Volume 285, number 2, 160–164 FEBS 09971 July 1991

ADONIS 001457939100681L

Minireview

On the origins of genetic variants

Ulrike Wintersberger

Department of Molecular Genetics, Institute for Tumorbiology and Cancer Research, University of Vienna, Borschkegasse 8a, A-1090 Wien, Austria

Received 16 May 1991

Two contrasting mechanisms responsible for the creation of genetic variants are described: one is the manifestation of the *limited* accuracy of the cellular machinery for DNA replication, the other results from the *ability* of cells to repair damaged DNA. *Replication-dependent variants* and those caused by episodical DNA damage enhance the probability that a small fraction of a cell population may survive a sudden (physical or biological) change of environmental conditions. *Replication-independent variants* arise during persistent but not immediately lethal stress (e.g. starvation) of a non-dividing population. The variants observed under these conditions are of selective advantage because they are able to cope with the particular stress situation. The molecular basis of their creation is a matter of intensive debate.

DNA replication fidelity; Fluctuation assay; Mutation; Gene conversion; Selection; Evolution

1. INTRODUCTION

The title of this review is partly borrowed from that of an article by Cairns and coworkers [1] which since its publication has provoked a stimulating scientific debate (see e.g. [2] and references therein). This debate arose as described below.

In the early 1940s when the biochemistry of the genetic material still was a mystery, Luria and Delbrück invented the 'fluctuation assay' to answer the question of whether bacterial mutants may arise spontaneously or specifically in response to a particular kind of stress [3]. A large population of bacteria sensitive to a certain strain of bacteriophages was divided into small subpopulations which were allowed to grow separately for several generations without any selective pressure before they were plated on Petri dishes containing bacteriophages. By means of statistical arguments it was shown that the colonies of phage-resistent bacteria which developed on the plates had arisen spontaneously during the prior growth of the small populations and not on the plates in response to the presence of the phages. From this classical experiment it was concluded that *genetic variants do originate randomly* and without selective pressure. However, the experiment was inappropriate to exclude the possibility that *in addition* to the spontaneous mutants, *genetic variants might be induced by selective pressure*. To test for that possibility Cairns and colleagues [1] applied experimental conditions which allowed mutants to emerge during a long period of nutritional stress. Again by statistically analyzing their data these authors claimed – to the surprise of parts of the scientific community – that a considerable portion of the advantageous mutations which they observed have taken place under the selective pressure of starvation. Molecular mechanisms for the creation of the 'directed mutations', as they were called, still remain to be detected.

With the present knowledge of the biochemistry of the genetic material and its replication, the mechanisms which had generated the Luria-Delbrück variants are much easier to imagine: they most likely arose from errors of DNA replication.

2. GENETIC VARIANTS RESULTING FROM INTRINSIC INFIDELITY OF GENOME DUPLICATION

Perfect reproduction is the essence of any living *individual*, whether it is a simple or complex organism, a virus or some virus-like entity. Because the accuracy of genome replication is limited by costs of increasing the precision of the copy-mechanism, perfect reproduction is, however, beyond reach. On the other hand, the Luria-Delbrück experiment demonstrates that natural *populations* are able to survive unexpected catastrophes *only because they allow variability* among the in-

Correspondence address: U. Wintersberger, Department of Molecular Genetics, Institute for Tumorbiology and Cancer Research, University of Vienna, Borschkegasse 8a, A-1090 Wien, Austria. Fax: (43) (222) 430790.

Published by Elsevier Science Publishers B.V.

dividuals at the expense of genetic identity between parents and progeny. Hence, an optimal balance between accuracy of genome duplication and allowance of variations may be decisive for the evolutionary success of a population.

The degree of variation between parents and progeny differs enormously with different forms of life. Within one form of life it depends on the stability of living conditions. Under stable conditions, undisturbed by any physical or biological threats, variants predominantly result from replication errors and for the following considerations the term 'spontaneous' variants will be confined to these kinds of mutants.

Template directed nucleotide-chain synthesis can proceed without enzyme catalysis. Under this condition the insertion of a nucleotide containing the correct base (according to the Watson–Crick rule of complementarity) relies exclusively on the difference in free energy ΔG between the Watson–Crick type base-pairing and other possibilities of pairing, and the error frequency amounts to about one mispaired nucleotide per 10 to 100 nucleotides incorporated. Three biological mechanisms are known which ensure an enormously higher accuracy of nucleic acid-replication. First, insertion of the next nucleotide at the growing end of the polynucleotide chain (in the 5′→ 3′ direction) is catalyzed by template-dependent polymerases which, in addition to catalyzing the formation of the phosphodiester bond, substantially amplify the discrimination between the correct nucleotide and the three incorrect ones; second, a non-complementary last nucleotide is removed preferentially by a 3′→5′ exonuclease activity (proofreading); and third, non-complementary bases can be excised by the mismatch repair mechanisms even after further nucleotides have been added to the growing chain. Whereas accuracy of RNA polymerization, according to present knowledge, exclusively depends on the fidelity of the polymerization reaction, DNA synthesis is considerably refined by proofreading and mismatch repair.

In vitro tests (using synthetic templates) and various assay systems, combining an in vitro enzyme reaction with testing the resulting polymerization product in vivo, have been worked out for the determination of the replication-fidelity of DNA polymerases and of complexes of DNA polymerases with accessory proteins [4]. Test systems which in addition to misincorporation of single bases detect other classes of mistakes of DNA polymerases, like frame shifts, small deletions etc., are also available [5]. In these tests the misincorporation rates of different purified DNA polymerases (separate from, or in combination with accessory proteins) vary over several orders of magnitude.

For example, the two DNA polymerases participating in the replication of the *Escherichia coli* chromosome (in vivo error frequency is 10^{-9}–10^{-10} per nucleotide incorporated), DNA polymerase III (Pol III) and DNA polymerase I (Pol I) are quite distinct from each other: Pol III is a multi-subunit enzyme and Pol I consists of a single polypeptide chain. The enzymatic activities, DNA-dependent DNA polymerization and 3′→5′ exonucleolytic phosphodiester-bond hydrolysis, the proof reading activity, reside in two separate subunits of Pol III and in two domains of the Pol I polypeptide. In the combined in vitro–in vivo test Pol III, the enzyme responsible for the majority of chromosomal replication, exhibited a misincorporation-rate lower than $1/10^7$ and Pol I which synthesizes only short stretches during DNA replication, a rate of about $1/10^6$. The in vitro assays showed that proofreading in the case of Pol III contributed a 100-fold, and in the case of Pol I a tenfold enhancement to accuracy [6]. Because polymerization and excision of nucleotides are competing processes their ratio not only depends on intrinsic enzymatic properties but also on the concentration of each of the four nucleoside-triphosphates, the so-called pool of nucleotides. High amounts of nucleoside-triphosphates will favour and speed up the polymerization reaction and thereby enhance the error rate; imbalance of the pool will lead to a further increase of mistakes. Logic expects that the replication of a small genome might tolerate a higher error rate than that of a large one. That this simple view does not always hold, is demonstrated by the bacteriophage T4 DNA polymerase (possessing a 3′→5′ exonuclease activity) which is more accurate than Pol III even though it is responsible for the replication of a comparatively small genome.

An enzyme, the fidelity of which has interested several research groups, is the reverse transcriptase of the human immunodeficiency virus (HIV). Like all reverse transcriptases (which most likely are in evolutionary terms the oldest DNA polymerases, see [7]) examined so far, the HIV-enzyme has no proofreading activity and exhibits a low fidelity in the DNA-dependent DNA polymerizing reaction, a finding probably explaining the rapid generation of viral variants [8–11]. Unfortunately the error-rates of reverse transcriptases during RNA-dependent DNA polymerization are still unknown.

Eukaryotic organisms possess four (yeast, and probably other lower eukaryotes) or five (higher eukaryotes) different DNA polymerases [12,13]. Those eukaryotic enzymes which have been tested in vitro show a much higher misincorporation rate than expected from spontaneous mutation rates in vivo [4], indicating that accessory proteins must have a decisive influence on the fidelity of eukaryotic genome replication. Interestingly, the primase activity (a DNA-dependent RNA polymerase) associated with calf-thymus DNA polymerase is the least accurate polymerase known [14]. As mistakes occurring during primer synthesis do not contribute to the mutation rate, cells need not invest in high accuracy of RNA primer synthesis.

3. GENETIC VARIANTS EMERGING IN RESPONSE TO DANGER OF LIFE

Having been accurately copied, a genome does not stay intact for ever. Nucleic acids are reactive molecules and therefore undergo structural changes under the influence of physical or chemical attack. To avoid the loss of the integrity of DNA genomes, cells are equipped with a battery of enzyme-systems for DNA repair [15,16], whereas for the evolutionarily older RNA genomes repair mechanisms have not yet been observed.

The obvious objective (or rather, justification) of the evolution of DNA repair is the restoration of the original status of the damaged genome-structure and the improvement of a cell's ability to survive. Nevertheless, some of the DNA repair processes are error-prone and thus the most important of all sources of genetic variants. Certain structural changes in DNA may mislead the DNA polymerase to catalyze the insertion of a non-complementary base, others may block the replication machinery, a life threatening situation to which *E. coli* cells respond by induction of the SOS-repair cascade. The SOS-repair system allows the restoration of DNA replication at the expense of reduced fidelity or even of completely random insertion of nucleotides if the template is unrecognizably damaged (see [17] for details and further references). Recombination-repair mechanisms responding to DNA breaks may patch up a genome in an order which is different from the original [18]. Naturally, a cell will prefer to survive as a variant to being killed as the original. Furthermore, like variants due to replication errors those induced by DNA damaging episodes will contribute to population survival by the increased repertoire of individuals fortuitously able to cope with unexpected environmental situations. In this sense damage to DNA, the most direct threat to life, induces evolution [19,20].

Our laboratory recently became interested in the dynamics of DNA damage-induced generation of variants. We started to study the kinetics of the appearance of genetic events during the proliferation of haploid yeast cell populations after attack by a DNA damaging agent. Because the frequencies of variants are usually determined with reference to the fraction of cells having survived the DNA damaging treatment, we first concentrated on the course of lethal events occurring during growth of the afflicted cell populations. We found that the simple idea of killed cells being diluted out of the population by the dividing survivors was fulfilled by X-irradiated yeast cells, but not by those treated with mutagenic chemicals, among which ethyl methanesulphonate and *N*-methyl-*N'*-nitrosoguanidine led to the most extreme deviations from the expected behaviour [21]. Under the influence of these latter drugs, populations contained many more non-viable individuals than expected after numerous generations of growth. A careful study of the progeny of individual cells (by pedigree analyses) showed that one source of delayed death was the ability for residual divisions of lethally mutated cells, which can be explained by stepwise dilution of the product encoded by the gene just mutated. To our surprise, a considerable fraction of dead cells, however, emerged newly from seemingly healthy ancestors [22]. Two different explanations for this result are possible. First, certain DNA lesions may have survived replication and during later generations may have been converted into lethal mutations. Presumably also non-lethal mutants, i.e. genetic variants, may hence arise during late post-treatment generations. Second, a mutagenic repair mechanism might stay induced over several population doublings and cause structural changes to the genome without the need for repair. To further investigate the dynamics of the formation of genetic variants we have started an additional series of experiments.

The fluctuation test, although originally devised by Luria and Delbrück to resolve the question of whether genetic variants arise randomly or as an adaptive response, has since been used for measuring the rates of genetic events in growing cell populations where these events are neutral and without adaptive advantage. This rate is defined as the number of genetic events per genome replicated and thus measures replication-dependent genome alterations. The basis for the determination of rates using data from a fluctuation assay is the mode of distribution of variants over a large number of small cell populations grown independently and then plated on selective medium, such that only the variants can develop into visible colonies (see Introduction). In the calculations as yet elaborated for the determination of a rate from the numbers of the colonies this rate has to be assumed as constant during the proliferation of the population under study (see e.g. [23,24]). However, we would be interested in the estimation of the time-course of a changing mutation- or recombination-rate of a cell population which had experienced a DNA damaging attack. In other words, we would like to discern the 'wave' of induced genome variability. Yet, no simple mode of mathematical treatment for varying rates is available. We therefore had to work out a computer simulation programme (Klein and Wintersberger, manuscript in preparation) which allows a search for rate kinetics which best fit the experimentally obtained distribution of variant-colonies. For a start we have studied the induction of an intrachromosomal gene-conversion in haploid yeast cell populations by a single X-irradiation or a single treatment with UV-light. Being aware of the limitations of the method we can conclude from the data we have obtained so far, that increased gene-conversion rates occur up to the fourth post-treatment population doubling [25]. In the same way we have examined the course

of a mutation rate (a reversion of a point-mutation) in haploid yeast. Again, the experimental data are consistent with the assumption that the UV-induced highly increased mutation rate lasts for about two-to-three generations. Thus, a single attack by X-ray or UV-light obviously, in haploid yeast, induces a short wave but not a prolonged state of enhanced generation of variants. The effects of chemical mutagens, especially of those which gave rise to the delayed occurrence of lethally mutated cells in the experiments described above, have yet to be investigated. Our suspicion is that certain structural alterations of DNA may persist during replication and may, not necessarily but only with a certain probability, be eventually transferred into bona fide mutations or rearrangements of DNA.

But, what should be the advantage of a prolonged state of genome instability for a cell population having already resumed growing, after the life threatening attack by the DNA damaging agent had been overcome? Is it not more probable that such a state should last during a period of *persistent danger of death*? Experiments clearly demonstrating that this is in fact the case were recently described by Hall [26]. This author observed that during prolonged incubation of tryptophan-dependent *E. coli* colonies on a medium lacking tryptophan, the number of revertants (to tryptophan independence) increased many-fold over that expected from the mutation rate as determined in growing cultures. Thus, starvation of bacterial cells inhibiting their proliferation and threatening the survival of the population had 'induced' replication-independent mutations of selective advantage! Furthermore the number of these mutants increased linearily with starvation time. One could imagine that under the stress of starvation, DNA breaks and gaps may occur which may entail the induction of the SOS DNA repair mechanism. That this so-called error-prone process, while repairing DNA under the condition of prolonged stress, may induce more genetic variants than absolutely necessary because of loss of instructive DNA sequences makes sense. The only choice is between evolution (by developing new variants) and extinction. What may be more difficult to swallow is the fact that in Hall's experiments the accumulation of the advantageous mutants did not simply reflect an increased overall mutation rate. Rigorous control experiments showed that mutants without selective advantage under the particular experimental conditions did not accumulate.

How could a non-dividing cell preferentially produce or retain useful mutations? Cairns and coworkers [1] suggested, among other possibilities, the reverse transcription of those variant mRNAs which in the pool of randomly occurring ones may be useful and somehow be stabilized. Induction of increased reverse transcription by environmental stress has been observed in *Drosophila* [27] and by UV-light in yeast [28], but not in *E. coli*. In the eukaryotic organisms the mobility of retroviral-like transposable elements was found enhanced and one could imagine that the reverse transcriptase encoded by these elements may accidentally copy also mRNAs. Transposition of retroviral-like sequences themselves is an important mechanism in evolution [29] and pseudogenes are evidence for reverse transcription and insertion into the genome of sequences derived from mRNA. What is so difficult to believe is that these processes may not work just randomly and accidentally but somehow in a direction towards adaptive advantage. A model suggested by Hall for the explanation of the *apparent* specificity of the stress-induced mutants is easier to accept in line with the present view about the abilities of a bacterial cell. During periods of nutritional (or perhaps also other forms of) stress, a fraction of cells might get into a hypermutable state because of extensive DNA damage. As long as a cell remains in this state many different mutations will occur randomly at a high rate. Only if among these mutations there is by chance a useful one, will the cell recover and resume metabolic and reproductive activity. Those hypermutable cells which experience only useless mutations will die and will not be registered by the observer. Because under persistently catastrophic conditions more and more cells will reach the hypermutable state (before they irreversibly lose viability) the replication-independent mutations will arise in a time-dependent mode. This model predicts that the surviving variants (because of their passage through the hypermutable state) have in addition to the selected mutation also an increased probability of harbouring other (neutral) mutations. Hall has shown earlier that multiple mutations do occur at high rates during severe stress [30]. But again, as mutations can be detected only in cells having survived the hypermutable state, those neutral mutations which are not in the 'same boat' as an advantageous one, disappear and the observer has the impression that they did not appear at all.

Thus, with the Hall model we, for the time being, escape the assumption of a directionality of the mutation event rescuing the bacterial cell from death by starvation. As saving the genome from being lost before having produced progeny is the vital task of a bacterial cell there might however have evolved still other mechanisms which are beyond our present imagination.

Acknowledgements: I thank Dorothea Büchinger for help with the literature search and Aloisia Kopp for preparing the manuscript. The work described from our own laboratory was supported by a grant from the 'Jubiläumsstiftung der Österreichischen Nationalbank' (grant 2993).

REFERENCES

[1] Cairns, J., Overbaugh, J. and Miller, S. (1988) Nature 335, 142–145.
[2] Stahl, F.W. (1990) Nature 346, p. 791.
[3] Luria, S. and Delbrück, M. (1943) Genetics 28, 491–511.

[4] Loeb, L.A. and Kunkel, T.A. (1982) Annu. Rev. Biochem. 51, 429-457.
[5] Kunkel, T.A. (1990) Biochemistry 29, 8003-8011.
[6] Kunkel, T.A. (1988) Cell 53, 837-840.
[7] Wintersberger, U. and Wintersberger, E. (1987) Trends Genetics 3, 198-202.
[8] Preston, B.D., Poiesz, B.J. and Loeb, L.A. (1988) Science 242, 1168-1171.
[9] Roberts, J.D., Bebenek, K. and Kunkel, T.A. (1988) Science 242, 1171-1173.
[10] Takeuchi, Y., Nagumo, T. and Hoshino, H. (1988) J. Virol. 62, 3900-3902.
[11] Bebenek, K., Abbotts, J., Roberts, J.D., Wilson, S.H. and Kunkel, T.A. (1989) J. Biol. Chem. 264, 16948-16956.
[12] Hamatake, R.K., Hasegawa, H., Clark, A.B., Bebenek, K., Kunkel, T.A. and Sugino, A. (1990) J. Biol. Chem. 265, 4072-4083.
[13] Thömmes, P. and Hübscher, U. (1990) Eur. J. Biochem. 194, 699-712.
[14] Zhang, S. and Grosse, F. (1990) J. Mol. Biol. 216, 475-479.
[15] Friedberg, E.C. (1984) DNA Repair, W.H. Freedman, New York.
[16] Sancar, A. and Sancar, G.B. (1988) Annu. Rev. Biochem. 57, 29-67.
[17] Echols, H. and Goodman, M.F. (1990) Mutation Res. 236, 301-311.
[18] Oishi, M. (1988) in: The Recombination of Genetic Material (K. Brooks Low ed.) pp 445-491, Academic Press, London.
[19] Echols, H. (1981) Cell 25, 1-2.
[20] Wintersberger, U. (1984) Adv. Enzyme Regul. 22, 311-323.
[21] Klein, F., Karwan, A. and Wintersberger, U. (1989) Mutation Res. 210, 157-164.
[22] Klein, F., Karwan, A. and Wintersberger, U. (1990) Genetics 124, 57-65.
[23] Von Borstel, R.C. (1978) Meth. Cell Biol. 20, 1-24.
[24] Klein, F. and Wintersberger, U. (1988) Curr. Genet. 14, 355-362.
[25] Klein, F. (1987) PhD thesis, University of Vienna, Austria.
[26] Hall, B.G. (1990) Genetics 126, 5-16.
[27] Strand, D.J. and McDonald, J.F. (1985) Nucleic Acids Res. 13, 4401-4410.
[28] Rolfe, M., Spanos, A. and Banks, G. (1986) Nature 319, 339-340.
[29] McDonald, J.F. (1990) BioScience 40, 183-191.
[30] Hall, B.G., Betts, P.W. and Wootton, J.C. (1989) Genetics 123, 635-648.