations would be far more frequent (by a factor 50, I estimated) than would have been predicted from the single mutation rates. This was an important result, because it opened a new way of thinking about the origin of compensatory mutations, about the evolution of codon usage, and related topics discussed above.

At that time I was working in the Jacques Monod Institute, in the same floor as Miroslav Radman. I discussed my ideas with him, and also with Maury Fox who happened to be a visiting scientist in Miro's lab. Having written an article describing my « transient mutator » concept, I submitted it to « Genetics », through its editor-in-chief, John Drake. John Drake immediately accepted to do the editorial job himself.

Reviewer 1 (John Cairns ?) recommended publication after revision, writing « I gained several new thoughts during my reading of the paper and I would like to see it in print ». He made a long report, making a number of objections, based on technical arguments, supported by references to published work. He also suggested me to remove the references to Cairns et al. and Hall, because my paper dealt « exclusively with mutations during growth, under conditions where there is no strong selection for a novel genotype ». In the revised version, I still quoted these authors, considering that it was « fair to state one's initial motivation ».

The second reviewer was a population geneticist. He did not find my paper very useful : « … what difference does it make whether transient mutators contribute little or much to double mutations ? (…).Transient mutators are not genetic and hence not inherited ». In the end of the report, he wrote « In summary, the result that most double mutations are produced by transient mutators is of interest, but by itself it is anecdotic and I am not sure that it warrants a complete publication, especially given the uncertainty of many of the parameters used in the estimate ». I suspect that this is still the opinion of the handful of population geneticists who have heard about this work.

Drake's own judgement was very positive, and from the very beginning, he discussed with me a number of points in the manuscript, making inquiries of his own to get the most reliable estimates of some parameters. He also did some rewriting in the end, after the acceptance of the revised version. So, some of the sentences in the article are of his own writing. Apparently, he does not regret, fifteen years later, his decision to accept the manuscript [32, 33]. The article was summarized as follows :

ABSTRACT of the transient mutator article.

A population of bacteria growing in a nonlimiting medium includes mutator bacteria and *transient mutators* defined as wild-type bacteria which, due to occasional transcription or translation errors, display a mutator phenotype. A semi-quantitative theoretical analysis of the steady-state composition of an *Escherichia coli* population suggests that true strong genotypic mutators produce about $3x10^{-3}$ of the single mutations arising in the population, while transient mutators produce at least 10% of the single mutations and more than 95% of the simultaneous double mutations. Numbers of mismatch repair proteins inherited by the offspring, proportions of lethal mutations and mortality rates are among the main parameters that influence the steady-state composition of the population. These results have implications for the experimental manipulation of mutation rates and the evolutionary fixation of frequent but nearly neutral mutations (e.g., synonymous codon substitutions).

Mutators and cancer.

In 1991, Lawrence Loeb published an article on mutations and cancer, in which he proposed that one of the mutations that occurred in the lineage from a healthy progenitor cell to a malignant cell was a somatic mutation conferring a mutator character. He called the cells carrying this character « phenotypic mutators » because they differed, genetically, from the cells in the germ line [34]. However these cells are not transient mutators. They are mutators by virtue of a genetic alteration that is stably transmitted to the their progeny.

The legacy of transient mutators.

It became rapidly clear that when one looked for double mutation events in bacteria, the contributing bacteria were not genotypic mutators (e.g., Hall [31]). Furthermore, mutants in the accuracy of transcription

[35] or the accuracy of translation [36] boosted the mutation rate, thus validating the estimates of the parameters in the transient mutator paper. There were also a large number of papers reporting DNA sequencing work, showing that there were “too many mutants with multiple mutations” [33]. However, the mechanism underlying the multiple mutations was not clear [33]. Since the detected multiple events were usually within a gene or so, the mutation mechanism must have involved an error-prone DNA polymerase.

The population geneticists learnt nothing from the accumulation of evidence. John Maynard Smith co-authored with a few colleagues, an article in Nature discussing the competition between standard and mutator bacteria, neglecting transient mutators [37] and later five French colleagues produced a detailed treatment [38], using non realistic parameters. Lynch and Abegg produced a computer simulation, tending to show that transient mutators could not play a role in evolution [39], but they used parameters that suited them, without respect for biological plausibility.

2.3 THE CONNECTEDNESS OF ERRORS

In May 1991, I was invited to an EMBO workshop held in France, in Arc et Senans on “Genetics of translation: the interconnectedness of things”. On this occasion, I generalized the back-of-the envelope calculations made for the Genetics paper : I tried to evaluate how translation, transcription and replication errors were tied in *E. coli*. For instance, if mutation rates are increased by a factor 10, what would be the expected effect on ribosomal accuracy ? The calculations were a bit tedious, but in the end, I drew clear conclusions relative to the problem of error-propagation, and what circumstances were needed to produce an error-catastrophe. Another interesting insight came from a consideration of tRNA mischarging errors due to erroneous tRNA sequences produced by transcription errors. There was a paradox because “With a transcription error-rate of 10^{-5} per nucleotide, there would be about 6 erroneous variants of one tRNA per cell, far less than the 240 potential variants. This implies that the errors made in a bacterial line of descent change at each generation. (…). Nevertheless, only one tRNA variant in a hundred is misacylated as though almost all the interactions between the aminoacyl-tRNA ligases and the tRNAs had been tested by evolution one after the other”.

My article [15], entitled “Connections between translation, transcription and replication error-rates” was published in a special issue of Biochimie in 1991. It was summarized as follows :

Abstract of the article on connections between error-rates.
The analysis of published data from *E. coli* suggests that in all three processes of translation, transcription, and replication, a minority of errors are produced by sub-classes of error-prone components. These add to the basal level of errors a noise of about 10 to 30%. Each one of the three processes contributes to the noisiness of the two others in a loose manner : a large increase in one error-rate produces a moderate increase in another error-rate. The strongest influence is that of transcription on translation errors. There it is possible that a majority of the misacylation errors are produced during the encounter of a correct amino acyl-tRNA ligase with a mistranscribed tRNA. Extreme mutator mutants are expected to produce a moderate increase in translation errors.

In 1997, I was invited to contribute to a special issue of « The Origins of Life » in the honour of Leslie Orgel’s 70th birthday. I decided to extend my back-of-the-envelope calculations to higher organisms, namely, yeast and humans. The data on mutation rates in the human species were curious. On one side most genetic studies suggested that the mutation rate was around 10^{-8} mutation per nucleotide in a haploid genome per generation, thus about $0.5x10^{-9}$ per nucleotide in the genome per year. On the other side, there were molecular evolution mutation rates (more precisely, mutation fixation rates) derived from sequence comparisons at neutral loci on phylogenetic trees that were about $2x10^{-9}$, four time higher than the first. So, mutations would be fixed in the population faster than they are produced! The Origins of Life article highlighted this paradox and discussed possible solutions. I suggested, among other possibilities, that there had been a recent slowdown of mutation rates in the human lineage (the main selection in the lineage being on foetal mortality, becoming too high, due to the accumulation of deleterious mutations).

The core of the Origins of Life article [16] was about error-propagation, as the Biochimie 1991 article. But it went further, by discussing the optimality of error-rates in evolutionary context. In particular, I discussed the implication of the fact that in higher organisms

such as humans, having hundreds of billions of cells, a few cells could be, for a given gene, two mutations away from the genotype. And the genes could produce proteins with two translation errors, thus with four amino acid differences from the standard product of the gene. This feature was highlighted in relation to the origin of prion diseases.

The 1997 Origins of Life article was published with the following summary :

Abstract of the article on "the evolutionary design of error-rates and the fast-fixation enigma".

Genetic and non-genetic error-rates are analyzed in parallel for a lower and a higher organism (*E. coli* and man, respectively). From the comparison of mutation with fixation rates, contrasting proposals are made, concerning the arrangement of error-rates in the two organisms. In *E. coli*, reproduction is very conservative, but genetic variability is high within populations. Most mutations are discarded by selection, yet single mutational variants of a gene have, on average, little impact on fitness. In man, the mutation rate per generation is high, the variability generated in the population is comparatively low, and most mutations are fixed by drift rather than selection. The variants of a gene are in general more deleterious than in *E. coli*.

There is a discrepancy in the published mutation rates : the rate of mutation fixations in human populations is twice or four times higher than the individual rate of mutation production, a feature that is not consistent with current population genetics models. Two, not mutually exclusive, hypotheses may explain this 'fast fixation enigma' : (i) Mutation rates have substantially decreased in recent human evolution and (ii) A substantial fraction of the fixed mutations were generated in a process – such as gene conversion – that violates the principle of independence of mutation events.

However, the discrepancy between the two estimates of the mutation rates did not perturb the specialists in molecular evolution. Commenting on a codon usage study by Eyre-Walker and Keightley [40], that led to a rather high estimate of the production of deleterious mutations in human populations, the population geneticist Jim Crow wrote [41]:

"Every deleterious mutation must eventually be eliminated from the population by premature death or reduced reproductive success, a

“genetic death”. That implies three genetic deaths per person! Why aren’t we extinct?”

Apparently it did not occur to him that the mutation rate he was favouring could be overestimated by a factor 2 or even 4 [41a]. It seems to me that the published mutation rates, derived from phylogenetic reconstructions have recently “evolved”. Whether or not the changes were prompted by an awareness of the discrepancy I had pointed out, or by reasons internal to the field remains unclear to me. In a series of papers (e.g., [42]) Woodruff advanced another possible origin for the discrepancy between mutation and fixation rates, based on the accumulation of mutations that “actually occur in the cell lineage before germ cell formation or meiosis”.

2.4 MULTIPLE MUTATIONS IN HIGHER ORGANISMS

In 1995, after my migration to Ecole Nomale Supérieure, Michel Veuille, a population geneticist at Pierre-et-Marie Curie university in Paris and Dino Yanicostas, a molecular biologist at Institut Jacques Monod got in touch with me. Yanicostas had made sequencing work on a drosophila gene, and found polymorphic variants that seemed hard to explain in terms of independent single mutations. So Veuille and Yanicostas asked me if my Genetics 1991 transient mutator idea might help explain the sequencing results. We met perhaps twice, then Dino disappeared, and I pursued the discussions with Michel Veuille. Michel Veuille’s input was an excellent background in population genetics. He knew all the papers about genetic polymorphism in natural populations in which the issue of non-independent mutations had been discussed, and he was aware of the attempts to explain with mathematical models the fixation of double mutations.

I was more concerned with molecular mechanisms that might generate multiple mutation events. The problem here was more difficult than in the case of bacteria. We needed a high mutation rate within a short stretch of DNA (about the length of a gene), and a standard mutation rate over most of the genome. My back-of-the-envelope calculations, using now parameters taken from higher organisms, suggested that error-prone ribosomes, error-prone RNA polymerases, or even error-prone DNA polymerases could not explain the production, in significant proportions, of double mutations within genes. DNA repair seemed to offer better prospects.

The paradox of repair inaccuracy. Now, there is a paradox about

DNA repair : It can work efficiently, yet be highly inaccurate ! I explain why in the next paragraph. I was aware of this mathematical paradox, but at the time of this work, in 1995, the idea seemed absurd to most biologists.

Let us consider a simple example : a DNA polymerase replicates a genome at a 10^{-6} error rate. Assume that *every* error is detected by a mismatch repair system, that degrades a « repair patch » of DNA one thousand nucleotides long around the error. Then a repair DNA polymerase resynthesizes the patch at a 10^{-4} error-rate (so, this DNA repair polymerase is one hundred times less accurate than the standard polymerase). Since the repair enzymes make on average one error every ten thousand nucleotide incorporations, the section of DNA corresponding to the repair patch has now 10% chances of containing an error. So the overall genomic error-rate becomes 10^{-7}. Using a repair polymerase that is 100 times less accurate than the correct polymerase produced an overall improvement in replication accuracy by a factor 10! This simple calculation is correct under the assumption that every error is detected prior to repair, so accuracy rests on the efficiency of error-detection, not the fidelity of DNA resynthesis.

Michel Veuille was not exceedingly enthusiastic with the idea of error-prone DNA repair, and I had to go over the quantitative argument several times to have him accept the idea. However, by that time, I had convinced myself that error-prone DNA repair was not sufficient to produce double mutations at a sufficiently high rate.

I then conceived another way to produce double mutations, which, if correct, had a potential to revolutionize population genetics. In my new scheme, double mutations did not arise simultaneously, but in successive generations within the population. I present first a variant of this idea, it is not very interesting, but can be accepted by most specialists without difficulty. Mutations do not occur with uniform probability along a genome. In particular, there may be « mutation hotspots ». These hotspots become loci in which mutations accumulate within the population, so they are loci in which double or multiple changes may be observed. The mutations may have arisen during successive generations within a same line of descent, or they may arise by recombination between chromosomes containing different variants of the initial gene.

I present now a second, far more interesting variant of this idea. It is much more speculative than the first, but if it is correct, It can have far-reaching consequences. Consider a population in which a gene is suboptimal. Several variants of the gene then coexist within the

population. There is some « single-nucleotide polymorphism » (SNP). Two single nucleotide variants, can, by recombination, produce genes that differ by two nucleotides from the initial gene. So far, there is nothing exceptionally new. But now, consider this possibility : Forget the « reshuffling » aspect of recombination that would have allowed two SNPs on different chromosomes to become simultaneously present on a single one. Focus instead on the molecular mechanism of recombination, that involves the degradation of a DNA patch on a chromosome, and its resynthesis as a copy of a homologous stretch of DNA of the other chromosome, as implied by most current recombination models (starting with Robin Holliday's model [43]). As a result of this DNA resynthesis, assuming it occurs without errors, the corresponding DNA sections become identical on the two chromosomes, producing the genetic phenomenon called « gene conversion ». I made two assumptions : (i) the DNA resynthesis involved in this process was rather inaccurate. So, if there were two different SNPs on the corresponding DNA patches on the two chromosomes, after DNA resynthesis there would be a single one on both chromosomes. But, DNA resynthesis being error-prone, there would be a substantial chance of having a *de novo* mutation on the DNA resynthesis patch. (ii) DNA resynthesis mediated by recombination would occur preferentially at loci in which there would be a few (say 2 or 3) differences between the two chromosomes.

If my idea was correct, mutations would arise, within a population, preferentially at polymorphic loci. More precisely, I found that such a mechanism of « polymorphism-induced mutations » could work in populations in which the average genomic divergence between two individuals would be around 10^{-4} or less.

Michel Veuille was not enthusiastic with this idea. In particular, the assumption of error-prone DNA resynthesis associated with recombination did not seem plausible enough to him. Yet, he agreed to be the co-author of an article in which I would introduce my molecular ideas, and in which he would provide the population genetics context. I felt this contribution as important. I was very disappointed by the fact that the populations geneticists had barely noticed my earlier « transient mutator » article, and I found it important, this time, to frame my idea in a way that would be appealing to them.

Writing the manuscript together was not an easy task. I wished to emphasize the most speculative model, involving « polymorphism-induced mutations », and show how it renewed our conceptions about polymorphism – no longer the witness of a passive accumulation of mutations, but an active hotspot for novel mutations. Michel Veuille

wished to make a strong case for the existence of correlated mutations, that was in his opinion, an already foolhardy position in population genetics, and we was not willing to take too much risks with speculative molecular models. We mentioned in the MS that MV preferred the error-prone mismatch repair model, and that I preferred the gene conversion model. The title of the MS was : « DNA processing and gene conversion as sources of clustered mutations ».

I sent the manuscript on June 27th, 1995 to John Drake, still Editor-in-Chief of « Genetics », but he was overloaded. He asked Patricia Foster to be the editor of the manuscript. Patricia handled the manuscript, and received two reviews that were not enthusiastic, to say the least, so she rejected the paper, adding nevertheless « On a more hopeful note, in private comments reviewer #2 thought that, with extensive rewrite, your paper might be suitable for J. Theoret. Biology».

Reviewer #1 was a population geneticist who obviously did not understand molecular genetics. For instance, he found very strange our statement « that the standard error rate of the DNA polymerase…is typically 300 times larger than the mutation rate ». And he believed that all that was needed to understand the fixation of double mutations could be found in Kimura's treatment of the subject.

Reviewer #2 was more favourable to the work. His report started with this sentence : « This manuscript addresses a very interesting and important problem in evolutionary biology by trying to construct models that can explain multiple mutations in short stretches of DNA ». It ended with the sentence : « Until a more convincing case can be made for one model in a specific system I would find such argumentation too speculative for Genetics ».

Michel Veuille received the reports almost with relief, writing to me [44] « I am not too surprised by this barrage of fierce criticism, because I considered this writing to be premature. At least, we obtained a list of authors we must read absolutely. »

Feeling that the rejection of the manuscript was not deserved, I decided to go ahead alone, pushing forward the most innovative model, that involved gene DNA resynthesis associated with recombination, and stripping the manuscript of all the population genetics context, considering that the population genetics community was not ready for such ideas.

The manuscript was then sent on October 30th, 1995 to « Molecular and General Genetics » through my colleague Raymond Devoret, who acted as editor. He was personally favourable to the article, and he obtained two reports. One was very favourable, and the

reviewer (John Cairns) wrote : « It is important that each reader should be able to understand *instantly* the nature of the argument. So I would like to offer the following much simpler version of the mismatch-repair model, starting two lines from the foot of page 4 ». And he thus contributed three paragraphs of his own writing. The other reviewer was broadly favourable, but he went into lengthy discussions about the most reasonable values of each parameter. Finally the article was accepted [2], and here is its summary :

Abstract of the article on mutations boosted by polymorphism: "gene conversion as a focusing mechanism for correlated mutations".

Ways of producing complex mutational events without substantially raising the primary mutation rate are explored. If the small amount of DNA that is resynthesized through the action of the mismatch DNA repair system is not subject to further repair, the incidence of double mutations can increase by a factor 100, while single mutations would increase by only 30%. Such a boost in the incidence of double mutations seems insufficient to meet the needs of higher organisms. For them, an alternative strategy would be to produce complex events by a succession of single mutations occurring in a correlated manner over several sexual generations. It is proposed that gene conversion may fulfil this role. Assuming that the resynthesis of DNA that occurs during gene conversion produces mutations in the conversion track, one predicts a tendency for close mutations in corresponding sequences in the two homologous chromosomes, to promote, during conversion, further mutations in their vicinity. Semi-quantitative calculations suggest that such a mechanism can be quite effective, provided the divergence between two paired chromosomes is around 10^{-4} or less. Such a mechanism might constitute an adaptive mutation strategy acting at the population level.

So far, this article did not attract much attention (but see [33, 45]), yet I cherish this contribution. I believe that if it is ultimately correct, it would stand as one of my most important contributions to molecular biology and genetics. In particular, it could explain the repeated observations of "independent mutations" in human populations (e.g. [46-47a]. It suggests that local heterozygosity can be an important factor in the generation of cellular mutations occurring in the lifetime of a complex organism, thus providing a possible clue on the origin of some cancers.

There was a bold assumption in the first version of the article, as it

was submitted to « Genetics ». The assumption was that DNA resynthesis associated with recombination was error-prone. Strathern [48] and Kuzminov's [49] papers showing that recombination was indeed mutagenic were published in 1995, but I was not aware of them, possibly because their titles did not speak to me.

William Amos enters the scene. In 2009, a reviewer attracted my attention to an article recently published in BioEssays [3] entitled "Heterozygosity and mutation rate: evidence for an interaction" in which Amos developed the concept of "heterozygosity induced" mutations. He reached this concept through his experimental work, first on variations in micro satellite DNA sequences and later on substitution mutations. He was not aware of my 1996 gene conversion paper. So the concept of HI induced mutations was discovered independently by Amos through his experimental work, and by me through theoretical considerations. I wrote to him, and felt that we were on the same side of the fence. After the BioEssays contribution, Amos continued to collect data on HI induced mutations [4, 5] thus strengthening the case. This is a great satisfaction for me.

Gene conversion induced mutations become fashionable. After the publication of a very convincing paper in PNAS by Barbara Arbeithuber et al. [6a] on gene conversion induced mutations, a team of Chinese scientists, with Dacheng Tian as senior author, and the english evolutionist Laurence D. Hurst, as co-author presented what looks like almost perfect evidence for HI-induced mutations. They sequenced the whole genomes of homozygous plants, crossed the parents to produce heterozygous progeny, and determined the mutation levels in the parents and their progeny. They found higher mutation levels in the heterozygotes, and more importantly, they found that the HI-induced mutations clustered close (about 100 bp) to the sites of heterogeneity, thus strengthening an underlying molecular mechanism of the gene conversion type. The authors did not have the decency to quote my 1996 Molecular and General Genetic paper, and they mention Amos at the end of their article, without quoting his 2013 paper [4]. In any event, we can take this positively, and hope that the concept of polymorphism – induced mutations will gain wider and wider acceptance in the near future. It might even be dreamed that within 20 years some population geneticists might perceive the theoretical implications of the phenomenon.

2.5 WHY MUST ALL DNA REPAIR BE ERROR-PRONE.

In May 1998, there was an EMBO meeting on recombination, organized by Steve West and Alain Nicolas in Seillac, France, and I thought that the meeting would offer me a good opportunity to gain a deeper understanding of how recombination worked, and also to present my own ideas about recombination and multiple mutation events. There were about 70 participants from all over the world, and among them a high proportion of people who had reached a high reputation in the field. I was very pleased to meet there people I admired such as Martin Gellert, a pioneer of DNA topoisomerases studies. Benedicte Michel whom I knew from the years at the Jacques Monod Institute was at the meeting, and she helped me a lot by explaining the subtleties of modern ideas, and telling me who did what in the field. I sympathised there with Ishizo Kobayashi, of Tokyo University. In fact I had an opportunity to visit him two years later, and he received me with extraordinary hospitality. I presented my ideas at the meeting. They did not elicit particular interest. Later, I wrote to Martin Gellert, and he reacted very favourably. Thomas Lindahl, whom I had met during the return journey by train told me about Berdal's work [50]. Berdal and co-workers studied nucleotide excision repair by a 3-methyladenine glycosylase. They had shown that correct nucleotides were sometimes excised, and that an excess of *repair* enzymes produced *mutagenic* effects.

The main consequence of the meeting was to plunge me into deep reflections. There were several elements turning into my head. I was concerned in particular with the mechanism of somatic mutations that produced antibody diversification. The evolutionary mechanism that I had proposed in the MGG 1996 article required DNA resynthesis at a locus in which there was some local heterozygosity between the two homologous genes on two chromosomes. However, nothing in the available data on the accumulation of somatic mutations suggested any role for local heterozygosity. So, I had to imagine a mechanism for producing mutations repeatedly, over several cellular generations, on a same DNA segment, regardless of the divergence between the gene copies on the two chromosomes.

Another line of thought was about the consequences of « gratuitous DNA repair ». In my previous back-of-the-envelope calculations of 1991 [1] and 1996 [2], I had considered events in which a mismatch repair system attacked a perfectly complementary stretch of DNA. Being accustomed to the idea that every process had limited reliability, I was naturally inclined to take into account such errors. Quite surprisingly the calculations had shown that the « gratuitous » repair

patches could outnumber the legitimate ones by a factor 10 in *E. coli* [1], or perhaps by a factor 5 in higher organisms (see appendix in [2]).

It struck me that there were perhaps local DNA sequences that would be preferential sites of attack for gratuitous repair. They would have a particular conformation such that a mismatch detecting system would sense it as though it contained mispaired residues, then would go through the degradation and resynthesis business. (The sequences would be « false positives » for the MMR system). Now, if indeed such sequences are attacked, and DNA is resynthesized there accurately, the sequences would be regenerated as they were, and thus would be the targets of further attacks, *ad infinitum*. If however there is some inaccuracy in DNA resynthesis, the local sequence evolves until it looses its false positive character. I named « illusory mismatches » the local sequences detected by the MMR as though they contained real mismatches, and calculated that “strong illusory defects may arise at the decanucleotide level”. An illusory defect “deliberately” created by recombination could be used to initiate somatic hypermutation pathways used in immunoglobulin diversification. The reviewers and the Editor’s reactions to the manuscript were quite interesting [51]. The work was summarized as follows [52] :

Abstract of the article “Illusory defects and mismatches: Why must DNA repair always be (slightly) error-prone?”.

There is growing evidence that recombination is mutagenic and that some forms of DNA repair synthesis are error prone. DNA-repair systems detect structural defects in DNA with high efficiency but they occasionally also strike at normal sections of DNA. Considering the diversity of local DNA structure, some DNA sections with complementary sequences are bound to act as preferential false targets for a repair system (i.e., as « illusory defects »). However, if the repair system never changes the sequence upon repair, it will be solicited again and again by the illusory defect, a potentially harmful situation. It is therefore advantageous for a repair system to be, to some extent, error prone. Strong illusory defects may arise at the decanucleotide level and could be the cause of local increases in mutation levels. They might be used to initiate somatic hypermutation pathways.

Illusory defects may provide a class of avoided sequences, that may perhaps be revealed one day, using bioinformatic tools. Actually,

Hodgkinson, Ladoudakis and Eyre-Walker did publish a sequence analysis paper [53], in which they demonstrated exactly the kind phenomena that I had predicted in my illusory mismatch article. I wrote to Eyre-Walker, pointing out this coincidence, but he did not reply.

2.6 The blind spots of population genetics.

In 2009, Susan Rosenberg then editor at PLoS Genetics invited me to write a short review for this journal. I wrote a pamphlet on the technical and conceptual limitations of populations genetics, and in particular its blindness to potential genetic mechanisms mediating innovative evolution. The article was edited by Ivan Matic and published with the title “Frail hypotheses in evolutionary biology”. Since it is freely available online, I do not discuss the details. I merely provide here the subheadings: Smart evolutionary devices?/ On mutation and fixation rates/ The multiple origins of point mutations/ Phenotypic versatility and innovative evolution , and repeat here the introductory paragraph:

Introduction to “Frail hypothesis in evolutionary biology”.
In the last decades, under the headings of “mutation strategies”, “evolvability” or “soft inheritance”, many ideas have been advanced on mechanisms assumed to promote innovative evolution beyond what one may anticipate from the classical model of random mutation and selection. Many population geneticists find these ideas superficially seducing but mathematically unfounded. While agreeing with the need to critically evaluate such proposals in the light of population genetics, I will argue that population geneticists are not immune to criticism. For instance, the “infinite site model” introduced by Kimura makes the unrealistic assumption that any neutral mutation arises only once during a neutral fixation episode, that leads, I propose, to an underestimation of the neutral fixation rates in large populations. Critical parameters such as mutation and recombination rates, effective population sizes or beneficial/deleterious mutation ratios are assigned convenient values, that may seem *ad hoc* to people outside the field. The lack of concern for the subtleties of genetic mechanisms is also criticized. Phenomena such as compensatory mutations, recurrent mutations, hot spots, polymorphism, which population geneticists treat in the mathematical context of neutral versus selective fixations can instead be interpreted in terms of genetic mechanisms for producing complex mutational events.

Finally, single nucleotide substitutions are often treated as the quasi-exclusive source of variations, yet they cannot help much once the genes are optimized with respect to these substitutions. I suggest that population geneticists should invest more effort in refining the numerical values of the critical parameters used in their models. They should take into account the recent proposals on how mutations arise. They should also pay more attention to phenotypic variations, and develop criteria to discriminate between proposed evolutionary mechanisms that can actually work, and others that cannot.

---

3. MOLECULAR APPROACHES TO EVOLUTION

Early in 1977, while I was on a training course in Michel Imbert's visual neurophysiology laboratory, I received a telephone call from an editor in Masson's publishing house. Masson was a venerable and powerful scientific publisher. Gérard Lucotte, acting as an editor, was the director of a collection of books on evolution. Lucotte proposed to me to write a book on molecular evolution. We met in Masson's headquarters, Lucotte said that Masson would initiate a new policy with my book, making it cheap in order to reach the widest audience. By then, I had written on the genetic code (e.g., [54) and on the origin of life [55]). I had also spent much time, during my thesis years, looking at protein sequences and comparing them, so I had some familiarity with phylogenetic reconstructions. But I did not belong in any way to the molecular evolution community. Nevertheless, I determined rapidly what I wished to include in a book on molecular evolution, proposed a table of contents, and the contract was signed rapidly.

In June 1977, I went to Stockholm for a stay of two months in Rudolf Rigler's laboratory, at the Karolinska Institute. Charles Kurland was the initiator of this training course. I shared an office with Måns Ehrenberg. We had frequent discussions on kinetic models in relation to accuracy, and I was thinking on several other topics in parallel. I bought a nice correspondence writing pad with 200 thin A4 sheets of thin, crisp white paper, and started writing the first draft of my book on molecular evolution when I was out of the lab. I would write in the evenings, or during week-ends, usually in the open air, and close to the water. In two months, I had finished the first draft, going to the end of the correspondence pad. There were plenty of details missing. Essentially, I had laid down the thread of the reasonnings. I am very fond of this pad, entirely hand-written.

After my return in France, in July 1977, I started making a second draft. This time I used a cheap and noisy but robust « Brother » typewriting machine. When the second draft was finished, I asked a few friends of whom I appreciated the critical mind to read the book, and have it discussed collectively. The friends were Pierre Roubaud, Jerôme Lavergne, Jean-Pierre Dumas, Michel Volovitch. We met in the luxurious apartment of Jeanine Rondest, then journalist in the popular science magazine “La Recherche”, in the presence of Jeanine Rondest and Gérard Lucotte. The reactions were positive, but Gérard Lucotte insisted on the necessity to condense the MS to resize it exactly to the length decided in the contract, without tolerance. So, I went into another cycle of writing, condensing it as much as I could without sacrificing anything important, with just one exception:

There was a chapter in the book describing what could be expected in “heterologous” systems, for instance when one was mixing the components of two translation apparatuses. I developed the thesis already presented in [56] that in the case of organelles, a smaller number of tRNAs would be used, with a more degenerate pattern of codon readings. Pushing the idea one step further, I indicated in a paragraph that such conditions could be favourable to an evolution of the genetic code. Alas, I dropped this paragraph in the last round of condensations. When the first deviation from the universal code was described for human mitochondria [57], I reproached myself having used the scissors. Only much later, when deviations in the nuclear code were discovered (e.g., [58]) I became less disappointed because I had not anticipated at all the variations in the nuclear code. Nevertheless, I thought that the deviations from non-universality in the nuclear code might have to do with problems of interferences between two translations apparatuses, that of the host with that of a symbiont, an invader, or an engulfed prey [59].

The construction of the book was unconventional. I did not follow a historical order (from the origins of life to man) and did not segregate the disciplines (molecular biology, population genetics…), but spread over the chapters the burden of acquiring elementary knowledge (for the laymen) and the pleasure of distillating new concepts. The turning point of the book was chapter 12 on “the sequence space”. John Maynard Smith had written pertinently on the concept of a “protein space” in 1968 [60]. This reference came to my attention reading Richard Grantham’s thesis, or one of his articles, and I was immediately seduced by the concept. This concept allowed me to connect ideas on molecular evolution – “a walk in the sequence space” [61], with ideas on accuracy

(e.g., protein synthesis errors as extending a gene to a whole neighbourhood of the gene in the sequence space). From there, I could speculate on the general properties of the regions of the sequence space in which evolution would lead us. All this was laid down and discussed in a quite innovative way, at a time at which the concept of sequence space had hardly caught the attention of the biologists, Michael Conrad [62]) being a rare exception.

The book appeared in French version in February 1979 [17]. Charles Kurland who was then advisor to the English publisher Pitman suggested to them to have the book translated, that was done excellently by Richard Lang, a biophysicist at the National Institute for Medical Research, Mill Hill, London. Pitman then sold the reproduction rights to Princeton University Press. I made a few changes for the Princeton Edition to reflect new discoveries in molecular biology , and the Princeton edition appeared in 1983, followed by a Japanese edition (Kinokunya press, translation by K. Nagano, a specialist in the origins of life) in 1984.

Although the background knowledge in molecular biology, genetics and evolution has much expanded since the publication of the book, I feel that the concepts in the book remain quite modern, and quite advanced with respect to much of what we read to-day on the same topics. Year after year, I see “original” theoretical papers, published in the best journals, that look like long developments on the themes of various ideas I had outlined in the book in a single paragraph. The book was influential, considering all the more recent work that connect with the ideas in the book. On the other hand, it was not quoted often, possibly due to the presence, in the book, of a harsh criticism of some high ranked scientists, most notably, Manfred Eigen.

The book was reviewed by several prominent scientists, including Francisco Ayala and Sidney Brenner. Samuel Karlin liked it, and invited me to an extraordinary congress that he organized in Israel with Eviatar Nevo. At this congress, I had the occasion to meet most of the scientists who had a name in molecular evolution. I shared a room with Walter Fitch, then with Gabriel Dover, talked to Tomoko Ohta, Susumi Ohno, Robert May, and many others.

The kindest review was an anonymous one, written in 1999, on the American Amazon books site. It runs as follows:

Review of “Molecular approach to evolution”
This is a seminal book in molecular evolutionary theory.,
September 26, 1999, By A Customer

“Ninio's "Molecular Approaches to Evolution" helped establish the concept of "sequence space" and encouraged rigorous, elegant, but fun, approaches to the study of molecular evolution long before it became a fad. The prose makes clear extremely difficult topics in a delightful, and deeply satisfying way. But be careful - reading this book is likely to make you a zealous molecular evolutionist.”

I would have loved to update the book year after year, but there was a single French edition. The book was put on the market with an indecently high price, ruining my hopes to have it read by a wide audience, and leading me to a very painful conflict with the French publisher. Then, in 2000, I started to write a new version, directly in English, to be published by Wiley. While I had nearly completed the second draft of the book, the United States were shaken by the September 11th, 2001 events, and closed themselves to external influences. From the stored material, I extracted two chapters that became reviews, the first on prebiotic replication and catalysis [63], the second on the hidden assumptions of population genetics [14] – see Section 2.6.

REFERENCES

[1] Ninio, J. (1991) Transient mutators: a semi-quantitative analysis of the influence of translation and transcription errors on mutation rates. Genetics 129, 957-962.

[2] Ninio, J. (1996) Gene conversion as a focusing mechanism for correlated mutations: a hypothesis. Molecular and General Genetics 251, 503-508.

[3] Amos, W. (2009) Heterozygosity and mutation rate: evidence for an interaction and its implications. BioEssays 32, 82-90.

William (Bill) Amos is professor of evolutionary genetics, department of zoology, Cambridge University.
The idea of « heterozygosity induced » substitution mutations or « HI mutations » was deduced from his experimental results. It extended logically his findings on microsatellite instabilities. The molecular mechanism behind the phenomenon was not yet clear. Error-prone DNA resynthesis close to a region of heterozygosity, as postulated in the gene

conversion model (Ninio, 1996 ref. 2) was unknown to William Amos, but compatible with his findings.

[4] Amos, W. (2013) Variation in heterozygosity predicts variation in human substitution rates between populations, individuals and genomic regions. PLoS ONE 8(4), e63048.

[5] To be filled later.

[6] S. Yang, L. Wang, J. Huang, X. Zhang, Y. Yuan, J.-Q. Chen, L. D. Hurst, and D. Tian (2015) Parent-progeny sequencing indicates higher higher mutation rates in heterozygotes. Nature 523, 463-467.

This is a massive genome sequencing work, performed by a team of the State Key Laboratory of Pharmaceutical Biotechnology, School of Life Sciences, Nanjing University. Laurence D. Hurst who cosigned the paper (how he participated to the work is not mentioned) is Professor of evolutionary genetics in the Department of Biology and Biochemistry at the University of Bath. He is also heading the recently founded the Milner Centre for Evolution. The article published in Nature provides excellent support for the concept of mutations induced by heterozygosity (Amos, 2009, ref. 3) and even, more precisely, for the class of molecular mechanisms postulated by Ninio, 1996 in Molecular and General Genetics (here, reference 2). Furthermore, there is a sentence in that they speak of the autocatalytic character of polymorphism, , in agreement with Ninio, 1996 and Amos, 2009.

[6a] Arbeithuber B., Betancourt, A.J., Ebner, T., and Tiemann-Boege, (2015) Crossovers are associated with mutation and biased gene conversion at recombination hotspots. Proc. Nat. Acad. Sci. USA 112, 2109-2214.

This article presents detailed and convincing evidence for gene-conversion induced mutations, but neither Ninio (1996) nor Amos (2009) are quoted. The authors are at the Institute of Biophysics, Johannes Kepler University, Linz, Austria, the institut für Populationsgenetik, Vetmeduni Vienna, the department of Gynecological Endocrinology, Linz, and the Institute of Human Genetics, Medical Univsersity of Graz, Austria. The article was edited for PNAS by …. Laurence Hurst, and published in the February 17th issue.

[7] Ninio, J. and Bokor, V. (1986) Stratégies d'adaptation moléculaire. La vie des Sciences 3, 121-136.

[8] Ninio, J. (1986) Diversity and accuracy in evolution: sketches past, present and future. In *Accuracy in Molecular Processes* (Kirkwood. T.B.L., Rosenberger, R. and Galas, D.J., eds) Chapman & Hall, London, pp. 381-391.

[9] Baldwin, J.M. (1896) A new factor in evolution. American Naturalist, 30, 441-451 et 536-553. See also Piaget, J. (1974) Adaptation vitale et psychologie de l'intelligence. Hermann, Paris.

[10] Bateson, P. (2006) The adaptability driver: Links between behaviour and evolution. Biological theory 1, 342-345.

[11] Radman, M., Matic, I. and Taddei, F. (1999) Evolution of evolvability. Ann. New York Acad. Sci. 870, 146-155.

[12] Fox Keller, E. (1999) Regulation of gene variation. Ann. New York Acad. Sci. 870, 290-291.

[13] Caporale, L. (2000) Mutation is modulated: implications for evolution. BioEssays 22, 388-395.

[14] Ninio, J. (2010) Frail hypotheses in evolutionary biology. PloS Genetics 6, issue 9, e1001067

[15] Ninio, J. (1991) Connections between translation, transcription and replication error-rates. Biochimie 73, 1517-1523.

[16] Ninio, J. (1997) The evolutionary design of error-rates and the fast fixation enigma. Origins of Life and Evolution of the Biosphere 27, 609-621.

[17] Ninio, J. (1983) Molecular approaches to evolution. Princeton University Press, Princeton, New-Jersey. Also : (1982) Pitman, London (1979) Approches moléculaires de l'évolution, Masson, Paris (1984) Japanese translation, Kinokunya Press, Tokyo.

[18] Drake, J.W. (1991) A constant rate of spontaneous mutation in DNA-based microbes. Proc. Nat. Acad. Sci. USA 88, 7160-7164.

[19] Drake, J.W., Charlesworth, B., Charlesworth, D., and Crow, J.F. (1998) Rates of spontaneous mutations. Genetics 148, 1667 – 1686.

[20] Kimura, M. (1968) Evolutionary rate at the molecular level. Nature 217, 624-626. Kimura, M. and Ohta, T. (1971) Protein polymorphism as a phase of molecular evolution. Nature 229, 467-469. Ohta, T. (1973) Slightly deleterious substitutions in evolution. Nature 246, 96-98. Kimura, M. (1983) The neutral theory of molecular evolution. Cambridge University Press, London.

[21] Stephan, W. and Kirby, D.A. (1993) RNA folding in drosophila shows a distance effect for compensatory fitness interactions. Genetics 135, 97-103. Piskol, R., and Stephan, W. (2008) Analyzing the evolution of RNA secondary structures in vertebrate introns using Kimura's model of compensatory fitness interactions. Molecular Biology and Evolution. 25, 2483-2492.

[22] Merlo, L.M.F., Lunzer, M. and Dean, A.M. (2007) An empirical test of the concomitantly variable codon hypothesis. Proc. Nat. Acad. Sci. USA 104, 10938-10943.

[23] Kondrashov, A.S., Sunyaev, S., and Kondrashov, F. A. (2002) Dobzhansky-Muller incompatibilities in protein evolution. Proc. Nat. Acad. Sci. USA 99, 14878-14883.

[24] Davis, B.H., Poon, A.F.Y., and Whitlock, M.C. (2009) Compensatory mutations are repeatable and clustered within proteins. Proc. Biol. Sci. 276, 1823-1827.

[24a] Callahan, B., Neher, R.A., Bachtrog, D., Andolfatto, P., and Schraiman, B. (2011) Correlated evolution of nearby residues in drosophilid proteins. PloS Genetics 7 (2) e1001315.

[25] Andersson, S.G.E. and Kurland, C.G. (1990) Codon preferences in free-living organisms. Microbiological Reviews 54, 198-210.

[26] Duret, L. (2000) tRNA gene number and codon usage in the *C. elegans* genome are co-adapted for optimal translation of highly expressed genes. Trends in Genetics 287-289.

[27] Rocha, E.P.C. (2004) Codon usage bias from tRNA’s point of view: Redundancy, specialization, and efficient decoding for translation

optimization. Genome Research 14, 2279-2286.

[28a] Kacser, H., and Burns, J. A. (1981) The molecular basis of dominance. Genetics 97, 639-666.

[28b] Cornish-Bowden, A. (1987) Dominance is not inevitable. Journal of Theoretical Biology 125, 333-338.

[29] Cairns, J., Overbaugh, J. and Miller, S. (1988) The origin of mutants. Nature 335, 142-145.

[30] Hall, B.G. (1990) Spontaneous point mutations that occur more often when advantageous than when neutral. Genetics 126, 5-16.

[31] Hall, B.G. (1991) Adaptive evolution that requires multiple spontaneous mutations : mutations involving base substitutions. Proc. Nat. Acad. Sci. USA 88, 5882-5886.

[32] Drake, J.W. (2007) Mutations in clusters and showers. Proc. Nat. Acad. Sci. USA 104, 8203-8204.

[33] Drake J.W. (2007) Too many mutants with multiple mutations. Crit. Rev. in Biochem. and Mol. Biol. 42, 247-258.

[34] Loeb, L. (1991) Mutator phenotype may be required for multistage carcinogenesis. Cancer Research 51, 3075-3079.

[35] Timms, A.R., and Bridges, B.A. (1993) Double, independent mutation events in the *rpsL* gene of *Escherichia coli*: an example of hypermutability? Molecular Microbiology 9, 335-342.

[36] Slupska, M. M., Baikalov, C., Lloyd, R., and Miller, J.H. (1996) Mutator tRNAs are encoded by the Escherichia coli mutator genes mutA and mutC: a novel pathway for mutagenesis. Proc. Nat. Acad. Sci. USA 93, 4380-4385.

[37] Taddei, F., Radman, M., Maynard-Smith, J., Toupance, B., Gouyon, P.-H., and Godelle, B. (1997) Role of mutator alleles in adaptive evolution. Nature 387, 700-702.

[38] Tenaillon, O., Toupance, B., Le Nagard, H., Taddei, F., Godelle, B. (1999) Mutators, population size, adaptive landscape and asexual

populations of bacteria. Genetics 152, 485-493.

[39] Lynch, M., and Abegg, A. (2010) The rate of establishment of complex adaptations. Molecular Biology and Evolution 27, 1404-1414.

[40] Eyre-Walker, A. and Keightley, P.D. (1999) High genomic deleterious mutation rates in hominids. Nature 397, 344-347.

[41] Crow, J. (1999) The odds of losing at genetic roulette. Nature 397-294-294.

[41a] Ninio, J. (1999) La fin de la lignée humaine ? Pour la science 262, 106-107.

[42] Woodruff, R.C., and Thompson jr, J.N. (2005) The fundamental theorem of neutral evolution: rates of substitution and mutation should factor in premiotic clusters. Genetica 125, 333-339.

[43] Holliday, R. (1964) A mechanism for gene conversion in fungi. *Genet. Res*. 5, 282-304.

[44] Cher Jacques,
Merci pour ton envoi. Je ne suis pas trop surpris par cette volée de bois vert, car je pensais que cette rédaction était prématurée. Au moins avons-nous récupéré le nom de quelques auteurs à lire absolument. Amicalement, Michel.

[45] Giver, C.R. and Grosovsky, A.J. (1997) "Single and coincident intragenic mutations attributable to gene conversion in a human cell line" *Genetics* 146, 1429-1439.

[46] Carrasquillo, M.M., Zlotogora, J., Barges, S. and Chakravarti, A. (1997) Two different connexin 26 mutations in an inbred kindred segregating non-syndromic recessive deafness: implications for genetic studies in isolated populations. Human Molecular Genetics 6, 2163-2172.

[47] Feingold, J. (1998) Mutations multiples d’un même gène dans les petites populations. Comptes-Rendus Acad. Sci., Sciences de la vie, 321, 553-555.

[47a] Patrinos, G.P., Kollia, P., Loutradi-Anagnostou, A., Loukopoulos,

D., and Papadakis, M. (1998) The Cretan type of non-deletional persistence hereditary persistence of fetal hemoglobin [Agamma-158C--> T] results from two independent gene conversion events. Human Genetics 102, 629-634.

[48] Strathern, J.N., Shafer, B.K. and McGill, C.B. (1995) DNA synthesis errors associated with double-strand break repair. Genetics 140, 965-972.

[49] Kuzminov, A. (1995) Collapse and repair of replication forks in *Escherichia coli*. Mol. Microbiol. 16, 373-384.

[50] Berdal, K.G., Johansen, R.F., and Seeberg, E. (1998) Release of normal bases from intact DNA by a native DNA repair enzyme. EMBO J. 17, 363-367.

[51]. See my web chapter on « Adventures with journal editors » in the « Miscellanea » section.

[52] Ninio, J. (2000) Illusory defects and mismatches : Why must DNA repair always be (slightly) error prone ? BioEssays 22, 396-401.

[53] Hodgkinson, A., Ladoukakis, E., and Eyre-Walker, A. (2009) Cryptic variation in the human mutation rate. PLoS Biology 7: e1000027.

[54] Ninio, J. (1973) Recognition in nucleic acids and the anticodon families. Progress in Nucleic Acids Res. Mol. Biol. 13, 331-337.

[55] Ninio, J. (1976) L'évolutionnisme et les origines de la vie. La Recherche 7, 325-334.

[56] Ninio, J. (1975) Considerations on the problem of the joint evolution of two different translation apparatuses within the same cell. In *Molecular biology of nucleocytoplasmic relationships* (Puiseux-Dao, C., ed.) pp. 31-39, Elsevier, Amsterdam. This chapter ends with an allusion to the evolution of the code, quoting Crick who himself was quoting Brenner.

[57] Barrell, B.G., Bankier, A.T. and Drouin, J. (1979) A different genetic code in human mitochondria. Nature 282, 189-194.

[58] Caron, F. and Meyer, E. (1985) Does *Paramecium primaurelia* use a different genetic code in its macronucleus? Nature 314, 185-188.

[59] Ninio, J. (1986) Divergence in the genetic code. Biochemical Systematics and Ecology 14, 455-457. Ninio, J. (1990) The revised genetic code. Origins of Life and Evolution of the Biosphere 20, 167-171.

[60] Maynard Smith, J. (1970) Natural selection and the concept of a protein space. Nature 225, 563-564.

[61] Ninio, J. (1981) Molecular evolution: A walk in the sequence space. In Food, Nutrition and Evolution (Walcher, D.N., and Kretschmer, N., eds) Masson Publishing, New-York, pp. 9-13.

[62] Conrad, M. (1978) Evolution of the adaptive landscape. In Theoretical approaches to complex systems. (Eds R. Heim and G. Palm). Lecture notes in mathematics, vol. 21, Springer-Verlag, Berlin, pp. 147-169.

[63]. Ninio, J. (2007) Errors and alternatives in prebiotic replication and catalysis. Chemistry & Biodiversity 4, 622-632 – see the forthcoming update of the web chapter on the origins of life.