REVIEW ESSAYS

Prospects & Overviews

BioEssays

www.bioessays-journal.com

Epigenetic and Transcriptional Variability Shape Phenotypic Plasticity

Simone Ecker, Vera Pancaldi, Alfonso Valencia, Stephan Beck, and Dirk S. Paul*

Epigenetic and transcriptional variability contribute to the vast diversity of cellular and organismal phenotypes and are key in human health and disease. In this review, we describe different types, sources, and determinants of epigenetic and transcriptional variability, enabling cells and organisms to adapt and evolve to a changing environment. We highlight the latest research and hypotheses on how chromatin structure and the epigenome influence gene expression variability. Further, we provide an overview of challenges in the analysis of biological variability. An improved understanding of the molecular mechanisms underlying epigenetic and transcriptional variability, at both the intra- and inter-individual level, provides great opportunity for disease prevention, better therapeutic approaches, and personalized medicine.

1. Introduction

No two cells in a cellular population are the same, and no two individuals of a multi-cellular species are identical—not even if they share the same genetic makeup like monozygotic twins or cloned animals. Even cells or model organisms with the same genotype that are grown under the exact same laboratory conditions can display variability in appearance and behavior.[1,2] Furthermore, all cells of a multi-cellular organism have essentially the same genome, but exhibit many different phenotypes. This is due to epigenetic and transcriptional differences that lead to the production of different proteins, which drive phenotypic diversity. Whether a gene is expressed in a specific cell at a given moment in time depends on a multitude of regulatory proteins and biochemical steps. Randomness and biological "noise" are present in every biochemical process, especially in the epigenetic modification of DNA and the transcription and translation of genes. Thus, epigenetic and gene expression variability are key contributors to phenotypic differences.

Here, we distinguish between different types of variability at different organizational levels, specifically: 1) cell-to-cell variability in a population of cells; 2) inter-individual variability of multi-cellular organisms; and 3) variability across populations and species. Cell-to-cell variability, for example, is important in shaping cell fate determination and plays a key functional role in cellular differentiation.[3,4] Also, it is thought to be required for population robustness and higher-level function of multi-cellular organisms.[5] For example, variability in a population of cells allows essentially binary decisions, such as undergoing cell death, to turn into more flexible and fine-tuned responses at the level of the cell population as a whole. This creates an adaptive advantage and provides benefits in survival.[6,7] These effects have mainly been investigated in unicellular organisms, but are known to also be relevant for human adaptation.[8] Presumably, they are a unifying feature of biological systems at all levels, with variability forming the basis for positive natural selection, thereby enabling evolution.[9] In fact, the above-described mechanisms may be comparable to how evolution makes ecosystems robust through the generation of biodiversity.[10]

The different levels of variability are related to each other.[11–13] There also exists a direct correspondence between the measurement of variability at one time point in a population of, for example, 1000 cells and the measurement of variability of one cell at 1000 time points[14]—a concept known as ergodic hypothesis. It has further been shown that cell-to-cell gene expression variability in yeast populations correlates with variability across populations, and—to a lesser extent—across species.[12,13,15] For example, Dong et al.[15] showed that fluctuations in gene expression between isogenic yeast cells correlate well with expression variation within individual cells, and this variability also correlated with variability between

Dr. S. Ecker, Prof. S. Beck
UCL Cancer Institute
University College London
72 Huntley Street, London WC1E 6BT, UK
E-mail: s.ecker@ucl.ac.uk

Dr. V. Pancaldi, Prof. A. Valencia
Barcelona Supercomputing Center (BSC)
C/ Jordi Girona 39-31, 08034 Barcelona, Spain

Dr. V. Pancaldi, Prof. A. Valencia
ICREA
Pg. Lluís Companys 23, 08010 Barcelona, Spain

Dr. D. S. Paul
MRC/BHF Cardiovascular Epidemiology Unit
Department of Public Health and Primary Care
University of Cambridge
Cambridge CB1 8RN, UK

Dr. D. S. Paul
Department of Human Genetics
Wellcome Trust Sanger Institute
Wellcome Genome Campus
Hinxton, Cambridge CB10 1HH, UK

DOI: 10.1002/bies.201700148

different yeast strains or species. The positive correlation between different levels of variability was maintained under varying environmental conditions. Similar results were obtained by a study investigating the relationship between transcriptional variation across mammalian individuals and species in limb development,[9] showing that variability in gene expression levels across four different species of mammals was correlated with intra-species expression variability among individual animals.

Recently, large collections of human epigenomic and transcriptomic data have become available, facilitated by consortia such as the NIH Roadmap Epigenomics project,[16] the International Human Epigenome Consortium (IHEC, http://ihec-epigenomes.org) and the associated BLUEPRINT project[17] (http://www.blueprint-epigenome.eu). In the context of BLUEPRINT, we analyzed differential variability across primary immune cells derived from healthy individuals and aimed to characterize the extent and functional implication of epigenetic and transcriptional variability in different immune cell types.[18]

In a previous study, we had investigated gene expression differences between the two main subtypes of chronic lymphocytic leukemia, known to show only minimal differential expression,[19,20] and observed strongly increased gene expression variability in the more aggressive subtype of the disease.[21] While epigenetic analyses of monozygotic twins discordant for type 1 diabetes revealed no differences in mean DNA methylation, we found substantial enrichment of hypervariable loci among the siblings with the disease.[22] Together, these studies were among the first to classify disease status or aggressiveness based on variability, where the classical comparison of mean DNA methylation or gene expression levels was not informative. The data highlight the importance of inter-individual epigenetic and transcriptional variability and its application to uncovering disease biology.

In this review, we focus on epigenetic variability (i.e., DNA methylation and chromatin structure) and transcriptional variability (i.e., gene-level expression variability). We do not discuss allele-specific expression or transcript and isoform variability. We distinguish between two main types of biological variability: 1) inter-individual variability, that is the differences between individuals; and 2) intra-individual variability or cell-to-cell variability, that is, differences across single cells of a population. We define sources and determinants of epigenetic and transcriptional variability and provide examples of their functions and implications in health and disease. Last, we discuss questions and challenges in the analysis of variability, and consider how these concepts and approaches could be applied to the development of new therapeutic approaches and personalized medicine.

2. Biological Variability Derives From Distinct Sources

There are many possible sources of epigenetic and transcriptional variability, which can be divided into three main categories: 1) individual-intrinsic factors; 2) environmental factors; and 3) random fluctuations, also referred to as stochasticity. These different sources of variability are further described below and summarized in **Figure 1**.

2.1. Individual-Intrinsic Factors

The expression variability of a gene is, in part, encoded by its genomic context (e.g., promoter DNA sequence)[23] and further controlled by the epigenome[11] (see Section 3). Additional, non-genetic, individual-intrinsic factors such as sex, age, and environmental factors, further influence both epigenetic and gene expression variability.

2.1.1. Genetic Variation

Most studies on gene expression variability thus far have focused on the mapping of genetic variants associated with gene expression changes across individuals,[13] so-called expression quantitative trait loci (eQTL). These studies quantify the effect that single-nucleotide polymorphisms (SNPs) and copy number variations (CNVs) have on gene expression. Atlases of cis- and trans-eQTLs across a vast number of cell types, tissues, and environmental conditions have been generated. Remarkably, up to 90% of expressed protein-coding genes have an eQTL in at least one tissue.[24] The amount of gene expression variation explained by genetic variability is typically small (<5% for the majority of genes[4,25,26]). Thus, common genetic variants only explain a small proportion of total expression variation. However, their effects are still stronger than those of major demographic factors such as age and sex.[25] Genetic variation also has an impact on the epigenome, and there is a high level of interaction between genetic, epigenetic, and transcriptional variability.[26–28] Of note, genetic variants can affect both the mean and variance of a quantitative phenotype.[6,29,30]

2.1.2. Sex

DNA methylation and gene expression differences between males and females have been reported for both autosomal genes and genes expressed on sex chromosomes.[31,32] These differences are particularly important in the context of the immune system, as women exhibit generally stronger immune responses than men, and many auto-immune diseases such as rheumatoid arthritis and multiple sclerosis have a higher incidence in females.[31,33]

2.1.3. Age

Besides sex, age is the most important non-genetic source of inter-individual variability (recently reviewed by Tejedor and Fraga[34]). It has been shown that epigenetic marks change over life to such an extent that an individual's age can be predicted from its DNA methylation profile.[35,36] Further, the methylome and transcriptome become more and more diverse with increasing age. This phenomenon of increased variability with age occurs both in genetically identical twins[37,38] and unrelated individuals[8,35,39] and is also referred to as "epigenetic drift."

A recent study by Jenkinson et al.[40] showed that DNA methylation entropy (a measure of disorder or randomness) increased in older individuals, with these changes being more

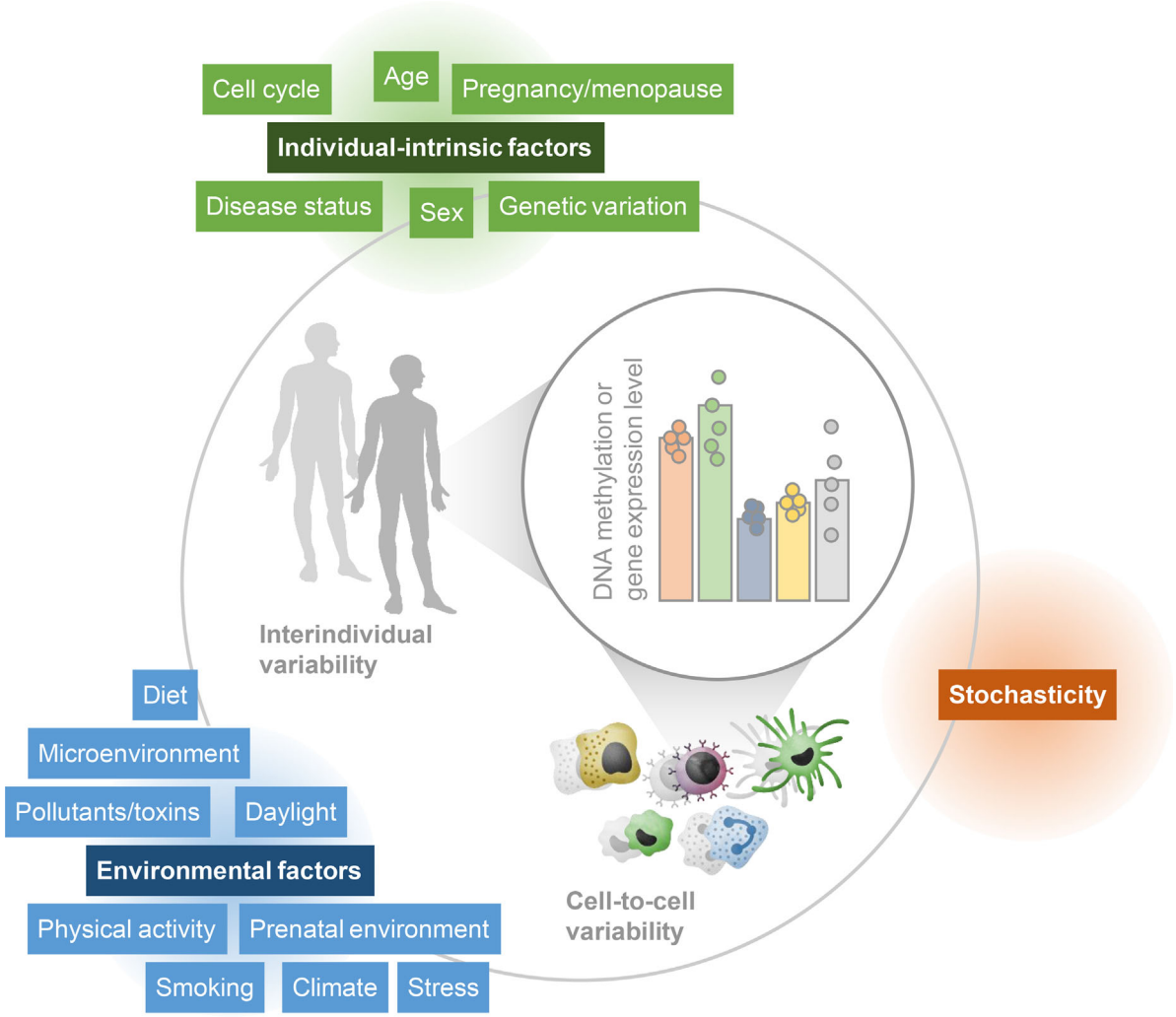

Figure 1. Overview of the different types and sources of epigenetic and transcriptional variability. Epigenetic and transcriptional variability exist between different individuals (i.e., inter-individual variability) and between different cells (i.e., cell-to-cell variability). The bar plot illustrates the levels of an epigenetic mark (e.g., DNA methylation) or gene expression measured in a specific cell type or individual. Sources of such variability, including individual-intrinsic and environmental factors, are also shown. Credits: The immune response, BigPicture (http://bigpictureeducation.com).

pronounced than age-associated differences in mean DNA methylation. An extraordinarily long-lived human population was shown to exhibit less pronounced epigenetic drift,[41] pointing to an important implication of biological variability in aging and its association with life- and healthspan.[42] The epigenetic component of accumulating environmental exposure, and its interplay with genetic and stochastic factors, provides an explanation for the frequently observed discordance of disease between monozygotic twins and the increase of common diseases with age.[43]

2.1.4. Other Non-Genetic Individual-Intrinsic Factors

Related to aging, variability also occurs over an individual's lifetime due to development, growth, pregnancy, and menopause. At the cellular level, the distinct stage of the cell cycle is another important source of variability.[44,45,46]

2.2. Environmental Factors

The epigenome forms the intersection between the genome and the environment. Many lifestyle and behavioral factors have been shown to influence the epigenome—and as a consequence—the transcriptome and the resulting phenotype. Such alterations can have a long-lasting impact, and may potentially even be passed on to subsequent generations.[47,48] Environmental factors impacting the epigenome and transcriptome include prenatal exposures, childhood adversities, nutrition, physical activity, exercise, stress, exposure to pollutants and toxins, smoking, climate, season, daylight, culture, education, socio-economic factors, and many others.[31,39,49–52]

At the cellular level, the micro-environment plays an important role; for example, hematopoietic stem cells from different micro-environments (i.e., the bone marrow, cord blood, and fetal liver), exhibit different DNA methylation profiles.[53] The location of a cell within a population and its local

crowdedness are also a source of variability, as well as cellular volume and mitochondrial content.[44,54]

2.3. Stochasticity

Stochasticity is particularly important at the level of cell-to-cell variability in biochemical processes. This stochasticity is thought to contribute to cell differentiation, adaptation to changing environmental conditions (e.g., rapid response to external stimuli), and population robustness.[3,7,12]

At the level of individual cells and the level of multi-cellular organisms, each of these sources of variability also leads to temporal variation. The accumulation of environmental exposures contributes to the strong epigenetic changes observed with increasing age.[44] Most of the variability is determined by a complex combination and interaction of different sources of variation. The relative contributions of different sources of variability can be different in distinct cell types, and are only starting to be revealed. For example, DNA methylation variability of T cells was found to be associated with donor age at many loci, while in monocytes, DNA methylation variability associated more often with season.[18] Interestingly, we found more than a thousand genes with small but significant sex-specific differential expression in neutrophils, while monocytes and T cells showed less sex-specific gene expression.[18] Thus, variability can be different for the same gene across cell types and distinct tissues, and can be influenced by distinct factors.

3. Chromatin Properties and Gene Regulatory Networks Control Gene Expression Variability

Expression variability for a single gene can be controlled by different mechanisms, including: 1) local genome properties such as the DNA sequence and epigenetic landscape; 2) 3D chromatin structure; and 3) the gene's position in regulatory networks (see **Figure 2**).

3.1. Epigenetic Mechanisms

Chromatin properties at gene promoters and bodies can impact transcriptional variation.[11,46,56] TATA boxes are sequence motifs that can be found at the promoters of genes displaying variability and plasticity in changing environments. The mechanism through which the presence of TATA boxes affects expression is partly dependent on the different nucleosome architecture favored by this motif.[57] A recent single-cell RNA sequencing study of mouse embryonic stem cells reported that active chromatin states in gene bodies led to reduced gene expression variability and that gene body chromatin marks are more important determinants of gene expression variability than promoter sequence features.[46] Alemu et al.[11] showed that reduced variability associated with active chromatin marks (such as H3K4me3 and H3K36me3, and DNaseI hypersensitivity) and increased variability associated with repressive marks such as H3K27me3 (a mark that is catalyzed by the Polycomb complex often found to mediate the silencing of developmentally important genes). Using single-cell transcriptomics, it was shown that a subset of genes that are active despite being marked by H3K27me3 and bound by Polycomb displayed higher variability in mouse embryonic stems cells.[58] This is thought to be due to different bursting dynamics characterizing transcription at these genes, which show high expression levels despite carrying repressive chromatin marks. Polycomb could modulate the frequency of transcriptional bursting, affecting expression variability independently of its mean levels.[58] This enhanced switching of the transcriptional state from "OFF" to "ON" in a subset of Polycomb targets could produce expression fluctuations in single cells, which have been observed in embryonic stem cells.[59] Polycomb was also shown to modulate methylation variability,[40] which, in turn, can lead to gene expression variability.[18,26,60]

Indeed, when we correlated the DNA methylation values of loci exhibiting high variability in DNA methylation with the corresponding gene expression values derived from the same samples, we observed a significant correlation between DNA methylation and gene expression levels in up to 33% of the hypervariable loci.[18] Interestingly, we also observed high gene expression variability for genes with very consistent (i.e., stable, not variable) DNA methylation at their promoters. The promoters with either highly variable or very consistent DNA methylation levels were also found to exhibit more transcription factor binding motifs than other promoters.[18] These findings indicate a complex regulatory control exerted on the genes found to be at the extreme ends of promoter DNA methylation variability (either very low or very high). Regarding DNA methylation variability at gene bodies, our data suggested that gene expression variability increases with DNA methylation variability.[18]

High mean methylation levels of gene bodies have been associated with reduced transcriptional noise,[61] possibly because methylation excludes the deposition of the H2A.Z histone variant, which was found to reduce expression and increase the variability and plasticity of expression regulation in *Arabidopsis*.[62] Deposition of H2A.Z in promoters has an impact on the nucleosome configuration[63] and DNA methylation control of H2A.Z deposition might further affect transcriptional variability through promoter architecture.

3.2. 3D Chromatin Structure

Recent technological advances have allowed us to map the 3D conformation of chromatin and to uncover the spatial organization of the genome inside the nucleus. These experiments have revealed the complexities of genome folding at different scales and levels. Chromatin loops pervade the genome connecting genomic regions at various linear distance ranges. Loops are thought to be organized in regions of dense chromatin interactions insulated by specific proteins from other similar regions, so-called topologically associated domains (TADs).[64] Chromatin loops are assembled into separate compartments characterized by distinct levels of activity and transcriptional output.[65] Although similar TADs were found across developmental stages, single cells, individuals, and even species, their assignment to different compartments is much more variable[66,67] and the compartments themselves are probably more related to transcriptional programs and phenotypes. Therefore, it is conceivable that variability in expression across single cells

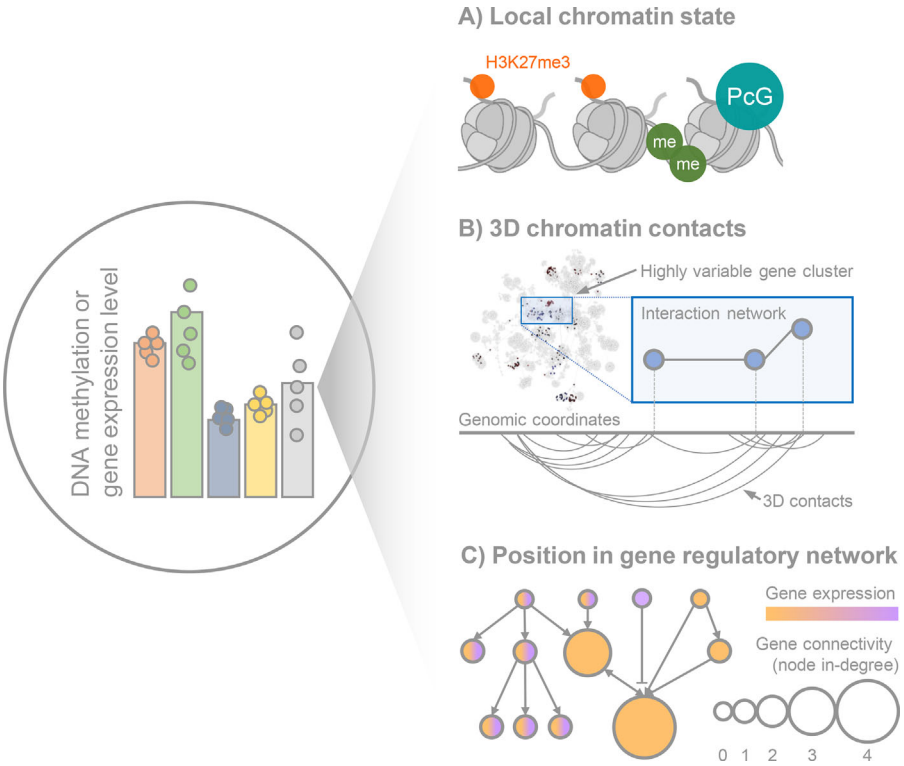

Figure 2. Epigenetic patterns, chromatin structure, and gene regulatory networks control gene expression variability. A) While active chromatin marks such as H3K4me3 and H3K36me3 and DNaseI hypersensitivity have been shown to associate with reduced gene expression variability, repressive marks such as H3K27me3, deposited by members of Polycomb, are associated with increased gene expression variability, possibly through an effect on transcriptional bursting. B) 3D chromatin contacts can be mapped using chromosome capture methods. These contacts form a network where nodes are chromatin fragments and edges are drawn when the fragments interact in 3D. Nodes that are colored have high gene expression variability. Using Chromatin Assortativity (ChAs),[55] we detected a tendency for highly variable genes to preferentially connect in 3D. C) In regulatory networks, central and highly connected genes have been found to show low gene expression variability.

arises as a consequence of the specific assignment of genomic regions to compartments with different levels of activity in different cells.

Motivated by these findings, we investigated gene expression variability of monocytes, neutrophils and T cells in the context of 3D chromatin interactions. To this aim, we leveraged promoter-centered chromatin contact maps generated using the Promoter-Capture HiC protocol.[68] These data provide information on 3D interactions between distant genomic regions. We used a novel statistical method (ChAs) that assesses the relationship between chromatin properties and 3D chromatin interaction networks through a measure called assortativity,[55] which we applied to expression variability. We found that chromatin fragments that overlap genes with high expression variability are preferentially connected with each other in the 3D contact map of the corresponding cell type (manuscript in preparation).

Our results support the hypothesis that chromatin structure can be related to transcriptional variability. Indeed, Kar et al.[58] reported that gene expression variability could be associated with specific regions along chromosomes. The availability of more detailed maps of 3D chromatin structure, combined with single cell transcriptomics, will allow us to further characterize the relationships between nuclear organization and variability at different levels.

3.3. Gene Regulatory Networks

An additional factor affecting expression variability is the gene's context within a regulatory network. For example, genes with high expression variability were found to be less central in signaling networks and less connected, especially in stem cells, whereas the most connected genes were shown to be the stable elements of pluripotent regulatory networks.[66,67] Specific network motifs, such as feed-forward loops, also affect variability of the regulated genes.[66] Moreover, the variability of genes could be directly affected by variability in their regulators. In fact, a recent model for the evolution of gene regulation in bacteria suggests that the propagation of variability of expression from a transcription factor to its target could be the most primordial type of gene regulation.[69]

On a more global level, integrating gene expression datasets with protein-protein interaction networks has made it possible to identify characteristic expression configurations related to disease, cancer stemness and intra-tumor heterogeneity.[70,71] Defining new concepts of network entropy or energy[40] allows the characterization of the association between the expression levels of single genes and the position of these genes in the interaction network. These analyses provide a connection between molecular data and phenotypes such as transcriptome

variability and plasticity, which are seen as global emergent properties of the gene network.

Taken together, the genomic and epigenetic context of a gene, as well as its position in interaction and regulatory networks, are important mechanisms impacting expression variability. Understanding how these mechanisms interact will shed light on how this biological variability is regulated and harnessed in evolution to ensure cell survival and adaptability in changing environments.

4. Variability Plays a Key Role in Human Health and Disease

Epigenetic and transcriptional variability play key roles in human physiology, particularly in the functioning of the human immune system.[31,33,72,73] Conversely, increased variability is also associated with pathogenesis, disease progression, and aggressiveness, and resistance to therapy.[21,22,74–77] Both aspects can be linked back to the same underlying biological principle of variability, providing an evolutionary advantage and thus leading to enhanced survival: cancer cell populations, for example, are known to be highly heterogeneous, and this heterogeneity is strongly associated with disease progression and severity, and therapeutic resistance of the cancer.[74,78] Thus, in this context, variability is detrimental to the host organism but beneficial to the population of malignant cells.

4.1. Variability in the Immune System

Variability is a crucial aspect of the human immune system,[33] at both the inter- and intra-individual level. The human immune system shows a high level of variability and constitutes an incredibly adaptive defense mechanism. This variability is particularly important for migratory cells of the immune system facing highly changeable environments such as infections and invasion of pathogens. Indeed, migratory immune cells, such as monocytes, neutrophils, and lymphocytes, exhibit strong phenotypic plasticity and adaptability.[72]

We found particularly high DNA methylation and gene expression variability in neutrophils compared to monocytes and T cells.[18] This inter-individual variability could possibly be related to intra-individual variability. The corresponding plasticity of these immune cells would enable rapid adaptation to changing external conditions and, importantly, an effective defense against invading pathogens. In our study, neutrophil-specific hyper-variable genes were enriched for functions critical to adaptability and rapid reactions to external stimuli, such as signaling and motility, supporting this hypothesis.

In summary, a high level of immune cell adaptability to environmental changes and functional diversity of immune cells are important for a healthy immune system.[72,73]

4.2. Variability in Cancer

Alongside variability at the genetic level, there is a substantial contribution of non-genetic variation to cancer development, from precursor cells through all stages of tumor progression to metastasis and disease relapse.[74,79,80] Increased DNA methylation variability in normal tissue can be predictive of neoplastic transformation years before the cancer develops.[75] Furthermore, tissue DNA methylation has been shown to explain 77% of the variance of the lifetime risk of cancer.[76]

A general loss of epigenetic stability leads to increased epigenetic and transcriptional variability in cancer.[77] Epigenetic alterations often target enhancer chromatin states that do not only affect the mean level of transcription, but can also increase its variability, possibly leading to intra-tumor heterogeneity.[80] Neoplastic transformation is thought to revert the epigenetic state to more closely resemble stem cells, with high self-renewal potential and/or cells similar to unicellular organisms that achieve adaptability through increased variability.[74,78] In support of this hypothesis, it has been reported that loci that are important in normal cellular differentiation show hyper-variability in cancer.[3] Cancer cells adapt rapidly to changing environments such as alterations in oxygen levels due to neovascularization or necrosis, or metastasis to new micro-environments.[77] Feinberg et al.[74,81] noted that the constitutive activation of enhanced epigenetic plasticity in cancer leads to classical cancer hallmarks such as inflammation, invasion, and proliferation, and recent literature suggests that epigenetic plasticity gives rise to all known hallmarks of cancer.[79]

Epigenetic alterations are important in tumor evolution. There is a widespread interdependence of genetic and epigenetic variation[82]; however, epigenetic alterations have been reported to diversify at rates that are orders of magnitude higher than those of somatic genetic alterations.[83] Considerable epigenetic reprogramming has also been observed between primary tumors and metastases with little genetic differences.[84]

Strikingly, all cancers that have been investigated thus far with respect to epigenetic or transcriptional variability, showed a strong increase of variability in tumor samples compared to matched normal tissue.[3,74,77] In addition, the differences in variability between cancer and normal samples are remarkably higher than differences in mean levels.[3,43,75,77]

4.3. Variability in the Emergence of Therapeutic Resistance

Heterogeneity in disease is also key in therapeutic resistance. Cancer relapse has traditionally been associated with genetic differences of cellular subclones, proliferative status, or the micro-environment enabling some cells to survive treatment.[85] However, there is accumulating evidence that the variability of cells in therapeutic response is governed by non-genetic factors.[85,86] Even within a subclone, cells display variability leading to differences in function and unequal responsiveness to therapy, and therefore allowing a subset of cells to escape treatment.[87]

Resistance arises through a combination of stochastic fluctuations, adaptation, and epigenetic inheritance.[88–90] For example, variability in protein levels of receptor proteins mediating cell death was linked to cells escaping tumor necrosis factor related apoptosis-inducing ligand (TRAIL) exposure.[85] Similarly, cell-to-cell variation in p53 expression was reported to enable cells to survive in response to chemotherapy.[91] Escaping cells are not only resistant to death but also protected against future death stimuli.[90]

The failure of most cancer therapies to achieve durable responses has typically been associated with intra-tumor heterogeneity,[89,91] yet variability and its prognostic value is not assessed in most clinical trials.[89] It has already been demonstrated that intra-tumor epigenetic heterogeneity is a valuable prognostic marker.[89,92] In chronic lymphocytic leukemia, for example, DNA methylation signatures provide important information about clinically relevant disease subgroups,[93,94] and Sheffield et al.[95] reported that DNA methylation patterns of Ewing sarcoma tumors reflected a continuous disease spectrum. Heterogeneity was not due to genetic variation in these diseases, and it has been hypothesized that differing DNA methylation profiles reflect the differentiation stage of the cell from which the tumors originate.[93,95] For many other cancer types, however, the picture is less clear.[89] In solid tumors, additional variability can come from tumor-adjacent stroma cells with different epigenetic and transcriptional profiles. Therefore, it is important to characterize the variability present in normal samples as well.[89] Taken together, both epigenetic and transcriptional variability may serve as effective biomarkers for personalized medicine.[75,80,89]

An improved understanding of both intra- and inter-individual variability and disentangling the relationship between different layers of variability and the underlying mechanisms will be key to achieve better therapeutic outcomes.[31] In particular, the integration of normal tissue epigenomics and variability across individuals will be essential to promote the understanding of onset and heterogeneity in disease.[80]

5. Current Challenges and Outlook

Given the increasing attention of the biomedical community on the topic of variability and the wealth of data being produced, we argue that a number of challenges in the field need to be considered and overcome. Below, we distinguish and discuss 1) methodological and technical challenges; 2) challenges in advancing fundamental biological understanding; and 3) challenges in translating this knowledge into the clinic and health care management.

5.1. Methodological and Technical Challenges

A critical methodological issue in the study of variability is teasing apart technical variability from biological variability, or in other words, improving ways to remove technical variability without underestimating or compromising the ability to detect biological variability. Even subtle differences in sample collection, handling, transport, storage, experimental procedures, reagents, instruments, and other acts of measuring will lead to technical variability. An intriguing example is described by Lithgow et al.[96] in a comment about the "long journey to reproducible results" across laboratories working on aging in worms. All factors potentially contributing to variability should be minimized as much as possible by applying a strict standardization of every implicated procedure, and by collecting a sufficient number of measurements and replicates. However, technical noise is also inherent to sample processing techniques such as microarray hybridization or sequencing. Thus, the successful detection and removal of technical variability is crucial for the subsequent computational data processing.

This is particularly relevant in the analysis of single cell data, which is still facing limitations due to low signal resolution and reduced signal-to-noise ratio compared to the analysis of data derived from traditional bulk sample approaches.[89,97] Major advances have already been achieved,[98,99] and as the field continues to grow, we expect these hurdles to be overcome in the near future. Methods integrating different layers of regulation in single cells have already been developed[100–102] and will provide unprecedented insights into cell-to-cell variability in molecular regulatory networks and the interactions and relationships between different layers of gene regulation and cellular phenotypes.

Another aspect often hindering the meaningful analysis of biological variability is reduced statistical power, as the sample size for variability analyses needs to be considerably increased compared to traditional studies of mean differences between disease and control groups, for example.[103] Furthermore, the correlation structure between variability and other features must be taken into account. For example, increased gene expression variability is associated with transcript length, while the number of expressed transcripts shows a strong negative correlation with gene expression variability.[46] Technical variability due to intrinsic experimental noise is greater for lowly expressed genes than for those expressed at high levels.[104,105] In the simplest case, this can be due to the law of large numbers: a difference of one or two reads has a bigger impact on a lowly expressed gene with few reads than on a highly expressed gene with many reads. For DNA methylation, there also exists a relationship between mean DNA methylation levels and variability, with increased variability generally observed for intermediate DNA methylation.[106] Thus, it is critical to obtain a measurement of variability that is independent of the mean to ensure to not confound changes in variability with shifts in mean.

5.2. Challenges in Advancing Fundamental Biological Understanding

Beyond the technical and methodological challenges described above, the analysis of variability also opens new questions related to the interpretation of the underlying biological mechanisms. For example, it remains difficult to disentangle the interdependencies and relationships between different layers of gene regulation and how, in combination, they lead to variability and plasticity in phenotypes.[31,107] How the organization of the genome inside the nucleus could potentially modulate the variability of methylation and expression in different parts of the genome also remains unknown. Variability might be affected by properties of the DNA at different levels, from the base-pair/nucleosome scale all the way to the general 3D organization of the DNA polymer inside the nucleus. Recent results in advanced microscopy suggest a very flexible and stochastic distribution of nucleosomes on the DNA, disfavoring the commonly accepted concept of a regular organization of DNA into a 30 nm-fiber structure.[108,109] The characteristics of nucleosome organization are likely to affect pluripotency and cell fate specification, which

suggests that the underlying biology that we are starting to discover will shed further light on new determinants of gene expression variability and phenotypic plasticity.

Another unresolved question is the relative contribution of different sources of variability to the phenotype. The mapping of environmental, genetic, and non-genetic factors has already revealed important associations with inter-individual variability of immune cell function.[18,26,110] As described above, variability is also observed due to stochasticity, which is thought to form the basis of adaptation and population robustness. However, a study by Battich et al.[44] argues that the contribution of stochastic events to variability might be smaller than previously thought. The authors reported that variability between genetically identical cells exposed to the same culture conditions is only minimally stochastic and can be predicted with multivariate models of the phenotypic state and population context of single cells. High-content imaging-based transcriptomics in single human cells showed that stochastic fluctuations in gene expression are buffered by nuclear retention, indicating that cellular compartmentalization confines transcriptional noise to the nucleus, where variability was shown to be higher than in the cytoplasm. Therefore, cellular compartmentalization acting as a passive noise-filter is an effective mechanism to buffer stochastic fluctuations arising from transcriptional bursts for most genes.[44,111] On the other hand, altered chromatin mobility between different sub-compartments of the nucleus has been associated with increased stochastic variation, along with dysregulated spatial separation between active and inactive chromatin environments.[80] The extent of the contribution of stochastic events to phenotypic variation of multi-cellular organisms remains to be determined.

The observations on buffering fluctuations through cellular compartmentalization could partly explain why the correlation between gene expression and protein levels is often weak.[8,112] The translation of RNA expressed from the same gene in a single cell has been shown to be very variable and to occur in bursts as for gene expression.[112] The regulation of protein synthesis could occur globally, or act on each RNA molecule individually and be context dependent,[113] for example, in time (e.g., cell cycle position) and space (e.g., subcellular location).[107,112] It has also been shown that eQTLs tend to have reduced effect size on protein levels, and protein QTLs often do not show effects on RNA and are thus likely to arise from post-translational regulation.[114]

5.3. Challenges in the Translation to the Clinic and Health Care Management

The ultimate challenge ahead is to translate these fundamental biological insights into the clinic and health care management. Gaining knowledge of the different layers of variability and their relation to the phenotype, as well as how variability is generated, controlled, and maintained mechanistically will be fundamental to broaden our understanding of human health and disease. For example, variability could be exploited to achieve a specific immune response against a pathogen or disease.[31] Another approach would be to reduce variability in diseases that thrive on increased variability and dysregulation. There is great potential for drugs reducing epigenetic and transcriptional variability via epigenetic modifications, as these are generally reversible.[89]

Furthermore, it is vital to take inter-individual variability into account in clinical research to achieve better patient stratification and more tailored therapies. For example, individual epigenetic and transcriptional profiles can be used to predict drug response.[115,116] Host-intrinsic factors such as age, sex, or genetic differences due to ethnicity can also cause very different responses to treatments and should be taken into account. There is a strong bias toward white Western participants in clinical trials, and a number of drugs have been shown to be harmful to other ethnic groups.[117,118] Thus, diversity of participants in research studies and clinical trials is essential, and we need to shift the focus onto the variability between individuals, instead of average responses.

Taken together, an increased understanding of biological variability and taking both inter- and intra-individual variability into account in healthcare management and therapeutic approaches will be fundamental to promote human health by personalizing prevention, diagnosis, and treatments.

6. Conclusion

The advancement of single-cell technology and the generation of large-scale multi-omics data enable deeper investigations into the essence of variability in biological systems at the intra- and inter-individual levels. Several factors might underlie variability, such as individual-intrinsic, environmental, and stochastic effects. We have discussed how genetic and epigenetic factors can affect variability in the expression of specific genes, and how this might be influenced by chromatin organization in the nucleus and the genes' context in interaction networks. Together, epigenetic and transcriptional variability ultimately affect phenotypic heterogeneity and plasticity, an essential characteristic in human health and disease.

A better understanding of variability in both health and disease will pave the way for improving human health care. Elucidating the interaction between genetic and non-genetic individual-intrinsic factors and the environment, as well as how they influence the phenotype through epigenetic and transcriptional effects, will be crucial.[110,119] A precise understanding of cell- and individual-intrinsic variability, the environmental sources of variability, and how these effects are propagated and accumulate throughout life will also allow to reveal to what extent environmental or behavioral intervention could have an impact on modulating disease risks.[33] This way, personalized monitoring and adaptation of lifestyle factors such as diet or physical activity[120,121] could enable the combination of precision medicine with disease prevention, the ultimate goal of successful health care management. That this is achievable has been successfully exemplified by the Pioneer 100 Wellness Project (P100)[121] and the Blue Zones Vitality Project,[122] which showed that modulating the environment and behavior of a population can indeed lead to improved health and life-expectancy.

Thus, by bringing variability into the focus, we hope to look forward into a future of improved therapeutic approaches, personalized medicine, and preventive health care management.

Acknowledgments

SE and SB are funded by the UK National Institute for Health Research UCLH Biomedical Research Center (BRC84/CN/SB/5984) and the H2020 Project MultipleMS (733161). VP and AV are funded by the Joint BSC-IRB-CRG Program in Computational Biology and Severo Ochoa Award (SEV 2015–0493). The MRC/BHF Cardiovascular Epidemiology Unit is supported by the UK Medical Research Council (MR/L003120/1), British Heart Foundation (RG/13/13/30194), and NIHR Cambridge Biomedical Research Centre. We thank Allan Orozco for providing critical comments on the manuscript, and Laura Phipps for proofreading the text.

Conflicts of Interest

The authors declare no conflict of interest.

Keywords

3D chromatin structure, biological noise, chromatin, DNA methylation, epigenetics, gene expression, genome architecture, heterogeneity, transcription, variability

Received: August 10, 2017
Revised: October 31, 2017
Published online: December 18, 2017

[1] K. Gärtner, *Lab. Anim*. **1990**, *24*, 71.
[2] M. B. Elowitz, A. J. Levine, E. D. Siggia, P. S. Swain, *Science* **2002**, *297*, 1183.
[3] E. Pujadas, A. P. Feinberg, *Cell* **2012**, *148*, 1123.
[4] A. P. Feinberg, R. A. Irizarry, *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 1757.
[5] H. Dueck, J. Eberwine, J. Kim, *Bioessays* **2016**, *38*, 172.
[6] A. M. Hulse, J. J. Cai, *Genetics* **2013**, *193*, 95.
[7] Z. Bódi, Z. Farkas, D. Nevozhay, D. Kalapis, V. Lázár, B. Csörgő, Á. Nyerges, B. Szamecz, G. Fekete, B. Papp, H. Araújo, L. Oliveira, G. Moura, M. A. S. Santos, T. J. Székely, G. Balázsci, C. Pál, *PLoS Biol*. **2017**, *15*, e2000644.
[8] H. Quach, L. Quintana-Murci, *J. Exp. Med*. **2017**, *214*, 877.
[9] K. E. Sears, J. A. Maier, M. Rivas-Astroza, R. Poe, S. Zhong, K. Kosog, J. D. Marcot, R. R. Behringer, C. J. Cretekos, J. J. Rasweiler, Z. Rapti, *PLoS Genet*. **2015**, *11*, 1.
[10] F. Isbell, D. Craven, J. Connolly, M. Loreau, B. Schmid, C. Beierkuhnlein, T. M. Bezemer, C. Bonin, H. Bruelheide, E. de Luca, A. Ebeling, J. N. Griffin, Q. Guo, Y. Hautier, A. Hector, A. Jentsch, J. Kreyling, V. Lanta, P. Manning, S. T. Meyer, A. S. Mori, S. Naeem, P. A. Niklaus, H. W. Polley, P. B. Reich, C. Roscher, E. W. Seabloom, M. D. Smith, M. P. Thakur, D. Tilman, B. F. Tracy, W. H. van der Putten, J. van Ruijven, A. Weigelt, W. W. Weisser, B. Wilsey, N. Eisenhauer, *Nature* **2015**, *526*, 574.
[11] E. Y. Alemu, J. W. Carl, H. Corrada Bravo, S. Hannenhalli, *Nucleic Acids Res*. **2014**, *42*, 3503.
[12] B. Lehner, K. Kaneko, *Cell Mol. Life Sci*. **2011**, *68*, 1005.
[13] J. Li, Y. Liu, T. Kim, R. Min, Z. Zhang, *PLoS Comput. Biol*. **2010**, *6*, e1000910.
[14] S. Huang, *Development* **2009**, *136*, 3853.
[15] D. Dong, X. Shao, N. Deng, Z. Zhang, *Nucleic Acids Res*. **2011**, *39*, 403.
[16] Roadmap Epigenomics Consortium, *Nature* **2015**, *518*, 317.
[17] H. G. Stunnenberg, The International Human Epigenome Consortium, M. Hirst, *Cell* **2016**, *167*, 1145.
[18] S. Ecker, L. Chen, V. Pancaldi, F. O. Bagger, J. M. Fernández, E. Carrillo de Santa Pau, D. Juan, A. L. Mann, S. Watt, F. P. Casale, N. Sidiropoulos, N. Rapin, A. Merkel, BLUEPRINT Consortium, H. G. Stunnenberg, O. Stegle, M. Frontini, K. Downes, T. Pastinen, T. W. Kuijpers, D. Rico, A. Valencia, S. Beck, N. Soranzo, D. S. Paul, *Genome Biol*. **2017**, *18*, 18.
[19] U. Klein, Y. Tu, G. A. Stolovitzky, M. Mattioli, G. Cattoretti, H. Husson, A. Freedman, G. Inghirami, L. Cro, L. Baldini, A. Neri, A. Califano, R. Dalla-Favera, *J. Exp. Med*. **2001**, *194*, 1625.
[20] P. G. Ferreira, P. Jares, D. Rico, G. Gomez-Lopez, A. Martinez-Trillos, N. Villamor, S. Ecker, A. Gonzalez-Perez, D. G. Knowles, J. Monlong, R. Johnson, V. Quesada, A. Gouin, S. Djebali, M. Lopez-Guerra, D. Colomer, C. Royo, M. Cazorla, M. Pinyol, G. Clot, M. Aymerich, M. Rozman, M. Kulis, D. Tamborero, P. Papasaikas, J. Blanc, M. Gut, I. Gut, X. S. Puente, D. G. Pisano, J. I. Martin-Subero, N. Lopez-Bigas, A. Lopez-Guillermo, A. Valencia, C. Lopez-Otin, E. Campo, R. Guigo, *Genome Res*. **2014**, *24*, 212.
[21] S. Ecker, V. Pancaldi, D. Rico, A. Valencia, *Genome Med*. **2015**, *7*, 8.
[22] D. S. Paul, A. E. Teschendorff, M. A. N. Dang, R. Lowe, M. I. Hawa, S. Ecker, H. Beyan, S. Cunningham, A. R. Fouts, A. Ramelius, F. Burden, S. Farrow, S. Rowlston, K. Rehnstrom, M. Frontini, K. Downes, S. Busche, W. A. Cheung, B. Ge, M.-M. Simon, D. Bujold, T. Kwan, G. Bourque, A. Datta, E. Lowy, L. Clarke, P. Flicek, E. Libertini, S. Heath, M. Gut, I. G. Gut, W. H. Ouwehand, T. Pastinen, N. Soranzo, S. E. Hofer, B. Karges, T. Meissner, B. O. Boehm, C. Cilio, H. Elding Larsson, Å. Lernmark, A. K. Steck, V. K. Rakyan, S. Beck, R. D. Leslie, *Nat. Commun*. **2016**, *7*, 13555.
[23] L. B. Carey, D. van Dijk, P. M. A. Sloot, J. A. Kaandorp, E. Segal, *PLoS Biol*. **2013**, *11*, e1001528.
[24] GTEx Consortium, *Nature* **2017**, *550*, 204.
[25] A. Battle, S. Mostafavi, X. Zhu, J. B. Potash, M. M. Weissman, C. McCormick, C. D. Haudenschild, K. B. Beckman, J. Shi, R. Mei, A. E. Urban, S. B. Montgomery, D. F. Levinson, D. Koller, *Genome Res*. **2014**, *24*, 14.
[26] L. Chen, B. Ge, F. P. Casale, L. Vasquez, T. Kwan, D. Garrido-Martín, S. Watt, Y. Yan, K. Kundu, S. Ecker, A. Datta, D. Richardson, F. Burden, D. Mead, A. L. Mann, J. M. Fernandez, S. Rowlston, S. P. Wilder, S. Farrow, X. Shao, J. J. Lambourne, A. Redensek, C. A. Albers, V. Amstislavskiy, S. Ashford, K. Berentsen, L. Bomba, G. Bourque, D. Bujold, S. Busche, M. Caron, S.-H. Chen, W. Cheung, O. Delaneau, E. T. Dermitzakis, H. Elding, I. Colgiu, F. O. Bagger, P. Flicek, E. Habibi, V. Iotchkova, E. Janssen-Megens, B. Kim, H. Lehrach, E. Lowy, A. Mandoli, F. Matarese, M. T. Maurano, J. A. Morris, V. Pancaldi, F. Pourfarzad, K. Rehnstrom, A. Rendon, T. Risch, N. Sharifi, M.-M. Simon, M. Sultan, A. Valencia, K. Walter, S.-Y. Wang, M. Frontini, S. E. Antonarakis, L. Clarke, M.-L. Yaspo, S. Beck, R. Guigo, D. Rico, J. H. A. Martens, W. H. Ouwehand, T. W. Kuijpers, D. S. Paul, H. G. Stunnenberg, O. Stegle, K. Downes, T. Pastinen, N. Soranzo, *Cell* **2016**, *167*, 1398.
[27] F. Grubert, J. B. Zaugg, L. M. Steinmetz, M. Snyder, F. Grubert, J. B. Zaugg, M. Kasowski, O. Ursu, D. V. Spacek, A. R. Martin, *Cell* **2015**, *162*, 1051.
[28] S. M. Waszak, O. Delaneau, A. R. Gschwind, H. Kilpinen, S. K. Raghav, R. M. Witwicki, A. Orioli, M. Wiederkehr, I. Nikolaos, A. Yurovsky, L. Romano-palumbo, A. Planchon, I. Padioleau, G. Udin, S. Thurnheer, D. Hacker, *Cell* **2015**, *162*, 1039.
[29] A. P. Feinberg, R. A. Irizarry, D. Fradin, M. J. Aryee, V. Gudnason, M. D. Fallin, *Sci. Transl. Med*. **2010**, *2*, 49ra67.
[30] Y. Lu, A. Biancotto, F. Cheung, E. Remmers, N. Shah, J. P. McCoy, J. S. Tsang, *Immunity* **2016**, *45*, 1162.
[31] P. Brodin, M. M. Davis, *Nat. Rev. Immunol*. **2017**, *17*, 21.
[32] N. S. McCarthy, P. E. Melton, G. Cadby, S. Yazar, M. Franchina, E. K. Moses, D. A. Mackey, A. W. Hewitt, *BMC Genomics* **2014**, *15*, 981.
[33] A. Liston, E. J. Carr, M. A. Linterman, *Trends Immunol*. **2016**, *37*, 637.

[34] J. R. Tejedor, M. F. Fraga, *Bioessays* **2017**, *39*, 1700055.
[35] G. Hannum, J. Guinney, L. Zhao, L. Zhang, G. Hughes, S. Sadda, B. Klotzle, M. Bibikova, J.-B. Fan, Y. Gao, R. Deconde, M. Chen, I. Rajapakse, S. Friend, T. Ideker, K. Zhang, *Mol. Cell* **2013**, *49*, 359.
[36] S. Horvath, *Genome Biol*. **2013**, *14*, 3156.
[37] M. F. Fraga, E. Ballestar, M. F. Paz, S. Ropero, F. Setien, M. L. Ballestar, J. C. Cigudosa, M. Urioste, J. Benitez, M. Boix-Chornet, D. Heine-Sun, A. Sanchez-Aguilera, C. Ling, E. Carlsson, P. Poulsen, A. Vaag, Z. Stephan, T. D. Spector, Y.-Z. Wu, C. Plass, M. Esteller, *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 10604.
[38] Q. Tan, B. T. Heijmans, J. V. Hjelmborg, M. Soerensen, K. Christensen, L. Christiansen, *Int. J. Epidemiol*. **2016**, *45*, 1146.
[39] M. J. Jones, S. J. Goodman, M. S. Kobor, *Aging Cell* **2015**, *14*, 924.
[40] G. Jenkinson, E. Pujadas, J. Goutsias, A. P. Feinberg, *Nat. Genet*. **2017**, *49*, 719.
[41] L. M. McEwen, A. M. Morin, R. D. Edgar, J. L. MacIsaac, M. J. Jones, W. H. Dow, L. Rosero-Bixby, M. S. Kobor, D. H. Rehkopf, *Epigenetics Chromatin* **2017**, *10*, 21.
[42] S. Maegawa, S. Gough, N. Watanabe-Okochi, Y. Lu, N. Zhang, R. J. Castoro, M. R. H. Estecio, J. Jelinek, S. Liang, T. Kitamura, P. Aplan, *Genome Res*. **2014**, *24*, 580.
[43] A. P. Feinberg, *Nature* **2007**, *447*, 433.
[44] N. Battich, T. Stoeger, L. Pelkmans, *Cell* **2015**, *163*, 1596.
[45] F. Buettner, K. N. Natarajan, F. P. Casale, V. Proserpio, A. Scialdone, F. J. Theis, S. A. Teichmann, J. C. Marioni, O. Stegle, *Nat. Biotechnol*. **2015**, *33*, 155.
[46] A. J. Faure, J. M. Schmiedel, B. Lehnre, *Cell Syst*. **2017**. https://doi.org/10.1016/j.cels.2017.10.003. [Epub ahead of print].
[47] R. Feil, M. F. Fraga, *Nat. Rev. Genet*. **2012**, *13*, 97.
[48] C. C. Y. Li, J. E. Cropley, M. J. Cowley, T. Preiss, D. I. K. Martin, C. M. Suter, *PLoS Genet*. **2011**, *7*, e1001380.
[49] N. D. Powell, E. K. Sloan, M. T. Bailey, J. M. G. Arevalo, G. E. Miller, E. Chen, M. S. Kobor, B. F. Reader, J. F. Sheridan, S. W. Cole, *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 16574.
[50] S. Voisin, N. Eynon, X. Yan, D. J. Bishop, *Acta. Physiol*. **2015**, *213*, 39.
[51] T. Bauer, S. Trump, N. Ishaque, L. Thürmann, L. Gu, M. Bauer, M. Bieg, Z. Gu, D. Weichenhan, J.-P. Mallm, S. Röder, G. Herberth, E. Takada, O. Mücke, M. Winter, K. M. Junge, K. Grützmann, U. Rolle-Kampczyk, Q. Wang, C. Lawerenz, M. Borte, T. Polte, M. Schlesner, M. Schanne, S. Wiemann, C. Geörg, H. G. Stunnenberg, C. Plass, K. Rippe, J. Mizuguchi, C. Herrmann, R. Eils, I. Lehmann, *Mol. Syst. Biol*. **2016**, *12*, 861.
[52] A. Azzi, R. Dallmann, A. Casserly, H. Rehrauer, A. Patrignani, B. Maier, A. Kramer, S. A. Brown, *Nat. Neurosci*. **2014**, *17*, 377.
[53] M. Farlik, F. Halbritter, F. Müller, F. A. Choudry, P. Ebert, J. Klughammer, S. Farrow, A. Santoro, V. Ciaurro, A. Mathur, R. Uppal, H. G. Stunnenberg, W. H. Ouwehand, E. Laurenti, T. Lengauer, M. Frontini, C. Bock, *Cell Stem Cell* **2016**, *19*, 808.
[54] R. Guantes, A. Rastrojo, R. Pires das Neves, A. Lima, B. Aguado, F. J. Iborra, R. Neves, A. Lima, A. Begoña, F. J. Iborra, *Genome Res*. **2015**, *25*, 633.
[55] V. Pancaldi, E. Carrillo-de-Santa-Pau, B. M. Javierre, D. Juan, P. Fraser, M. Spivakov, A. Valencia, D. Rico, *Genome Biol*. **2016**, *17*, 152.
[56] J. K. Choi, Y.-J. Kim, *Nat. Genet*. **2009**, *41*, 498.
[57] J. B. Zaugg, N. M. Luscombe, *Genome Res*. **2012**, *22*, 84.
[58] G. Kar, J. K. Kim, A. A. Kolodziejczyk, K. N. Natarajan, E. T. Triglia, B. Mifsud, S. Elderkin, J. C. Marioni, A. Pombo, S. A. Teichmann, *Nat. Commun*. **2017**, *8*, 36.
[59] N. Moris, C. Pina, A. M. Arias, *Nat. Rev. Genet*. **2016**, *17*, 693.
[60] J. R. Wagner, S. Busche, B. Ge, T. Kwan, T. Pastinen, M. Blanchette, *Genome Biol*. **2014**, *15*, R37.
[61] I. Huh, J. Zeng, T. Park, S. V Yi, *Epigenetics Chromatin* **2013**, *6*, 9.
[62] D. Coleman-Derr, D. Zilberman, *PLoS Genet*. **2012**, *8*, e1002988.
[63] S. Rudnizky, A. Bavly, O. Malik, L. Pnueli, P. Melamed, A. Kaplan, *Nat. Commun*. **2016**, *7*, 12958.
[64] B. Bonev, G. Cavalli, *Nat. Rev. Genet*. **2016**, *17*, 661.
[65] E. Lieberman-Aiden, N. L. Van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, R. Sandstrom, B. Bernstein, M. A. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L. A. Mirny, E. S. Lander, J. Dekker, *Science* **2009**, *326*, 289.
[66] G. Chalancon, C. N. J. Ravarani, S. Balaji, A. Martinez-Arias, L. Aravind, R. Jothi, M. M. Babu, *Trends Genet*. **2012**, *28*, 221.
[67] E. A. Mason, J. C. Mar, A. L. Laslett, M. F. Pera, J. Quackenbush, E. Wolvetang, C. A. Wells, *Stem Cell Rep*. **2014**, *3*, 365.
[68] B. M. Javierre, O. S. Burren, S. P. Wilder, R. Kreuzhuber, S. M. Hill, S. Sewitz, J. Cairns, S. W. Wingett, C. Várnai, M. J. Thiecke, F. Burden, S. Farrow, A. J. Cutler, K. Rehnström, K. Downes, L. Grassi, M. Kostadima, P. Freire-Pritchett, F. Wang, H. G. Stunnenberg, J. A. Todd, D. R. Zerbino, O. Stegle, W. H. Ouwehand, M. Frontini, C. Wallace, M. Spivakov, P. Fraser, *Cell* **2016**, *167*, 1369.
[69] L. Wolf, O. K. Silander, E. van Nimwegen, *Elife* **2015**, *4*, e05856.
[70] K. Ibáñez, M. Guijarro, G. Pajares, A. Valencia, *Data Min. Knowl. Discov.* **2016**, *30*, 226.
[71] C. R. S. Banerji, S. Severini, C. Caldas, A. E. Teschendorff, *PLoS Comput Biol* **2015**, *11*, e1004115.
[72] M. Busslinger, A. Tarakhovsky, *Cold Spring Harb. Perspect. Biol*. **2014**, *6*, a019307.
[73] R. Satija, A. K. Shalek, *Trends Immunol*. **2014**, *35*, 219.
[74] W. Timp, A. P. Feinberg, *Nat. Rev. Cancer* **2013**, *13*, 497.
[75] A. E. Teschendorff, A. Jones, H. Fiegl, A. Sargent, J. J. Zhuang, H. C. Kitchener, M. Widschwendter, *Genome Med*. **2012**, *4*, 24.
[76] M. Klutstein, J. Moss, T. Kaplan, H. Cedar, *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 2230.
[77] K. D. Hansen, W. Timp, H. C. Bravo, S. Sabunciyan, B. Langmead, O. G. Mcdonald, B. Wen, H. Wu, Y. Liu, D. Diep, E. Briem, K. Zhang, R. A. Irizarry, A. P. Feinberg, *Nat. Genet*. **2011**, *43*, 768.
[78] D. A. Landau, S. L. Carter, G. Getz, C. J. Wu, *Leukemia* **2014**, *28*, 34.
[79] W. A. Flavahan, E. Gaskell, B. E. Bernstein, *Science* **2017**, *357*, eaal2380.
[80] A. P. Feinberg, M. A. Koldobskiy, A. Göndör, *Nat. Rev. Genet*. **2016**, *17*, 284.
[81] A. P. Feinberg, *J. Intern. Med*. **2014**, *276*, 5.
[82] T. Mazor, A. Pankov, B. E. Johnson, C. Hong, E. G. Hamilton, R. J. A. Bell, I. V. Smirnov, G. F. Reis, J. J. Phillips, M. J. Barnes, A. Idbaih, A. Alentorn, J. J. Kloezeman, M. L. M. Lamfers, A. W. Bollen, B. S. Taylor, A. M. Molinaro, A. B. Olshen, S. M. Chang, J. S. Song, J. F. Costello, *Cancer Cell* **2015**, *28*, 307.
[83] K. D. Siegmund, P. Marjoram, Y.-J. Woo, S. Tavaré, D. Shibata, *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 4828.
[84] A. P. Makohon-Moore, M. Zhang, J. G. Reiter, I. Bozic, B. Allen, D. Kundu, K. Chatterjee, F. Wong, Y. Jiao, Z. A. Kohutek, J. Hong, M. Attiyeh, B. Javier, L. D. Wood, R. H. Hruban, M. A. Nowak, N. Papadopoulos, K. W. Kinzler, B. Vogelstein, C. A. Iacobuzio-Donahue, *Nat. Genet*. **2017**, *49*, 358.
[85] S. L. Spencer, S. Gaudet, J. G. Albeck, J. M. Burke, P. K. Sorger, *Nature* **2009**, *459*, 428.
[86] S. M. Shaffer, M. C. Dunagin, S. R. Torborg, E. A. Torre, B. Emert, C. Krepler, M. Beqiri, K. Sproesser, P. A. Brafford, M. Xian, E. Eggan, I. N. Anastopoulos, C. A. Vargas-Garcia, A. Singh, K. L. Nathanson, M. Kerlyn, A. Raj, *Nature* **2017**, *546*, 431.
[87] A. Kreso, J. E. Dick, *Cell Stem Cell* **2014**, *14*, 275.
[88] T. Day, *Mol. Ecol*. **2016**, *25*, 1869.
[89] T. Mazor, A. Pankov, J. S. Song, J. F. Costello, *Cancer Cell* **2016**, *29*, 440.
[90] F. Bertaux, S. Stoma, D. Drasdo, G. Batt, *PLoS Comput. Biol*. **2014**, *10*, 1003893.

[91] A. L. Paek, J. C. Liu, A. Loewer, W. C. Forrester, G. Lahav, *Cell* **2016**, *165*, 631.
[92] D. A. Landau, K. Clement, M. J. Ziller, P. Boyle, J. Fan, H. Gu, K. Stevenson, C. Sougnez, L. Wang, S. Li, D. Kotliar, W. Zhang, M. Ghandi, L. Garraway, S. M. Fernandes, K. J. Livak, S. Gabriel, A. Gnirke, E. S. Lander, J. R. Brown, D. Neuberg, P. V. Kharchenko, N. Hacohen, G. Getz, A. Meissner, C. J. Wu, *Cancer Cell* **2014**, *26*, 813.
[93] M. Kulis, S. Heath, M. Bibikova, A. C. Queirós, A. Navarro, G. Clot, A. Martínez-Trillos, G. Castellano, I. Brun-Heath, M. Pinyol, S. Barberán-Soler, P. Papasaikas, P. Jares, S. Beà, D. Rico, S. Ecker, M. Rubio, R. Royo, V. Ho, B. Klotzle, L. Hernández, L. Conde, M. López-Guerra, D. Colomer, N. Villamor, M. Aymerich, M. Rozman, M. Bayes, M. Gut, J. L. Gelpí, M. Orozco, J.-B. Fan, V. Quesada, X. S. Puente, D. G. Pisano, A. Valencia, A. López-Guillermo, I. Gut, C. López-Otín, E. Campo, J. I. Martín-Subero, *Nat. Genet*. **2012**, *44*, 1236.
[94] A. C. Queirós, N. Villamor, G. Clot, A. Martinez-Trillos, M. Kulis, A. Navarro, E. M. M. Penas, S. Jayne, A. Majid, J. Richter, A. K. Bergmann, J. Kolarova, C. Royo, N. Russiñol, G. Castellano, M. Pinyol, S. Bea, I. Salaverria, M. López-Guerra, D. Colomer, M. Aymerich, M. Rozman, J. Delgado, E. Giné, M. González-Díaz, X. S. Puente, R. Siebert, M. J. S. Dyer, C. López-Otín, C. Rozman, E. Campo, A. López-Guillermo, J. I. Martín-Subero, *Leukemia* **2014**, *29*, 598.
[95] N. C. Sheffield, G. Pierron, J. Klughammer, P. Datlinger, A. Schönegger, M. Schuster, J. Hadler, D. Surdez, D. Guillemot, E. Lapouble, P. Freneaux, J. Champigneulle, R. Bouvier, D. Walder, I. M. Ambros, C. Hutter, B. Liegl-atzwanger, B. Huppertz, A. Leithner, G. De Pinieux, P. Terrier, V. Laurence, J. Michon, R. Ladenstein, W. Holter, R. Windhager, U. Dirksen, P. F. Ambros, O. Delattre, H. Kovar, C. Bock, E. M. Tomazou, *Nat. Med*. **2017**, *23*, 386.
[96] G. J. Lithgow, M. Driscoll, P. Phillips, *Nature* **2017**, *548*, 387.
[97] R. Bacher, C. Kendziorski, *Genome Biol*. **2016**, *17*, 63.
[98] S. Linnarsson, S. A. Teichmann, *Genome Biol*. **2016**, *17*, 97.
[99] E. Papalexi, R. Satija, *Nat. Rev. Immunol*. **2017**. https://doi.org/10.1038/nri.2017.76. [Epub ahead of print].
[100] I. C. Macaulay, C. P. Ponting, T. Voet, *Trends Genet*. **2017**, *33*, 155.
[101] C. Bock, M. Farlik, N. C. Sheffield, *Trends Biotechnol*. **2016**, *34*, 605.
[102] C. Angermueller, S. J. Clark, H. J. Lee, I. C. Macaulay, M. J. Teng, T. X. Hu, F. Krueger, S. A. Smallwood, C. P. Ponting, T. Voet, G. Kelsey, O. Stegle, W. Reik, *Nat. Methods* **20516**, *13*, 229.
[103] B. Phipson, A. Oshlack, *Genome Biol*. **2014**, *15*, 465.
[104] O. Stegle, S. A. Teichmann, J. C. Marioni, *Nat. Rev. Genet*. **2015**, *16*, 133.
[105] C. A. Vallejos, S. Richardson, J. C. Marioni, *Genome Biol*. **2016**, *17*, 70.
[106] P. Du, X. Zhang, C.-C. Huang, N. Jafari, W. A. Kibbe, L. Hou, S. M. Lin, *BMC Bioinformatics* **2010**, *11*, 587.
[107] B. Snijder, L. Pelkmans, *Nat. Rev. Mol. Cell Biol*. **2011**, *12*, 119.
[108] M. Ricci, C. Manzo, M. García-Parajo, M. Lakadamyali, M. Cosma, *Cell* **2015**, *160*, 1145.
[109] H. D. Ou, S. Phan, T. J. Deerinck, A. Thor, M. H. Ellisman, C. C. O'Shea, *Science* **2017**, *357*, eaag0025.
[110] R. ter Horst, M. Jaeger, S. P. Smeekens, M. Oosting, M. A. Swertz, Y. Li, V. Kumar, D. A. Diavatopoulos, A. F. M. Jansen, H. Lemmers, H. Toenhake-Dijkstra, A. E. van Herwaarden, M. Janssen, R. G. van der Molen, I. Joosten, F. C. G. J. Sweep, J. W. Smit, R. T. Netea-Maier, M. M. J. F. Koenders, R. J. Xavier, J. W. M. van der Meer, C. A. Dinarello, N. Pavelka, C. Wijmenga, R. A. Notebaart, L. A. B. Joosten, M. G. Netea, *Cell* **2016**, *167*, 1111.
[111] T. Stoeger, N. Battich, L. Pelkmans, *Cell* **2016**, *164*, 1151.
[112] X. Yan, T. A. Hoek, R. D. Vale, M. E. Tanenbaum, *Cell* **2016**, *165*, 976.
[113] Y. Liu, R. Aebersold, *Mol. Syst. Biol*. **2016**, *12*, 856.
[114] A. Battle, Z. Khan, S. H. Wang, A. Mitrano, M. J. Ford, J. K. Pritchard, Y. Gilad, *Science* **2015**, *347*, 664.
[115] Z. A. Gurard-Levin, L. O. W. Wilson, V. Pancaldi, S. Postel-Vinay, F. G. Sousa, C. Reyes, E. Marangoni, D. Gentien, A. Valencia, Y. Pommier, P. Cottu, G. Almouzni, *Mol. Cancer Ther*. **2016**, *15*, 1768.
[116] F. Azuaje, *Br. Bioinform*. **2017**, *18*, 820.
[117] N. J. Schork, *Nature* **2015**, *520*, 609.
[118] E. G. Cohn, G. E. Henderson, P. S. Appelbaum, *Genet. Med*. **2017**, *19*, 157.
[119] J. L. Pappalardo, D. A. Hafler, *Cell* **2016**, *167*, 894.
[120] X. Li, J. Dunn, D. Salins, G. Zhou, W. Zhou, S. M. Schüssler-Fiorenza Rose, D. Perelman, E. Colbert, R. Runge, S. Rego, R. Sonecha, S. Datta, T. McLaughlin, M. P. Snyder, *PLoS Biol*. **2017**, *15*, e2001402.
[121] N. D. Price, A. T. Magis, J. C. Earls, G. Glusman, R. Levy, C. Lausted, D. T. McDonald, U. Kusebauch, C. L. Moss, Y. Zhou, S. Qin, R. L. Moritz, K. Brogaard, G. S. Omenn, J. C. Lovejoy, L. Hood, *Nat. Biotechnol*. **2017**, *35*, 747.
[122] D. Buettner, S. Skemp, *Am. J. Lifestyle Med*. **2016**, *10*, 318.