

Are we Genomic Mosaics? Variations of the Genome of Somatic Cells can Contribute to Diversify our Phenotypes

P.A. Astolfi¹, F. Salamini² and V. Sgaramella^{*,2}

¹Department of Genetics and Microbiology, University of Pavia, Italy; ²PTP, Lodi, Italy

Abstract: Theoretical and experimental evidences support the hypothesis that the genomes and the epigenomes may be different in the somatic cells of complex organisms. In the genome, the differences range from single base substitutions to chromosome number; in the epigenome, they entail multiple postsynthetic modifications of the chromatin. Somatic genome variations (SGV) may accumulate during development in response both to genetic programs, which may differ from tissue to tissue, and to environmental stimuli, which are often undetected and generally irreproducible. SGV may jeopardize physiological cellular functions, but also create novel coding and regulatory sequences, to be exposed to intraorganismal Darwinian selection. Genomes acknowledged as comparatively poor in genes, such as humans', could thus increase their pristine informational endowment. A better understanding of SGV will contribute to basic issues such as the "nature vs nurture" dualism and the inheritance of acquired characters. On the applied side, they may explain the low yield of cloning *via* somatic cell nuclear transfer, provide clues to some of the problems associated with transdifferentiation, and interfere with individual DNA analysis. SGV may be unique in the different cells types and in the different developmental stages, and thus explain the several hundred gaps persisting in the human genomes "completed" so far. They may compound the variations associated to our epigenomes and make of each of us an "(epi)genomic" mosaic. An ensuing paradigm is the possibility that a single genome (the ephemeral one assembled at fertilization) has the capacity to generate several different brains in response to different environments.

Received on: April 22, 2010 - Revised on: May 01, 2010 - Accepted on: May 07, 2010

Keywords: Copy number variation, aneuploidy, genomics, epigenomics, transposons, retroposons.

"Where will genomics be ten years from now? As sequencing capacity increases globally and the data quality improves, we will move beyond the current goal of one genome per person by sequencing multiple genomes per person... The genomic revolution is just beginning." (J. C. Venter, *Nature* 464:676-677, 2010).

INTRODUCTION

It is known that the famous physicist Erwin Schrödinger, Nobelist in 1933 "for the discovery of new productive forms of atomic theory", pioneered the advent of molecular biology with a widely quoted booklet [1] which stimulated many colleagues of different expertise to converge on the study of DNA. Less known is a statement where Schrödinger serendipitously anticipated a violation of a tenet to be formulated after the unveiling of the DNA double helix. In particular Schrödinger stated: "...all the cells of our body are exactly identical for what pertains their chromosomes", and in a footnote added: "But biologists will forgive me if I do not mention the exceptional cases of mosaics". It is probable that the "mosaics" alluded to here were the products of intrauterine exchanges of embryonic cells between fraternal twins.

The tenet at risk of violation descended directly from the "central dogma of molecular biology", promulgated in the early 1950s by Francis Crick, codiscoverer of the double helix and admittedly one of the many physicists converted to molecular biology by Schrödinger. Crick's dogma prescribed

that in living cells the genetic information flows from DNA to RNA and hence to proteins, but not backwards; in the 1970s it had to be revised because of the discovery of reverse transcription of RNA into DNA. But by and large the dogma still stands as a cornerstone of the biosciences, especially with regard to the logically consequent invariability (a.k.a. continuity or stability) of the somatic genomes. A significant statement has appeared recently in a Nature commentary [2]: "Whereas the hard-wired genome is virtually identical in all of a person's roughly 250 different types of tissue and is essentially stable during that person's lifetime, the epigenome changes during development such that each cell type has its own characteristic set of marks. These marks change with age and may also change, possibly in a heritable way, in response to environmental stress".

It is also known that epigenomics studies the complex but reversible modifications of chromatin (*i. e.*, DNA plus the wealth of accompanying proteins and non-coding small RNA) [3], whereas genomics addresses the base sequences of the DNA present in the living cells. But if it is questionable that epigenomic marks are "directly dependent" on DNA sequence, as Ptashne *et al.* state in a recent note to Nature [4], the two are not unrelated: "In contrast to the DNA sequence which is identical in all tissues, the patterns of chromatin modifications and DNA methylation are tissue specific. Thus our genome contains two layers of informa-

*Address correspondence to this author at the Sezione di Genetica e Biologia Molecolare, Parco Tecnologico Padano, Via A. Einstein 1, c/o Cascina Codazza, 26900 Lodi (LO), Italy; Tel: ++39 0371 4662 647; Fax: ++39 0371 4662 349; E-mail: vittorio.sgaramella@tecnoparco.org

tion: the DNA sequence inherited by our parents which is fixed throughout life and identical to most extent in all the cells and tissues of our body, and chromatin and methylation patterns which are cell and tissue specific. It is clear that gene function and thus the phenotype could be influenced not only by the gene sequence but also by the epigenetic programming of the sequence.” [5]. Even more explicitly it has been stated that: “The proper development of multicellular organisms entails the distinct specification of disparate cell types. Despite having identical genomic sequences, different cell types exhibit substantially different profiles of gene expression, and their cellular identity must be conserved during somatic cell divisions. How then are cell-specific gene-expression patterns specified and maintained? It is now recognized that the key is ‘epigenetics’: the stable and heritable information that is distinct from DNA sequences and fostered by specialized mechanisms.” [6].

We propose here that there is a third layer of information, represented by the variations affecting the somatic genomes throughout development: the ensuing genomic mosaicism responds not only to epigenetic fluctuations often traceable to the environment [5-8], but also to genetic determinism, hence to both chance and necessity. For the transactions affecting DNA sequences we suggest the expression Somatic Genome Variations (or SGV), as preferable to the common Copy Number Variations (CNV), given that, *e. g.*, non-replicative transpositions and inversions are not *stricto sensu* CNV.

At present SGV are perceived no longer as syndromic to disease, but also as marks of differentiation. The acceptance of their existence, and even of their importance, has gone through the stepwise and haphazard additions of bits of evidence, but is steadily consolidating. The picture has been aptly summarized as follows: “Fortunately, in the face of ever-increasing level of variation being discovered in the human genome, there is at least one touchstone: the genomes of all the cells in a cancer-free individual are the same. However, even in this bedrock, a few cracks have appeared. For example, it is well-known that immune cells undergo rearrangements and deletions at immunoglobulin loci. Cases are also known of genomic mosaicism, even as far as aneuploidy in sub-population of cells [...]. But, aside from these peculiar exceptions, and barring mutations accumulated as part of the ageing process, the genomes in all of your cells are identical. Or are they?” [9]. In order to help answering this rhetorical question we must critically review the main notions supporting the invariability of the somatic genome: they include the extraordinary accuracy recognized to *in vitro* DNA replication, the stability of DNA borne by somatic and germinal nuclei, the constancy of the 2:1 ratio of their respective masses, the persistence of familiar traits through different generations. The seal to these notions was provided in the early 1960s by cloning *via* somatic cell nuclear transfer (SCNT) in amphibians: their main finding was that the somatic nuclei transferred into enucleated unfertilized or fertilized oocytes could produce healthy organisms; their main conclusion was that such nuclei should be considered totipotent. But those experiments had at least three caveats: (a) only embryonic and fetal cells would provide functional nuclei; those from adult cells would fail, at least in those early days; (b) yields have since remained close to 1% of

transferred nuclei, essentially regardless of their donor cells; (c) the resulting totipotency does not necessarily imply full identity of the involved genomes, as shown by the few clones grown to adulthood, which were taken as genotypically “identical” to the nuclei donors, but resulted phenotypically different, at least health-wise.

But the rhetorical question raised above by Dear [9] cogently underscores the complexities of the SGV issue. The debate over the invariability of the somatic genome is more than a century old: for a review see, *e. g.*, [10]. At the beginning it was thought that multicellular organisms would develop their manifold traits either through a differentially regulated expression of a constant genome (Delage, Spemann and others), or alternatively through genomic structural variations which would lead to different patterns of gene expression (Driesch, Weismann and others). Mutations were thought to occur in single cells, where they would disrupt gene structure and function, and thus occasionally trigger dramatic genomic alterations, uncontrolled growth of clonal progeny cells into anomalous and eventually diseased tissues [11]. Shortly, the notion of the variation of the somatic genome could be accepted, as long as it was associated to pathology. That under normal conditions all the somatic cells of an organism would harbor the same genome had become almost a second dogma: it predicated its invariability in the somatic cells of any complex organism throughout, and in spite of, its differentiation and development. But its tenet had to coexist with some contrasting evidences:

(a) in living cells the genetic information flows also from RNA back to DNA, and does so both spatially (*i.e.*, in different cells), and temporally (*i.e.*, at different developmental stages): the apostle of this tenet was Howard Temin, who studied RNA viruses and stated that “This normal process of information transfer in cells could not exist only for its ability to give rise to viruses. It must exist as a result of its role in normal cellular processes, for example cell differentiation, antibody formation and memory” [12]; remarkable prediction indeed, especially for what concerns memory and more in general cognitive functions. But still not enough to secure an official *imprimatur* to Temin’s doctrine even long after the subsequent extensive characterization of retroviruses [13] and in general of retroelements [14];

(b) the eukaryotic genomes host a variety of transposable elements (TE), at least in part responsible for improving the response of the organisms to the environment; the apostle of this tenet was Barbara McClintock, who clearly summarized her ideas as follows: “I believe that there is little reason to question the presence of innate systems that are able to restructure a genome. It is now necessary to learn of these systems and to determine why many of them are quiescent and remain so over long periods of time, only to be triggered into action by forms of stress, the consequences of which vary according to the nature of the challenges to be met. In addition to modifying gene action, these elements can restructure the genome at various levels, from small changes, involving a few nucleotides, to gross modifications, involving large segments of chromosomes, duplications, deficiencies, inversions and other more complex reorganizations.” [15] Both Temin and McClintock were eventually awarded fully deserved Nobel Prizes, but their colleagues are still taking a

while to accept their teachings, and hesitate to recognize that the integration of DNA copies of RNA sequences into the nuclear genome, as well as the chromosomal vagaries of TE, are incompatible with the stability of the somatic genomes and must lead to their restructuring.

Probably easier to be accepted are the concepts put forward by Zeh *et al.* [16] with regard to the causative relationship of epigenetics with regard to transpositions especially in evolution: “According to the *“epi-transposon hypothesis”*, physiological stress, associated with major climatic change or invasion of new habitats, disrupt epigenetic silencing, resulting in TE reactivation, increased TE expression and/or germ-line infection by exogenous retroviruses. Mobilized TE rapidly restructure the genome and alter gene expression patterns...”. However their correlations to development are steadily emerging, as revealed by the following quotations from Kazazian and Ostertag [17]: “Time and further research will tell whether McClintock’s hypothesis that TE have a significant role in an organism’s development can be extended from maize to humans, and specifically to the functions of human neurons.”. Discussing nervous system disorders, Lee & Lupski [18] display a more open view and hypothesize that “gene copy number variations due to structural alteration may be responsible for both normal and abnormal behavioral phenotypes”.

Many other instances of intraindividual somatic genome variations can be predicted at the theoretical level; several have indeed been described and they do not necessarily violate consolidated enunciations, since they act through recombination and similar notorious and legitimate transactions. But apart from mechanisms, their impact on the concept that all our cells harbour the same genome may threaten the central dogma and question some extant major projects.

STATE OF THE ART

It is still accepted that at each cellular division the structure of the genome as assembled at fertilization is faithfully duplicated together with the rest of the cell [6], apart from epigenetic changes and occasional ploidy variations [19]. In this way the genetic information is transmitted essentially unchanged to all the cells of our developing body, in its germinal as well as somatic lineages. Thus the phenotypes of the organism were seen as the gradual decoding of the information carried by a constant genome and expressed in the many features proper of our ~250 kinds of tissues, with all the subroutines of sophisticated programs of gene activation/inactivation. In cells where changes are the rule, one notable exception would therefore stand out: the genome, admittedly a fundamental element of our soma. Thus, throughout development the genome would remain unchanged. But several documented objections are raised against this view: they provide support to frequent, different and substantial manifestations of SGV, which go from single base substitutions or modifications to chromosome-scale events and may differ from traditional mutations since they are at least partly programmed. Some of the relevant objections to the invariability of the genome stem out of theoretical considerations: take, *e. g.*, the genome huge size (in terms of bp man has $\sim 6 \cdot 10^9$) which makes it an easy target for damages ($\sim 10^4$ per cell per day according to Jackson &

Bartek [20]), the number of cells contributing to the steady state of a human organism ($\sim 10^{14}$), the high number of mitoses ($\sim 10^{16}$) necessary for the organism’s full maturation and susceptible to errors and mishaps. Apart from these quantitative data, in the genome we have an extraordinary abundance of all sorts of transposable/unstable sequences, often exceeding 50% of the total DNA [16], and a reasonable confidence that most of these considerations apply to both animals [21] and plants [22]. Finally, there is the paradox of the low number of the human coding genes, 22,287 as acknowledged in 2006 by the Ensemble 34rd Catalogue of the IHGSC, as such close to those of a nematode [23] and to one third of a protist [24]; it cannot be excluded that such a gene paucity could be partially compensated through the intervention of SGV.

The strength of all these arguments contrasts with the weakness of the extant experimental data: beyond the classical cases of the gene reorganization controlling the immune response in vertebrates, of the macro/micro-nuclei transactions in protozoa, and of the telomere attritions punctuating the development of most eukaryotes, reports detailing tissue-specific and age-related variations in the genomes of eukaryotic cells are preciously few [24, 25].

Many recent data are, however, coalescing and paving the way to a desirable and feasible experimental characterization of SGV. One is the reported low success rate of cloning *via* SCNT, which is more compatible with rearrangements suffered by the relevant somatic cells genomes [26], than with technical hurdles [27]. Suggestive indications can be excerpted from recent genomic data: thus, the similarity of the estimated numbers of SGV in a given genome and of the gaps persisting in the sequences of the “quasi-completed” genomes available thus far [28, 29] point to an involvement of SGV. When the genomes of a mixture of cells of the same lineage of the same individual are being sequenced, at the SGV sites their genomes are likely to host heterogeneous sequences typical, or even unique, of each cell under study. It is in fact possible that some cells could bear alterations which result from untemplated, error-prone DNA repair at the sites of their occurrence, as in the assembly of the immunoglobulins and related genes. The genomic revolution [30], the unpredicted discoveries of systems similar to myelomas and the zooming on some of the most obvious of such sites will help to clarify the issue in the future. Advancements are expected from the increasingly powerful “next generation” sequencing tools. Discussing the interesting discovery that the jawless sea lamprey *Petromyzon marinus* loses close to 20% of their somatic genome during development, Eichler and colleagues underline the programmed nature of their loss which includes at least one transcribed gene, and conclude: “Thus, the loss of DNA could result in the generation of coding, promoter, enhancer, etc. sequences that impart unique functionality to somatic cell lineages.” [31].

As to their insurgence mechanisms, SGV may be triggered by genetic programs responding also to external stimuli/stresses [8, 32, 33], DNA replication/repair errors [34], exposure to variably processed reverse transcripts of viral genomes [35] and of pseudogenes [36], horizontal gene transfers [37], etc. In some cases the environmental circumstances may elicit a programmed cellular death (apoptosis). Alternatively, the somatic genomes may adapt to these

events *via* gene copy number variations, TE repositioning and representation, expansion or contraction of repeats (long or short, tandem or dispersed), chromosomal aberrations. Ultimate measures may be the loss of entire chromosomes and eventually the elimination of the whole nucleus, as in mammalian erythrocytes. All, *nota bene*, under the aegis of a rigorous somatic Darwinian selection.

Within the same organism different cell lineages, and possibly within the same lineages, single cells may display peculiar instances of SGV. Remarkably, cleavage stage human embryos are reported to suffer sudden dramatic insurgence of severe complex chromosomal aberrations, admittedly absent in the participating gametes [38, 39; see also the accompanying contribution of Robberecht *et al.* in this issue] and eventually responsible of undetected abortions of zygote-derived structures [40]; later on, the emerging germ lines may enjoy a peculiar protection against the insurgence of SGV as compared to the rest of the soma. Conversely, a preferential susceptibility to SGV may be found associated with cells positioned next to the boundary between environment and organism and progressing to a postmitotic stage, such as functional neurons [41, 42]. It has been proposed that the ensuing DNA rearrangements tend to occur close to the chromosomal sites involved in segmental duplications and to act as hotspots for further rearrangements and recombinations [43]. The genomic alterations caused by SGV may involve up to ~20% of the human genome [44]. This may affect many functions and support an inordinately increase in their frequency and size throughout development and aging [45]. An important message emerging from these data is that SGV may feature in supposedly “normal” cells and not only represent symptoms of extant or incumbent disease, from cancer to severe neurodegenerative disorders like Alzheimer’s disease and ataxia telangiectasia [46]. This last consideration gains credibility thanks to the blurred demarcation between physiology and pathology, ostensibly occurring in derangements of the immune system [47] and, probably, of the nervous system [40].

By altering the organization of the genome of affected cells, and hence their coding capacity as well as their regulation, SGV may account for a good share of the organismal phenotypic differences. This hypothesis is becoming attractive and in the last few years SGV have been enlisted within the traditional elements concurring to the shaping of the phenotype: these include the zygotic genome, as assembled at fertilization; the environment, in its macro-components, surrounding the whole organism as it develops, as well as its micro-components, surrounding each cell; and the epigenome, *i. e.*, the sum of the transient, inducible and partly inheritable modifications of the chromatin of the cells. These concepts have been elaborated in a graphical representation for which the term “ontogenic pyramid” has been proposed (Fig. 1), to suggest an idea and a model that in due course will probably undergo stringent criticism and re-elaboration.

A productive experimental effort aimed to the characterization of SGV must first detail out a broad strategic plan and then an accurate description of the operations in term of materials and methods. The optimal experimental organism(s) should display several useful traits, such as a complex developmental program leading to lineages as differentiated as

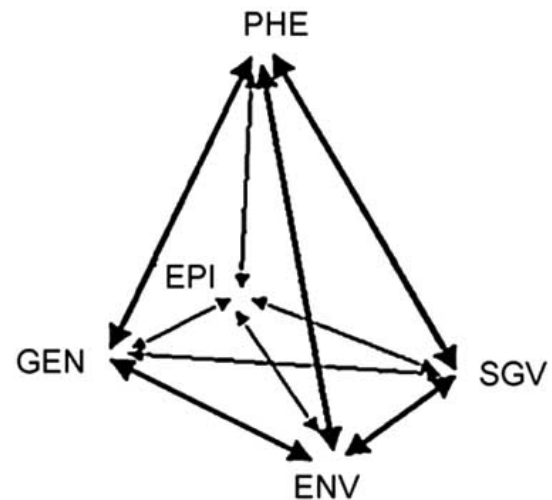


Fig. (1). The “ontogenic pyramid”. The phenotype (PHE) may be seen as the result of four variables: the original zygotic genome (GEN), the environments (ENV), the epigenome (EPI) and the somatic genome variations (SGV). The two headed arrows are meant to suggest complex reciprocal interactions between them. From Sgaramella & Astolfi [39], with permission.

possible; representative cells should derive from tissues originating from each of the three layers of an animal gastrula or from plant embryo structures; cells easy to be isolated as pure and uncontaminated as possible would be preferable; a reliable “reference” genome should be available for comparing the changes emerging during development. A reasonable prediction as to which sequences may be more amenable to SGV could first consider (retro)transposons, then elements such as LINES and SINEs, especially those sequences known or suspected to assume anomalous non-B conformations [44, 48]. It has been reported that, in the spectrum of the events responsible for human genome variations, deletions are more frequent than insertions or other rearrangements [49]: this is reasonable and may reflect a compensatory removal of sequences following synthetic (retro)transpositions otherwise responsible of an unbearable overgrowth of the genome size [50]. Relevant to this is the observation that during development: “*The set of genes deleted is not random in the (human) genome; there is a clear association with segmental duplications among larger deletion events. In addition to previously reported associations with immunity, defence, chemosensation and drug detoxification, other functional gene categories emerge in these studies, such as signal transduction, sex hormone metabolism and cancer susceptibility.*” [51]. Somehow ironically, indications for cells more susceptible to SGV may turn out to be a major contribution to science of the once much hyped organismal cloning: the finding that animal tissues (*e. g.*, neurons) are more recalcitrant than others (*e. g.*, cumulus) at producing mouse clones *via* SCNT could betray some tissue specificity of SGV [26, 52, 53] rather than technical hurdles. A second criterion for spotting cells more prone to SGV and thus more useful for their characterization, could be the extent of their epigenetic modifications: these changes control the structure of chromatin, and hence possibly vary the sequence of its DNA. This may occur through transcription,

then reverse transcription of some resulting RNA followed by the genomic integrations of their retrotranscripts as retroposons and pseudogenes [5, 16, 35, 54].

In a concerted approach to the study of SGV particular attention may go to: (1) their manifold nature, (2) their time-course during growth, (3) their molecular mechanisms, (4) their significance *vis-à-vis* the physiological as well as the pathological development of the organism, (5) their relevance on applied science as it regards, *e. g.*, reproductive and regenerative cloning, induction of pluripotency, DNA analysis, etc.

It would be desirable to focus on plant and animal models, better if endowed with relatively short life spans: this would facilitate the monitoring of the time course of SGV from early developmental stages to senescence and death. Among plants, a suitable system may be represented by *A. thaliana* [55], whereas among animals mice may be preferable, in spite of the somehow lower fraction of repeats in their genome [56]. Needless to say that humans are the ultimate objective.

From a practical viewpoint, it would be proper to study SGV in as few cells as possible, given the peculiar, if not the unique, nature of SGV. Single cell genome analysis would be highly desirable and may be reached in a not too distant future [57].

The study of SGV rests heavily on the use of efficient sequencing technologies [26], to be applied to whole genomes, or to parts selected because of their higher predisposition to SGV, as based either on prior knowledge, *e. g.*, retroposons, as discussed by Sen *et al.* [58], or on acquired experimental evidence (*e. g.*, gene duplication sites), as discussed by Rohlin *et al.* [30], and by many others [59-62]. Possible concomitant epigenetic changes are expected to augment the complexities of the experimental approaches to SGV characterization.

In a note appeared almost five years after the first “completion” of the Human Genome Project it was remarked that “*Uncovering the genetic basis of human phenotypic differences requires a comprehensive understanding of all forms of genetic variation*”, and the following explicit conclusions were reached: “*First, no single optimized approach has been developed yet to systematically capture all structural variation in the human genome. Second, it is likely that several thousand additional common structural variants await discovery. Their abundance, enrichment in environmental-interaction genes and their apparent attributes in terms of natural selection suggest that these genetic lesions will be important in genetic disease. A more systematic, unbiased approach to discover and genotype such variation is required. A Human Genome Structural Variation Project dedicated to the characterization of not only deletions but also of insertions and inversions should become a priority for the human genetics and human genome sequencing communities in an effort to further widen the spectrum of human genetic variation*” [51]. The relevance of these remarks has only been increasing in these last few years. If we were to place the adjective “somatic” before the substantives “genetics” and “genome”, as they recur in Eichler’s note, and to use the prefix “(epi-)” ahead of both of them, and if we were to en-

large the scope of his exhortations beyond disease to normal human traits, we would find a close match to the subject and the scope of this paper.

CONCLUSIONS

In shaping the physical, behavioural and cognitive human phenotypes, nature and nurture cooperate, probably not as independent entities. Rather they interact in a complex manner: accordingly, the genotype, the environmental stresses and/or stimuli, the epigenome [8] and finally SGV talk to each other so that eventually the genomes of any somatic cells may end up as unique patchworks and as such differ not only from the ephemeral genome of the zygote, but also among its many tissutal or even cellular versions [7, 43].

In a sense we are what we are because of our genes, but our genes are what they are because their roles are played in different macro- and micro-environments. Both the genes we inherit and the environments we are placed in result from random samplings out of countless alternatives. The dialogue between genes and environment dictates the necessity of our genetic/genomic mosaicism, its augmentation during development, its derangement during disease, possibly of its exaltation in forging our mind. Such interactions remind some of the ideas Lamarck published in 1809, the same year of Darwin’s birth [63]. This underscores the desirability of an analysis of the genome and the epigenome of different cell lineages at different developmental stages, and, possibly, at the single cell level, if we want to understand reliably the biologic determination of what we are genomically and epigenomically at any given point of our life. In this regard, somehow similarly to their colleagues physicists who in the last century integrated Heisenberg’s “indetermination principle” in the characterization of the nuclear elementary particles, today geneticists may be advised to consider an “indeterminability” of the genotype-phenotype relationship, given the ever-changing nature of both partners [23, 43, 64]. If all this is confirmed, when facing the ancient dilemma concerning the interactions of nature *vs* nurture in shaping our traits, we may find that the limits of the definition of the multifarious contributions of nurture (*i. e.*, diet, education, conscious and subconscious stresses and/or stimuli, etc) are no less severe than the contributions of nature (*i. e.*, our many possible genomes), as recently discussed by several authors [32, 33, 65]. The challenge becomes daunting when we envisage a role of SGV in shaping our neurons to produce just one brain (ours) out of the many possible ones which our inherited genome could have borne out by interacting with its unarguably changing environments: “*If correct, this hypothesis predicts that individual neural cells will, in fact, have distinctive spatially and temporally defined genomic sequences and chromatin structure.*” [66]. Nothing new under the sun: this prediction simply adds more appeal to the teachings untiringly dispensed over the years by Temin [12], McClintock [15] and Edelman [67], and will perhaps stimulate some re-evaluation of the significance of the notion that “*brain plasticity, while considerable, is constrained by genetics.*” [68].

The increasing overlapping of the broad fields of genomics and epigenomics emphasizes the importance of a large-scale approach to the study of the development-related varia-

tions of chromatin (the cellular structure which 140 years ago was dubbed “nuclein” by Miescher and launched molecular genetics), as well as of their phenotypic and behavioural consequences. The size of the endeavours and the expected value of the results recommend an integrated approach, but the absence of a rigorously defined listing of the variables, combined with the possibility of single groups to follow independent lines of research with accessible instrumentation would rather encourage the promotion of several investigators-driven projects.

If the pursuit of God’s particle, under the species of Higgs’ boson, as approved and enacted by the scientific community, needs a 27-km long, 5 billion €, 10.000 researchers LHC at CERN, Geneva, the Holy Grail of human genetics, *i. e.* the deciphering of our phenotype(s) from the spelling of the A, C, G and T of our genome(s), does not require a similar deployment of investigative weapons. Even assuming it can be found at all, it is possible that similarly relevant advancements would be secured thanks to less structured small scale studies, with minor overall risks if something goes out of control, experimentally as well as conceptually.

REFERENCES

- [1] Schrödinger, E. *What is Life — the Physical Aspect of the Living Cell*. Cambridge University Press; **1944**.
- [2] Abbott, A. Human genome at ten: The human race. *Nature*, **2010**, *464*(7289), 668-669
- [3] Ho, L.; Crabtree, G. R. Chromatin remodelling during development. *Nature*, **2010**, *463*, 474-484.
- [4] Ptashne, M.; Hobert, O.; Davidson, E. Questions over the scientific basis of epigenome project. *Nature*, **2010**, *464*(7288), 487.
- [5] Szyf, M. The early life environment and the epigenome. *BBA*, **2009**, *1790*, 878-885.
- [6] Margueron, R.; Reinberg, D. Chromatin structure and the inheritance of epigenetic information. *Nat. Rev. Genet.*, **2010**, *11*, 285-296.
- [7] Perry, G.H.; Dominy, N.J.; Claw, K.G.; Lee, A.S.; Fiegler, H.; Redon, R.; Werner, J.; Villanea, F.A.; Mountain, J.L.; Misra, R.; Carter, N.P.; Lee, C.; Stone, A.C. Diet and evolution of human amylase gene copy number variations. *Nat. Genet.*, **2007**, *39*, 1256-1260.
- [8] Turner, B. M. Epigenetic responses to environmental change and their evolutionary implications. *Phil. Trans. Royal Soc.*, **2009**, *364*, 3403-3418.
- [9] Dear, D. H. Copy-number variation: the end of the genome? *Trends Biotechnol.*, **2009**, *27*, 448-455.
- [10] Beetschen, J.-C.; Fischer, J.-L. Yves Delage (1854-1920) as a forerunner of modern nuclear transfer experiments. *Int. J. Dev. Biol.*, **2004**, *48*, 607-612.
- [11] Beroukhi, R.; Mermel, C. H.; Porter, D.; Wei, G.; Raychaudhuri, S.; Donovan, J.; Barretina, J.; Boehm, J. S.; Dobson, J.; Urashima, M.; Mc Henry, K. T.; Pinchback, R. M.; Ligon, A. H.; Cho, Y. J.; Haery, L.; Greulich, H.; Reich, M.; Winckler, W.; Lawrence, M. S.; Weir, B. A.; Tanaka, K. E.; Chiang, D. Y.; Bass, A. J.; Loo, A.; Hoffman, C.; Prensner, J.; Liefeld, T.; Gao, Q.; Yecies, D.; Signoretti, S.; Maher, E.; Kaye, F. J.; Sasaki, H.; Tepper, J.E.; Fletcher, J. A.; Taberner, J.; Baselga, J.; Tsao, M. S.; Demichelis, F.; Rubin, M. A.; Janne, P. A.; Daly, M. J.; Nucera, C.; Levine, R. L.; Ebert, B. L.; Gabriel, S.; Rustgi, A. K.; Antonescu, C. R.; Ladanyi, M.; Letai, A.; Garraway, L. A.; Loda, M.; Beer, D. G.; True, L. D.; Okamoto, A.; Pomeroy, S. L.; Singer, S.; Golub, T. R.; Lander, E. S.; Getz, G.; Sellers, W. R.; Meyerson, M. The landscape of somatic copy-number alteration across human cancers. *Nature*, **2010**, *463*, 899-905.
- [12] Temin, H. M. The DNA Provirus Hypothesis. The Establishment and Implications of RNA-Directed Synthesis. **1975**, Nobel Lecture.
- [13] Griffith, D. J. Endogenous retroviruses in the human genome sequence. *Genome Biol.*, **2001**, *2*(6), 1017.1-1017.5.
- [14] Gingeras, Th. Implications of non-cp-linear transcripts. *Nature*, **2010**, *461*, 206-211
- [15] McClintock, B. The significance of the response of the genome to challenge. **1983**, Nobel Lecture.
- [16] Zeh, D.W.; Zeh, J.A.; Ishida, Y. Transposable elements and an epigenetic basis for punctuated equilibria. *Bioessays*, **2009**, *31*, 715-726.
- [17] Ostertag, E. M.; Kazazian, H. H. L1 integration in a transgenic mouse model. *Nature*, **2005**, *435*, 810-811.
- [18] Lee, J.A.; Lupski, J.R. Genomic rearrangements and gene copy number alterations as a cause of nervous system disorders. *Neuron*, **2006**, *52*, 103-121.
- [19] Storchova, Z.; Pellman, D. From polyploidy to aneuploidy, genome instability and cancer. *Nat. Rev. Mol. Cell Biol.*, **2004**, *5*, 45-54.
- [20] Jackson, S.P.; Bartek, J. The DNA-damage response in human biology and disease. *Nature*, **2009**, *461*, 1071-1078.
- [21] Shapiro, J. A. Revisiting the Central Dogma in the 21st century. *Ann. N. Y. Acad. Sci.*, **2009**, *1178*, 6-28.
- [22] Morgante, M. Plant genome organisation and diversity: the year of junk! *Curr. Opin. Biotechnol.*, **2006**, *17*, 168-173.
- [23] Sgaramella, V.; Salamini, F. Gene paucity, genome instability, clonal development: has an individual genome the potential to encode for more than one brain? *DNA Repair*, **2006**, *5*, 531-533.
- [24] The Human Genome Variation Working Group. Completing the map of human genetic variation. *Nature*, **2007**, *447*, 161-165.
- [25] Gaubatz, J.W. Extrachromosomal circular DNAs and genomic sequences plasticity in eukaryotic cells. *Mutat. Res.*, **1990**, *237*, 5-6.
- [26] Panelli, S.; Damiani, G.; Galli, C.; Sgaramella, V. Rearranged genomes of bovine blood cells can allow the development of clones till late stages; but rare unrearranged genomes have greater potential and lead to adulthood. *Gene*, **2004**, *334*, 99-103.
- [27] Tachibana, M.; Sparman, M.; Sritanauudomchai, H.; Ma, H.; Clepper, L.; Woodward, J.; Li, Y.; Ramsey, C.; Kolotushkina, O.; Mitralipov, S. Mitochondrial gene replacement in primate offspring and embryonic stem cells. *Nature*, **2009**, *461*, 367-72.
- [28] Garber, M.; Zody, M. C.; Arachchi, H. M.; Berlin, A.; Gnerre, S.; Green, L. M.; Lennon, N.; Nusbaum, C. Closing gaps in the human genome using sequencing by synthesis. *Genome Biol.*, **2009**, *10*, RR60 (doi:10.1186/gb-2009-10-6-r60).
- [29] Drmanac, R.; Sparks, A. B.; Callow, M. J.; Halpern, A. L.; Burns, N. L.; Kerami, B. G.; Carnevali, P.; Nazarenko, I.; Nilsen, G. B.; Yeung, G.; Dahl, F.; Fernandez, A.; Staker, B.; Pant, K. P.; Bacchash, J.; Borcherding, A. P.; Brownley, A.; Cedeno, R.; Chen, L.; Chernikoff, D.; Cheung, A.; Chirita, R.; Curson, B.; Ebert, J. C.; Hacker, C. R.; Hartlage, R.; Hauser, B.; Huang, S.; Jiang, Y.; Karpinchyk, V.; Koenig, M.; Kong, C.; Landers, T.; Le, C.; Liu, J.; McBride, C. E.; Morenzoni, M.; Morey, R. E.; Mutch, K.; Perazich, H.; Perry, K.; Peters, B. A.; Peterson, J.; Pethiyagoda, C. L.; Pothuraju, K.; Richter, C.; Rosenbaum, A. M.; Roy, S.; Shafto, J.; Sharanhovich, U.; Shannon, K. W.; Sheppy, C. G.; Sun, M.; Thakuria, J. V.; Tran, A.; Vu, D.; Zaranek, A. W.; Wu, X.; Drmanac, S.; Oliphant, A. R.; Banyai, W. C.; Martin, B.; Ballinger, D. G.; Church, G. M.; Reid, C. A. Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science*, **2010**, *327*, 78-81.
- [30] Rohlin, A.; Wernersson, J.; Engwall, Y.; Wiklund, L.; Björk, J.; Nordling, M. Parallel sequencing used in detection of mosaic mutations: comparisons with four diagnostic DNA screening techniques. *Hum. Mutat.*, **2009**, *30*, 1012-1020.
- [31] Smith, J.J.; Antonacci, F.; Eichler, E.E.; Amemiya, C.T. Programmed loss of millions of base pairs from a vertebrate genome. *PNAS*, **2009**, *106*, 11212-11216.
- [32] Hilbricht, T.; Varotto, S.; Sgaramella, V.; Bartels, D.; Salamini, F.; Furini, A. Retrotransposons and siRNA have a role in the evolution of desiccation tolerance leading to resurrection of the plant *Craterostigma plantagineum*. *New Phytol.*, **2008**, *179*, 877-887.
- [33] Martienssen, R. Great leap forward? Transposable elements, small interfering RNA and adaptive Lamarckian evolution. *New Phytol.*, **2008**, *179*, 570-572.

- [34] Friedberg, E.C.; Walker, G.C.; Siede, W. *DNA Repair and Mutagenesis*. ASM, 1995.
- [35] Maksakova, I.A.; Mager, D.L.; Reiss, D. Keeping active endogenous retroviral-like elements in check: the epigenetic perspective. *Cell. Mol. Life Sci.*, **2008**, *65*, 3329-3347.
- [36] Tam, O.H.; Aravin, A.A.; Stein, P.; Girard, A.; Murchison, E.P.; Cheloufi, S.; Hodges, E.; Anger, M.; Sachidanandam, R.; Schultz, R.M.; Hannon, G.J. Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature*, **2008**, *453*, 534-539.
- [37] Bock, R. The give-and-take of DNA: horizontal gene transfer in plants. *Trends Plant Sci.*, **2009**, *15*, 11-15.
- [38] Vanneste, E.; Voet, T.; Le Caignec, C.; Ampe, M.; Konings, P.; Melotte, C.; Debrock, S.; Amyere, M.; Vikkula, M.; Schuit, F.; Fryns, J.P.; Verbeke, G.; D'Hooghe, T.; Moreau, Y.; Vermeesch, J. R. Chromosome instability is common in human cleavage-stage embryos. *Nat. Med.*, **2008**, *15*, 577-583.
- [39] Ledbetter, D.H. Chaos in the embryo. *Nat. Med.*, **2008**, *15*, 490-491.
- [40] Vorsanova, S.G.; Kolotii, A.D.; Iourov, I.Y.; Monakhov, V.V.; Kirillova, E.A.; Soloviev, I.V.; Yurov, Y.B. Evidence for high frequency of chromosomal mosaicism in spontaneous abortions revealed by interphase FISH analysis. *J. Histochem. Cytochem.*, **2005**, *53*, 375-380.
- [41] Iourov, I. Y.; Vorsanova, S. G.; Liehr, T.; Kolotii, A. D.; Yurov, Y. B. Increased chromosome instability dramatically disrupts neural genome integrity and mediates cerebellar degeneration in the ataxia-telangiectasia brain. *Hum. Mol. Genet.*, **2009**, *18*, 2656-2669.
- [42] Westra, J.W.; Peterson, S.E.; Yung, Y.C.; Mutoh, T.; Barral, S.; Chun, J. Aneuploid mosaicism in the developing and adult cerebellar cortex. *J. Comparat. Neurol.*, **2008**, *507*, 1944-1951.
- [43] Sgaramella, V.; Astolfi, P. Somatic genome variations interact with environment, genome and epigenome in the determination of the phenotype: A paradigm shift in genomics? *DNA Repair*, **2010**, *9*, 470-473.
- [44] Conrad, D. F.; Pinto, D.; Redon, R.; Feuk, L.; Gokcumen, O.; Zhang, Y.; Aerts, J.; Andrews, T. D.; Barnes, C.; Campbell, P.; Fitzgerald, T.; Hu, M.; Ihm, C. H.; Kristiansson, K.; Macarthur, D. G.; Macdonald, J. R.; Onyiah, I.; Pang, A. W.; Robson, S.; Stirrups, K.; Valsesia, A.; Walter, K.; Wei, J.; Wellcome Trust Case Control Consortium, Tyler-Smith, C.; Carter, N. P.; Lee, C.; Scherer, S. W.; Hurles, M. E. Origins and functional impact of copy-number variations in the human genome. *Nature*, **2009**, *464*, 704-712.
- [45] Vijg, J. *Ageing of the genome*. Oxford: Oxford University Press; **2007**.
- [46] Iourov, I.Y.; Vorsanova, S.G.; Liehr, T.; Yurov, Y.B. Aneuploidy in the normal, Alzheimer's disease and ataxia-telangiectasia brain: differential expression and pathological meaning. *Neurobiol. Dis.*, **2009**, *34*, 212-220.
- [47] Vega, F.; Medeiros, L.J. V(D)J recombination and staggered DNA breaks: guilty again. *Blood*, **2010**, *115*, 2121-2122.
- [48] Bacolla, A.; Wells, R. D. Non-B conformations as determinants of mutagenesis and human disease. *Mol. Carcinog.*, **2009**, *48*, 273-285.
- [49] Motl, J. A.; Chalker, D. L. Subtraction by addition: domesticated transposases in programmed DNA elimination. *Genes Dev.*, **2009**, *23*, 2455-2460.
- [50] Sen, S. K.; Han, K.; Wang, J.; Lee, J.; Wang, H.; Callinan, P.A.; Dyer, M.; Cordaux, R.; Liang, P.; Batzer, M.A. Human Genomic Deletions Mediated by Recombination between *Alu* Elements. *Am. J. Hum. Genet.*, **2006**, *79*, 41-53.
- [51] Eichler, E.E. Widening the spectrum of human genetic variation. *Nat. Genet.*, **2006**, *38*, 9-11.
- [52] Wakayama, T.; Perry, A.C.; Zuccotti, M.; Johnson, K.R.; Yanagimachi, R. Full-term development of mice from enucleated oocytes injected with cumulus cell nuclei. *Nature*, **1998**, *394*, 369-373.
- [53] Gurdon, J. Nuclear reprogramming in egg. *Nat. Med.*, **2009**, *15*, 1141-1144.
- [54] Whitelaw, E.; Martin, D.K. Retroposons as epigenetic mediators of phenotypic variations in mammals. *Nat. Genet.*, **2001**, *27*, 361-365.
- [55] Ossowski, S.; Schneeberger, K.; Lucas-Lledó, J.I.; Warthmann, N.; Clark, R.M.; Shaw, R.G.; Weigel, D.; Lynch, M. The rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana*. *Science*, **2010**, *327*, 92-96.
- [56] Waterston, R. H.; Lindblat-Toh, K.; Birney, E.; Rogers, J.; Abril, J.F.; Agarwal, P.; Agarwala, R.; Ainscough, R.; Alexandersson, M.; An, P.; Antonarakis, S.E.; Attwood, J.; Baertsch, R.; Bailey, J.; Barlow, K.; Beck, S.; Berry, E.; Birren, B.; Bloom, T.; Bork, P.; Botcherby, M.; Bray, N.; Brent, M.R.; Brown, D.G.; Brown, S.D.; Bult, C.; Burton, J.; Butler, J.; Campbell, R.D.; Carninci, P.; Cawley, S.; Chiaromonte, F.; Chinwalla, A.T.; Church, D.M.; Clamp, M.; Clee, C.; Collins, F.S.; Cook, L.L.; Copley, R.R.; Coulson, A.; Couronne, O.; Cuff, J.; Curwen, V.; Cutts, T.; Daly, M.; David, R.; Davies, J.; Delehaunty, K.D.; Deri, J.; Dermizakis, E.T.; Dewey, C.; Dickens, N.J.; Diekhans, M.; Dodge, S.; Dubchak, I.; Dunn, D.M.; Eddy, S.R.; Elnitski, L.; Emes, R.D.; Esvara, P.; Eyras, E.; Felsenfeld, A.; Fewell, G.A.; Flicek, P.; Foley, K.; Frankel, W.N.; Fulton, L.A.; Fulton, R.S.; Furey, T.S.; Gage, D.; Gibbs, R.A.; Glusman, G.; Gnerre, S.; Goldman, N.; Goodstadt, L.; Grafham, D.; Graves, T.A.; Green, E.D.; Gregory, S.; Guigó, R.; Guyer, M.; Hardison, R.C.; Haussler, D.; Hayashizaki, Y.; Hillier, L.W.; Hinrichs, A.; Hlavina, W.; Holzer, T.; Hsu, F.; Hua, A.; Hubbard, T.; Hunt, A.; Jackson, I.; Jaffe, D.B.; Johnson, L.S.; Jones, M.; Jones, T.A.; Joy, A.; Kamal, M.; Karlsson, E.K.; Karolchik, D.; Kasprzyk, A.; Kawai, J.; Keibler, E.; Kells, C.; Kent, W.J.; Kirby, A.; Kolbe, D.L.; Korf, I.; Kucherlapati, R.S.; Kulbokas, E.J.; Kulp, D.; Landers, T.; Leger, J.P.; Leonard, S.; Letunic, I.; Levine, R.; Li, J.; Li, M.; Lloyd, C.; Lucas, S.; Ma, B.; Maglott, D.R.; Mardis, E.R.; Matthews, L.; Mauceli, E.; Mayer, J.H.; McCarthy, M.; McCombie, W.R.; McLaren, S.; McLay, K.; McPherson, J.D.; Meldrum, J.; Meredith, B.; Mesirov, J.P.; Miller, W.; Miner, T.L.; Mongin, E.; Montgomery, K.T.; Morgan, M.; Mott, R.; Mullikin, J.C.; Muzny, D.M.; Nash, W.E.; Nelson, J.O.; Nhan, M.N.; Nicol, R.; Ning, Z.; Nusbaum, C.; O'Connor, M.J.; Okazaki, Y.; Oliver, K.; Overton-Larty, E.; Pachter, L.; Parra, G.; Pepin, K.H.; Peterson, J.; Pevzner, P.; Plumb, R.; Pohl, C.S.; Poliakov, A.; Ponce, T.C.; Ponting, C.P.; Potter, S.; Quail, M.; Reymond, A.; Roe, B.A.; Roskin, K.M.; Rubin, E.M.; Rust, A.G.; Santos, R.; Sapojnikov, V.; Schultz, B.; Schultz, J.; Schwartz, M.S.; Schwartz, S.; Scott, C.; Seaman, S.; Searle, S.; Sharpe, T.; Sheridan, A.; Shownkeen, R.; Sims, S.; Singer, J.B.; Slater, G.; Smit, A.; Smith, D.R.; Spencer, B.; Stabenau, A.; Stange-Thomann, N.; Sugnet, C.; Suyama, M.; Tesler, G.; Thompson, J.; Torrents, D.; Trevaskis, E.; Tromp, J.; Ucla, C.; Ureta-Vidal, A.; Vinson, J.P.; Von Niederhausern, A.C.; Wade, C.M.; Wall, M.; Weber, R.J.; Weiss, R.B.; Wendl, M.C.; West, A.P.; Wetterstrand, K.; Wheeler, R.; Whelan, S.; Wierzbowski, J.; Willey, D.; Williams, S.; Wilson, R.K.; Winter, E.; Worley, K.C.; Wyman, D.; Yang, S.; Yang, S.P.; Zdobnov, E.M.; Zody, M.C.; Lander, E.S. Initial sequencing and comparative analysis of the mouse genome. *Nature*, **2002**, *420*, 520-565.
- [57] Panelli, S.; Damiani, G.; Espen, L.; Sgaramella, V. Ligation overcomes terminal underrepresentation in multiple displacement amplification of linear DNA. *BioTechniques*, **2005**, *39*, 174-179.
- [58] Sen, S. K.; Han, K.; Wang, J.; Lee, J.; Wang, H.; Callinan, P. A.; Dyer, M.; Cordaux, R.; Liang, P.; Batzer, M. A. Human genomic deletions mediated by recombination between *Alu* elements. *Am. J. Hum. Genet.*, **2006**, *79*, 41-53.
- [59] Bovee, D.; Zhou, Y.; Haugen, E.; Wu, Z.; Hayden, H. S.; Gillett, W.; Cooper, G. M.; Sampas, N.; Phelps, K.; Levy, R.; Morrison, V. A.; Sprague, J.; Jewett, D.; Buckley, D.; Subramaniam, S.; Chang, J.; Smith, D. R.; Olson, M. Z. V.; Eichler, E. E.; Kayl, R. Closing gaps in the human genome with fosmid resources generated from multiple individuals. *Nat. Genet.*, **2008**, *40*, 96-101.
- [60] Freeman, J. L.; Perry, G. H.; Feuk, L.; Redon, R.; McCarroll, S. A.; Altshuler, D. M.; Aburatani, H.; Jones, K. W.; Tyler-Smith, C.; Hurles, M. E.; Carter, N.P.; Scherer, S. W.; Lee, C. Copy number variation: New insights in genome diversity. *Genome Res.*, **2006**, *16*, 949-961.
- [61] Innan, H.; Kondrashov, F. The evolution of gene duplication: classifying and distinguishing between models. *Nat. Rev. Genet.*, **2010**, *11*, 97-107.
- [62] Zufall, R. A.; Robinson, T.; Katz, L. A. Evolution of developmentally regulated genome rearrangements in eukaryotes. *J. Exp. Zool. B Mol. Dev. Evol.*, **2005**, *304B*, 448-455.

- [63] Lamarck, J-B. *Philosophie zoologique, ou exposition des considérations relatives à l'histoire naturelle des animaux*. 1809. Paris, Dent.
- [64] Robertson, M. Editorial: Gene regulation, evolvability, and the limits of genomics. *J. Biol.*, **2009**, 8, 94.
- [65] Koonin, E.; Wolf, Y.I. Is evolution Darwinian or Lamarckian? *Biol. Direct*, **2009**, 4, 42-56.
- [66] Mattick, J. S.; Mehler, M. F. RNA editing, DNA recoding and the evolution of human cognition. *Trends Neurosci.*, **2008**, 31, 227-230.
- [67] Edelman, G.M.; Tononi, G. *A Universe of Consciousness. How Matter Becomes Imagination*. Basic Books, **2000**.
- [68] Gross, C.G. Making sense of printed symbols. Review of the book "Reading in the brain", by S. Dehaene, Viking NY 2009. *Science*, **2010**, 327, 524-525.