

PERSPECTIVE ARTICLE

A Postgenomic Perspective on Molecular Cytogenetics

Henry H. Heng^{1,2,*}, Steven D. Horne¹, Sophia Chaudhry¹, Sarah M. Regan¹, Guo Liu¹, Batoul Y. Abdallah¹ and Christine J. Ye³

¹Center for Molecular Medicine and Genetics, Wayne State University School of Medicine, Detroit, MI, USA; ²Department of Pathology, Wayne State University School of Medicine, Detroit, MI, USA; ³The Division of Hematology/Oncology, University of Michigan Comprehensive Cancer Center, Ann Arbor, MI, USA

Abstract: Background: The postgenomic era is featured by massive data collection and analyses from various large scale-omics studies. Despite the promising capability of systems biology and bioinformatics to handle large data sets, data interpretation, especially the translation of -omics data into clinical implications, has been challenging.

Discussion: In this perspective, some important conceptual and technological limitations of current systems biology are discussed in the context of the ultimate importance of the genome beyond the collection of all genes. Following a brief summary of the contributions of molecular cytogenetics/cytogenomics in the pre- and post-genomic eras, new challenges for postgenomic research are discussed. Such discussion leads to a call to search for a new conceptual framework and holistic methodologies.

Conclusion: Throughout this synthesis, the genome theory of somatic cell evolution is highlighted in contrast to gene theory, which ignores the karyotype-mediated higher level of genetic information. Since “system inheritance” is defined by the genome context (gene content and genomic topology) while “parts inheritance” is defined by genes/epigenes, molecular cytogenetics and cytogenomics (which directly study genome structure, function, alteration and evolution) will play important roles in this postgenomic era.

ARTICLE HISTORY

Received: October 21, 2016

Revised: January 29, 2017

Accepted: February 03, 2017

DOI:

10.2174/1389202918666170717145716

Keywords: Systems biology, Genome theory, System inheritance, Parts inheritance, Karyotype coding, Fuzzy inheritance, Genome re-organization, Genome chaos, Heterogeneity.

1. INTRODUCTION

The success of sequencing the human genome in the early 2000s marked a historical transition of genomic research from the pre- to post-genomic era [1]. Perhaps also due to a common misconception that finishing the genome sequencing phase means the mission for genome-level research has been accomplished, the term post-genome era has been used, as if decoding DNA by sequencing is equal to decoding the genome itself [2-4]. Furthermore, some have even suggested that the goal of this post-genome era is to search for molecular mechanisms beyond the genome. In fact, “beyond the genome” has become a buzzword, even though it is clear that the genome is not just a bag of all genes and DNA sequences, and the functional organization of the genome is virtually unknown. Nevertheless, based on the rationale of searching for the answer beyond the genome, many -omics studies, featured by large-scale and high-throughput technological platforms, have been introduced, armed with both cutting edge computational tools and advanced bioinformatics. Together, these efforts have pushed systems biology into an exciting new period [5-12].

It is understandable why so many researchers eagerly move to the next type of -omics beyond the genome. Besides passion for scientific novelty and the influence of the biotechnology industry resulting in a push to acquire and apply the newest technological platform, the increased disappointment in genomics plays an important role [13-19]. Obviously, the major promises of sequencing the human genome did not pay off in terms of solving the mystery of life or providing an understanding of the genetic basis for most common human diseases. Such a disappointment has quickly led to the common viewpoint that: “If the answer cannot be obtained from genome sequencing, why not move beyond the genome?” On the surface, this seems logical. However, this view is based on a deep misunderstanding of genetic organization, because the genome system is defined by genomic context rather than gene content [2]. Therefore, one cannot decode the genome system by just sequencing its parts (DNA), or illustrating the pattern of protein interaction, or understanding the potential of epigenetic dynamics. The molecular profiles of DNA, proteins and RNA represent some important features of the emergent system where the genome is the structural and functional platform [3, 4]. In other words, many features of bio-systems, such as protein interaction pattern and network dynamics (including systems-specific boundaries), are defined by the genome context. Characterization of the genome’s features or behavior is not beyond genomic research; it is the very core of it.

*Address correspondence to this author at 3226 Scott Hall, 540 E. Canfield, Detroit, MI 48201, USA; Tel: 313-577-9544; Fax: 313-577-6100; E-mail: hheng@med.wayne.edu

Over ten years have passed since various -omics have been introduced. Following initial expectations and many promising publications, the excitement seems to have reached its plateau for some, and yet, additional new -omics continue to emerge. It is true that much more data has been collected, from transcriptomics, to proteomics, and to metabolomics, but so far, there has been no fundamental change at the conceptual level, as the relationship among genomics and other -omics simply reflects the relationship of information layers between genotype and phenotype. Furthermore, many fundamental explanations of various -omics data are still based on traditional genetics/genomics principles, which are within the framework of pre-genomics. Even more puzzling, a large amount of data has further challenged (rather than supported) much of our basic thinking in genetics and protein biology. For example, most of the large-scale whole genome scanning effort has failed to deliver, despite publicized success [20-24]; the highly anticipated personal genome project has generated more genetic variations than we can explain; most transcripts do not contribute to protein synthesis; a large portion of newly synthesized proteins immediately go through the protein degradation process; the once very exciting protein microarray has revealed more non-specific “noise” than the patterns we believed to exist, especially in clinical samples; many key proteins have thousands of potential interaction partners, and the significance of molecular specificity (*e.g.* the degree of affinity) is drastically reducing; key metabolic pathways can be quickly altered under stress, and types of metabolic signatures are often the results of evolutionary selection rather than simply molecular availability; drug discovery based on large-scale genetic screening plus computational design has negatively impacted the drug pipeline for pharmaceutical companies [4]. Even the general attitude toward systems biology has changed. Just a decade ago, systems biology offered so much hope. Now, having the capability to display different proteins in different colorful hairballs, and with the knowledge that almost every protein is linked and can be classified into hub, link and functional modules, it is still very challenging to apply systems biology to predicting a system’s response and/or evolutionary direction under stress. Due to its dynamic nature, it is often difficult even to repeat the same pattern of protein networks by repeating experiments. Moreover, the current network is constructed using average profiles, and it is known that average cells are not the key players for macro-cellular evolution, while outliers are highly relevant for disease conditions [4, 25]. No wonder the clinical implications are extremely limited. Even some leading cancer biologists are now questioning the promise of relying on large amounts of data plus powerful bioinformatics analysis [26].

Clearly, the various -omics communities in the post-genomic era need to face this reality check and ask critical questions, rather than just collecting more data based on good wishes. In particular, why have most of the promises of -omics not yet been realized? Why is it that while many impressive individual research papers have come out, and more data have been accumulated, the clinical implications of the field have so far disappointed? What have we missed in our efforts to continually push data generation and analyses?

To address these questions, we have compared some important surprises and confusions following various -omics studies, especially large-scale genomic studies, in the context of disease research. These syntheses have led to three realizations: First, the century-long believed power of the gene is significantly reducing, and the importance of the genome is regaining deserved attention [4, 13, 14, 27, 28]. As the correlation between individual gene and disease phenotype has reduced (not just many gene mutations have been linked to each common and complex disease type, but a high number of mutations are also detected from normal tissues) [29, 30], identifying the correct target has become a priority. While most researchers are now looking for epigenetic mechanisms and non-coding regions, we proposed to re-focus on karyotype dynamics. Second, overwhelming multiple levels of genetic and non-genetic heterogeneity-mediated complexity exist, and the higher genome level can often overpower the lower gene level when there is a conflict [31-37], which explains why the gene-centric-defined genetic determinists’ approach is not going to work. However, despite the fact that the importance of heterogeneity in cancer has been known for decades, most molecular researchers have been more or less ignoring this issue, due to the belief that gene mutation specificity will be dominant in cancer. Now, the cancer genome sequencing project has proven otherwise, and interest in studying heterogeneity is strongly coming back. Interestingly, molecular cytogenetics approaches can play an important role in studying the mechanism of heterogeneity [37-40]. Third, there is an urgent need to rethink the genetic/genomic theories, as well as evolutionary theory, in the light of cellular evolution and disease initiation and progression. In particular, we must consider how genetic information is passed, why the separation of germline and somatic cells is essential for balancing micro- and macro-evolution, and what determines the dynamic relationship between short-term adaptation and long-term survival. As we will discuss in a later section, the knowledge of heterogeneity and how the outlier’s behavior drives evolution will significantly change the traditional clonal expansion concept [4, 25, 37, 41-43]. Together, these three realizations call for a new conceptual framework, as we have the data, and will have much more every day, but we do not have the correct framework to interpret them.

In the past two decades, we and other researchers have worked hard to search for a new framework to depart from the gene-centric view of current genomics [2]. In fact, the three realizations discussed above are direct results of pushing karyotype evolutionary research within the context of somatic cell evolution. Even though it is timely to introduce the viewpoints of systems biology, bioinformatics, and computational tools into molecular cytogenetics [12, 44], on this occasion, we would like discuss more about how to use new cytogenetics concepts and strategies to establish the correct genetic and evolutionary foundation of many -omics research topics. To achieve this goal, we refer to familiar and convenient data sets to prove a point (rather than mention a diverse and comprehensive array of literature sources; we thus owe an apology for not citing other similar data sets). A brief review will discuss how karyotype evolutionary studies have pinpointed the limitation of gene mutations in understanding cancer, identified new types of inheritance, and

even discovered the new pattern of cellular evolution. These discussions will clearly point out the importance of integrating molecular cytogenetics with current functional genomics and human disease studies, and how cytogenetics can provide the genomic basis or platform for various -omics studies where genetics and evolution are intimately involved.

2. CONTRIBUTIONS OF MOLECULAR CYTOGENETICS/CYTOGENOMICS IN PRE- AND POST-GENOMIC ERAS:

Cytogenetics represents a major subfield of genetics/genomics. Since the chromosome has long been considered as a vehicle of genes, it has been closely associated with gene research. In fact, the initial gene theory is mainly based on chromosomal research. Prior to the completion of the sequencing of the human genome, molecular cytogenetics played an important role for helping determine the position of candidate genes for certain diseases. Clonal, chromosomal abnormalities (such as translocations, large-size deletions and duplications detected by cytogenetic methods) have helped narrow down the genomic regions that could host the potential gene, which is crucial to isolating the candidate gene. Examples include the cloning of Bcr-Abl gene from chromosomal translocation t(9;22)(q34;q11) for Chronic Myelogenous Leukemia (CML) [45], the cloning of the Duchenne muscular dystrophy gene from the X chromosome [46, 47], and the identification of the gene for polycystic kidney disease (*PKDI*) from translocation t(16;22)(p13.3;q11.21) [48]. A large number of cancer genes have been cloned with the help of cytogenetic characterization of chromosomal abnormalities in interesting patient cases. From the 1980s to the 1990s, cancer cytogenetics has played a crucial role for gene identification [49]. Interestingly, however, using information from chromosomal abnormalities in order to clone cancer genes has been most successful in blood cancer types, due to the special features of recurrent clonal expansion and different patterns of evolution [50]. For more detailed information regarding a large number of fusion genes and their corresponding translocations, please visit The Atlas of Genetics and Cytogenetics in Oncology and Haematology (a peer-reviewed Internet journal/encyclopaedia/database) (<http://AtlasGeneticsOncology.org>), and the Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer (<http://cgap.nci.nih.gov/Chromosomes/Mitelman>). Interestingly, as soon as the gene is cloned, the chromosomal abnormality itself becomes less attractive to some molecular researchers who favor the value of specific gene over genome. Unfortunately, for most sporadic cancer cases with high levels of genome alterations, individual genes have limited prediction power compared with chromosomal abnormalities. This observation suggests that a chromosomal abnormality not only has an impact on the directly interrupted genes (such as the fusion gene itself), but also changes the genomic context for many genes across the entire genome [2-4].

Another major contribution of molecular cytogenetics was to use effective tools such as Fluorescent *in situ* hybridization, or FISH, to fill in the gaps between genes and chromosomes [51]. Examples include gene mapping along chromosomes [52, 53], validating the location of YACs

(Yeast Artificial Chromosomes) or BACs (Bacterial Artificial Chromosomes) and their chimeric status [54], and helping the construction of a physical map for a specific region of the genome or entire chromosome [55]. In particular, high-resolution fiber FISH was used to estimate the gaps of the physical map [56, 57] and to confirm the correct order for assembled DNA sequencing. Furthermore, FISH was frequently used to study chromosomal structure (of both mitotic and meiotic chromosomes) [58], identify the insert sites for transgenic animals and their chromatin organization [59, 60], illustrate the specific binding of protein and DNA elements [61, 62], analyze meiotic chromosome segregation [63], and help determine the clinical characteristics and diagnoses for human diseases [64].

In the post-genomic era, molecular cytogenetics has pushed higher resolution and large-scale platforms. Compared with molecular biology, cytogenetics has a lower resolution. It is also time-consuming, and karyotype information often lacks specific molecular mechanisms. There was a push to use more molecular platforms to reduce classical cytogenetic analyses; however, as we will discuss later, the karyotype level of analysis is actually essential, and we must not replace it with more gene-centric technologies, as the karyotype represents a new layer of genetic information. In fact, it is now known that for many cancer types, high-resolution sequencing data have limited correlation with clinical outcomes. In contrast, the cytogenetic profile offers the best prediction value. This observation fully supports the importance of using genome-level alterations in cancer prediction, as genetic importance is often disassociated with molecular resolution, and the key is to find the right level of genetic organization [65].

One of the best achievements of molecular cytogenomics in the post-genomic era was the discovery of copy number variations through the use of array CGH [66]. This discovery has generated the excitement of monitoring this level of genetic variations and its biological meaning in understanding human diseases [67]. Various arrays, such as single nucleotide polymorphism genomic microarray, have also been applied to cloned genes involved in chromosome translocations [68, 69], and many fusion genes have been effectively identified. Studying copy number variations also offers hope for identifying the missing heritability [70, 71]. With their high-throughput capability, various array technologies have drastically changed clinical diagnosis. As genomic arrays have been routinely used in many cytogenetic diagnostic laboratories, it has become clear that there is an overwhelming amount of heterogeneity in copy number variations among patients and normal controls, which also challenges the interpretation of these sub-chromosomal variations. Recently, next generation sequencing has promised even higher resolutions and efficiencies for genetic detection. Paradoxically, pushing for this highest resolution might just miss the point if the ignored overall genome landscape is actually more important than individual gene defined genetic components [4]. Furthermore, cytogenetic methods offer comparative information concerning both the individual cell and the cell population, which can effectively monitor the adaptive potential. In contrast, current sequencing efforts are mainly based on average molecular profiling [25].

At the forefront of studying chromatin/chromosome function, research on interphase chromatin domains and their interaction is very active [29, 72, 73], especially when combined with technologies to monitor DNA-protein and protein-protein interactions in living cells. For example, GFP fusion proteins, FLIP (fluorescence loss in photobleaching) or FRAP (fluorescence recovery after photobleaching), FRET (fluorescence resonance energy transfer), and FCS (fluorescence correlation spectroscopy) were used for this purpose [74, 75]. Further visualization methods, such as halo FISH, have been used to study chromatin loop size and its relationship with gene expression [76-78]. It was discovered that anchor sites are very important for the formation of functional unity of the chromatin, and these anchor sites are highly dynamic, which provides the basis for gene regulation. The imaging of various levels of genetic organization represents a major implication for various molecular cytogenetic technologies. We have previously summarized the research activities into four types of studies, comprised of visualization: 1) at the macro-molecular complex level; 2) at the chromatin loop domain level; 3) at the chromosome level to monitor structure and behavior, and 4) at the level of the entire genome (including direct visualization of karyotype changes as the chromatin/chromosome behavior, the chromosomal territories, and the karyotypic changes caused by physiological, pathological and environmental challenges) [38, 79].

A number of unique developments are worthy of being mentioned. First is the usage of cross-species reciprocal painting to study the phylogenetic relationship among different species and to compare the established evolutionary tree based on sequencing data [12, 80-83]. Interestingly, this approach is most informative for species within but not between placental mammals, monotremes, marsupials and birds [84]. Such comparison underscores the importance of focusing on the karyotype level to study the evolution among species and the important mechanism of reorganizing the genome in evolution [2-4, 85].

Second are mechanistic studies of aneuploidy (both the genes/pathways and general stresses that lead to aneuploidy, and the consequences of the aneuploidy being studied) [86]. Interestingly, some of these investigators are yeast biologists, whose research has brought aneuploidy research into the spotlight. For example, aneuploidy can effectively rescue key gene mutations [27], and it is directly linked to transcription and adaptive potential [87]. A similar conclusion has also been illustrated in human cancer cells [88, 89]. In addition to studies that illustrate the relationship between polyploidy and aneuploidy, the common link between chromosomal instability or CIN and diverse molecular mechanisms has been addressed [35, 90-96]. Similarly, cancer drug resistance has been linked to genome-level alterations.

Third, with the increased understanding of the importance of genome-level heterogeneity in cancer evolution, it is necessary to monitor all types of chromosomal abnormalities. Surprisingly, despite that there are many types of chromosomal abnormalities, many of them are still unclassified and have not even been reported or generally accepted, due to the fact that these abnormalities are highly diverse and non-clonal. Since all types of chromosomal aberrations can

be linked to CIN, their quantitative degree can be used to monitor overall genome instability. Recently, CIN has been linked to different types of human diseases, as elevated NCCAs have been detected from many common diseases [37]. Based on stress-elicited, genome alteration-mediated adaptation, and the trade-off of that gain *versus* the loss of genomic stability as pre-conditions for various diseases or illnesses, a general model was proposed of how genetic changes drive somatic cell evolution; this process can both improve cellular function but at the same time increase the risk of disease [97-99]. Under this realization, it is important to study all types of genomic variations and their impact on human variations and diseases. That is the reason why it is important to study the genome chimeric issue [100, 101] as well as a large number of cases with small supernumerary marker chromosomes (sSMC) [102]. It should be noted that, even though a few investigators are pushing the practice of using NCCAs to study genome instability, the majority of the cytogenetic researchers continue to ignore NCCAs with the mindset that if these aberrations are not recurrent, they must not be important. In fact, this alone is the policy of cytogenetic diagnosis. The same idea is very popular in current systems biology.

Fourth, although it has started a bit slower compared to other fields, there is an increased effort to apply bioinformatics into molecular cytogenetic analyses. Currently, examples include the use of various molecular cytogenetic and cytogenomic databases in the interpretation of data, and in comparing cytogenetic data and sequencing data, as well as copy number variation data, as a way to make sense of phenotypes. For example, *in silico* molecular cytogenetics was used to prioritize research among genes [12, 44]. However, it is challenging to understand how to use the genome alteration data. Currently, when there are many genome-level alterations (such as when the genome is chaotic), most researchers only cherry-pick one or a few mutations or pathways to explain the story; they do not realize that the individual gene story no longer makes sense when there are large numbers of drastic genome-level changes.

3. NEW CHALLENGES FOR POSTGENOMIC RESEARCH

3.1. It has Become much Easier to Obtain Data but Incredibly Difficult to Interpret them

Prior to the post-sequencing era, most researchers focused on a limited number of individual genes and associated potential biological and medical implications. A large number of investigators around the globe studied “famous” genes that were thought to have high clinical significance. Such genes were studied in the following manner: linkage and physical mapping prior to gene identification, molecular cloning and characterization, promoter region analysis, mutation spectrum study, identification of genetic modifiers, establishment of transgenic or knockout animal models, developmental and physiological features, up-stream and down-stream interaction partners, and theoretical developments including gene therapy and drug discovery. Due to the limited number of genes that were considered to be critical for the subject of research, competition was extremely high, and the gene cloning associated with major diseases was often an intense race among researchers [103, 104].

In the post-sequencing era, the landscape of research including popular methodologies has drastically changed. In the past, the hardest step was mapping and cloning the gene of interest, which represented a major achievement. The rest of characterization just followed the logical steps of analyses. Now, it is much easier and quicker to obtain deep sequencing data and expression information of known and unknown genes or genetic markers. This is because an individual's genome can be sequenced in days, and most gene information can be found in numerous databases, eliminating the intense work of screening libraries and validating cloned genes. Therefore, new challenges revolve around how to make sense of the high volume of data.

Paradoxically, the massive amount of data only made research more complicated. When a major disease gene was cloned 20-30 years ago, scientists confidently claimed that they would soon find a cure. For the past few decades, the dominating conceptual framework in cancer research has been the combination of gene mutation theory and clonal evolutionary viewpoints. In this theory, it is thought that normal cells become cancerous through a gradual accumulation of genetic changes [20, 37, 39, 41, 42, 91, 105]. These cancer cells are then able to expand and metastasize. The basis for this theory was further concentered with the discovery of the Philadelphia chromosome, the reciprocal translocation between chromosomes 9 and 22 that is commonly associated with Chronic Myelogenous Leukemia (CML), which gives rise to the fusion gene BCR-ABL. This gene is used as a biomarker for detection and treatment for CML and in fact serves as an archetype for the entire field to follow [45, 50]. This discovery has led to an influx of research concentrating on identifying other fusion genes, and there are many success stories regarding leukemias and lymphomas but much less for most solid tumors. The list of oncogenes and tumor suppressor genes continues to grow after decades of effort; however, their value as biomarkers and therapeutic targets in a clinical setting has not been as fruitful as expected due to the high sample heterogeneity observed compared to more homogenous experimental models. Now, scientists realize their prediction that finding biomarkers would essentially lead to the curing of all cancers was wishful thinking, as the genetic landscape of most cancers is highly dynamic, especially under therapeutic intervention. For example, there are many gene mutations, copy number variations and genome level alterations in addition to cancer-specific genes [106]. Not surprisingly, whole genome scanning has identified large numbers of involved loci, but most of them only contribute to the overall genetics in a very moderate fashion. This is consistent with the large number of gene mutations detected from the cancer sequencing project. Similarly, whole genome expression studies have suddenly complicated the understanding of specific pathways, as expression profiles showed that so many unexpected genes are clearly involved in specific pathways for unknown reasons. Even for some well-characterized genes such as BRCA1 and BRCA2, there are dozens of other genes that influence their contributions to disease conditions, making clinical prediction based only on BRCA1 and BRCA2 gene status less precise. Furthermore, the fact that many "normal" healthy individuals display a large number of gene mutations including some very serious ones (such as DNA repair gene mutations) paradoxically adds another layer of complexity.

This situation is far beyond cancer research. For example, comparative genome analysis also has complicated the field of phylogenetics. In the past, due to the lack of whole genome sequencing, phylogenies of species were based on single genes or an accumulation of as many sequenced genes as possible. What resulted was mass confusion and disputes within the scientific communities as to which phylogenies were considered to be correct. So when whole genome sequencing became a feasible tool, there was an expectation that the question of the real phylogeny of life would be answered. However, there are still disputes as to the relationship of Eukaryotes, Bacteria, and Archaea, as well as which domain includes the organism that is considered to be the Last Unknown Common Ancestor (LUCA). Obviously, the lack of appreciation towards the karyotype (or genome topology) defined genomic history of evolution contributes to such confusion, as speciation studies need to be more focused on macro-evolution [2, 4, 85, 107]. Simply put, new discoveries associated with the post-sequencing era challenge many basic predictions about the gene theory of human disease, which has forced the scientific community to challenge the very foundation of current genetics and genomics.

3.2. Inconsistencies of Genetic Information at Multiple Levels

The challenge of data interpretation extends beyond the large amount of data available, as the data obtained from different levels of genetic organization (*e.g.* gene, epigene, and chromosomal levels) often conflict with each other. First, there is no simple linear relationship between the knowledge gained at lower and higher levels. For example, it is not the case that the accumulation of gene-based knowledge can automatically lead to an understanding of the genome, as there is a knowledge gap between "parts inheritance" and "system inheritance" [65]. The pattern of global gene expression can be drastically altered by simply introducing an additional chromosome 21, meaning the pattern of gene expression for a large number of genes has changed, most of which are not located on chromosome 21!

Perhaps the most important example of how there is a conflict between genes and chromosomes comes from the reinterpretation of the main function of sexual reproduction [107, 108]. For over a century, the dominating theory of the function of sex has been to provide necessary diversity for evolution, as the process of meiosis and stochastic sperm and egg genetic mixing will generate different combinations of genes. By comparing the punctuated and stepwise evolution using cancer models, we have realized that the major function of sex is to maintain the purity of the genome rather than increase genetic diversity. Despite that the secondary function of sex is to increase gene level diversity, the primary function is to reduce the diversity at the genome level. As insightfully pointed out by Wilkins and Holliday, "The conclusion is surprising: the initial function of chromosome pairing was to *limit*, not enhance, recombination" [109]. Our further synthesis suggests that the function of sex is conflicted by stability at the genome level and diversity at the gene level. The meiotic process plays an important role in separating germline and somatic cells, which allows the increased constraint and dynamics for evolution [4].

A similar conflict also can be observed between individual cell and cellular population, another collaborative yet conflicting pair. The average rule is assumed to happen in the biological world, so current statistical analysis is thought to be an essential predictor as to how systems work. One of the main issues with cancer population studies is the statistical findings, which revolve mainly around the average. Extensively using mixed cells for DNA/RNA/protein profiles (such as Western blots) represents such an example. However, averages are unable to account for the heterogeneity of cancer cells. This is due to the fact that average-based studies do not incorporate outliers (exceptions to the rule), which is very important for cancer evolution, especially after drug treatment. Our recent study has illustrated that in unstable cancer cell populations, there is no 'average profile,' and the outliers often determine the direction of cancer evolution. In this sense, the application of biostatistics could be misleading. This idea can be extended to other heterogeneous systems involving gene mutations, transcription regulation, pathway switching and cell death, especially under high levels of stress [90, 110, 111]. This is why single cell profiling has become increasingly important [112, 113]. Furthermore, as mentioned, there is decreased predictability between short-term profiling and longer term outcomes when somatic cellular evolution is involved, and especially when the macro-evolution is dominant [4, 110, 111].

In summary, we now have the capability to collect massive amounts of data using cutting edge -omics technologies, but the rationale behind obtaining data at different levels of genetic organization and time windows is less clear. It has become challenging to interpret and synthesize these data, let alone to apply this new information towards the development of useful clinical understandings. For all these diverse issues, the common feature is that these seemingly convincing, elegant and feel-good assumptions do not reflect the heterogeneous nature of biological systems, especially when dealing with disease conditions where the law differs from normal physiological conditions due to evolutionary selection under high levels of stress [2-4]. This leads to the paradoxical situation that many individual papers make sense of the observations and provide an explanation; however, when many are put together, the stories do not make sense due to obvious conflict. Pure basic research is good based on the selected model and data presentation, but this knowledge has major limitations when it comes to explaining the reality of diseases.

3.3. Important Role of Systems Biology and its own Limitations

At first, the pairing of systems biology with bioinformatics seemed to be a logical solution. If finding dominant gene mutation patterns was not possible, what about finding dominant patterns of pathways or networks? It was hypothesized that this approach would reveal the convergence of a few important pathways that could be handled in the clinic. If there were so many genes that could be potentially involved for a given trait, and it is challenging to pick the winner, why not use bioinformatics tools to identify the pattern from the noise? If there are so many data points, why not use quantitative meanings to wash out the insignificance and just report the important ones? If the identification of key driver

gene mutations are complicated by large numbers of passenger gene mutations, why not compare even more samples to reduce the noise? Many system biologists, with the help of powerful computational software packages, could test a large number of hypotheses and let the data speak for themselves. Different types of filters were designed to dissect the same dataset, and the data presentation became very colorful. It was thought that bias would be reduced by using these computational tools.

If all biological problems could be solved this way, our life in the post-genomic era would be much easier than it is. However, reality is always more complicated than computational models can handle. In fact, bioinformatics' analyses depend on the correct biological concept, specific hypothesis, technical assumptions, and different filters that need to be precisely reflected on the system behavior (including dynamics, heterogeneity, and different conditional states). In a sense, most computational analyses are highly biased based on pre-assumptions and the rationale of specific filters. One of the biggest challenges is using mathematics to simplify highly heterogeneous biological systems. For example, there are many famous assumptions that are clearly not reflected in biology. Genome heterogeneity is a perfect explanation as to why it is becoming difficult to validate published research. According to *Science Magazine*, two-thirds of published psychology research in highly respected journals cannot be replicated. Cancer biology studies have also been under scrutiny because only 6 out of 53 high-profile papers can be reproduced [114]. A large amount of cancer studies have looked into mouse models to understand how particular pathways play a role in cancer. Due to the homogeneity of inbred mouse strains, it is difficult to translate what is found in an experimental system to clinics because of the increase in heterogeneity.

Obviously, current experimental systems and some key practices of data collection are responsible for such gaps, as without facing biological realities (such as the time issue and heterogeneity-mediated complexity) [4, 115], scientists in fact do not yet have a good understanding of how many different variables are involved in network interaction, even for some well-established pathways, let alone how to deal with genome scale holistic interaction. Knowing that most individual cells can function differently, it can quickly reduce the initial trust we have placed in systems biology. The following examples further illustrate the challenges we are facing: In regards to building a gene function network, there is an assumption that all genes are equal; when searching for patterns, the overwhelming variation is often assumed as insignificant "noise". When studying the function of a specific gene, it is assumed that the state of a system without such a gene by using knock out models reflects the original function of the gene. Upon completion of the Human Genome Project, gene expression profiling by way of whole genome microarrays became a popular tool. It was thought that with gene expression profiles, defects or patterns found in the cellular pathways result in diseases such as cancer. However, it has been found that due to the heterogeneity of the cancerous tumors, there is a significant difference in gene expression between individual samples. Gene expression profiles are also used to identify a group of similarly expressed genes that can be used for prognosis. But, these gene

signatures are highly heterogeneous. When comparing four high-grade serous ovarian cancer (HGS) studies, overlap between two studies is no greater than 34%, overlap between three studies is no greater than 8%, and there is no relationship found between all four studies [116]. This indicates that gene expression profiles of gene signatures are patient-specific or progression-specific. Knowing that the genetic landscapes of cancer cells are highly dynamic, and the average profile is not very useful for predicting the emergence of outliers, detecting the historical genetic scar might have limited usage to predict the longer term evolution.

4. FUTURE DIRECTION: SEARCH FOR A NEW CONCEPTUAL FRAMEWORK AND HOLISTIC METHODOLOGIES

Considering the limitations and challenges of postgenomic research we have discussed, we must decide which new directions to take and determine what, if anything, we are overlooking. With the increased lack of specificity that we see with larger sample sizes and application of higher resolution technologies, it is becoming clearer that the reductionist's approach does not apply to systems biology. The key problem is the lack of a conceptual framework and proper methodologies. Due to the scope of this article, we will address these issues by mainly focusing with a molecular cytogenetics and cytogenomics perspective. Furthermore, we must take a stance as to how much data are enough, as the more data researchers accumulate, the cloudier the picture becomes.

4.1. Some Considerations for Systems Biology

The current paradigm of systems biology is based on four principal components or steps: 1. Define and collect the list of biological parts that participate in a cellular process; 2. Study the interactions between these parts, reconstruct the genetic circuits/network; 3. Use a mathematical format to describe the reconstructed network, and generate computer models to analyze, interpret and predict the biological functions based on the reconstructed network; and 4. Propose specific hypotheses based on the model prediction and test them experimentally to improve the model [10]. Based on these key steps, the influence of the reductionist is obvious, even though systems biology is supposed to work under more holistic principles. Despite some degree of certainty, such as genetic information that can be understood from gene to protein, and that the protein interaction network can be established based on chemical laws and experimentation, some key limitations are difficult to overcome. These include: most genotypes cannot be dissected into genes from the holistic genome (only a small portion of genes display high penetrance when associated with phenotypes); the information of genomic topology/heterogeneity/fuzziness is lacking for current efforts of reconstructing the network; and the evolutionary process that is based on genome level selection (beyond the selection of large-scale modules) is ignored.

The phenotype of a biological system is highly dynamic due to the following facts: 1). Any individual molecule can have many functions; 2). The genetic information carried by genetic identity is often less precise or fuzzy [4]; and 3). The function of genetic circuits depends on the context of the entire organism [10]. Because of this, it is difficult to predict the relationship between systems phenotype and cellular

phenotype, even though it is easier to explain a specific phenotype using network information (a specific systems phenotype). The key is to apply evolutionary analyses to study the system behavior rather than individual pathways or even systems phenotype. Some recent developments are indeed encouraging. By combining metabolic networks and evolution, more exciting observations have been made. One example is that there are many potential systems phenotypes which can be linked to the same cellular phenotype. Also, cellular phenotype dynamics are linked to the cellular evolutionary process. Following studies of the various levels of systems, it was proposed that mechanistic multi-scale models can be used to illustrate the relationship among various key players in systems biology. These are, namely: genes, genetic circuits, systems phenotype, cellular level phenotype, and organismal level phenotype.

Since the system phenotype can be highly dynamic (as it is linked to a huge number of pathways and similar cellular phenotypes), this leads to a more profound question: are profiles lower than the cellular level really as useful as expected? Interestingly, the search for such frameworks was unexpectedly initiated from cytogenetic studies. Specifically, the overwhelming karyotype changes observed in cancer have cried out for a departure from the gene-centric approach to systems biology.

4.2. Genome Mediated Macro-somatic Cell Evolution

A shift to a holistic framework requires accurate definition of the system, and this can be achieved with a deep understanding of how genomics and evolution work. The two phases of cancer evolution were originally discovered based on karyotypic changes from an immortalization model where a pattern of clonal and non-clonal expansions were detected. Such a pattern has recently been confirmed using cancer genome sequencing including single cell genome sequencing [39, 117-120]. Cancer evolution represents a series of genome-mediated system replacements consisting of NCCAs and CCAs occurring within two evolutionary phases. During the stepwise phase, the majority of cells are clonal across generations, and karyotypic diversification is traceable. In contrast, the punctuated phase is defined by a high frequency of NCCAs and massive/rapid genome reorganization (genome chaos), which break multiple system constraints (*e.g.* genome integrity, tissue architecture). Cancer progression thus consists of both macro-cellular (genome system replacement) and micro-cellular (genome system modification) evolution [4].

Evolution involves the contributions of multiple genetic levels (genome, gene, epigene), however, their influences vary sharply. Gene-level change, for example, modifies an existing system. Recent work by our group and others has provided ample support that the karyotype defines and governs the genetic system, representing the selection unit in evolution [20, 35]. Karyotypic change has been linked with transcriptome dynamics in cancer and yeast studies [87-89], supporting that genome topology change rapidly creates new systems. Together, the genome or karyotype defined "system inheritance" and gene/epigene defined "parts inheritance" provide the genetic basis for macro- and micro-cellular evolution.

Genetic systems under the context of evolutionary selection in real disease conditions are also drastically different from linear experimental conditions and models. Stepwise processes can be easily observed under strict, experimental conditions with conditional, more homogenous models, allowing for the easy and obvious identification of key contributors (e.g. genes, pathways) by researchers. However, these understandings rarely translate effectively to the clinic due to the much higher degrees of heterogeneity observed in patients coupled with more complex evolutionary dynamics. Moreover, our recent research has demonstrated that within a heterogeneous population, it is the outlier (not the sought-after average signal from samples/models) that directs overall cell population growth and drives disease progression and evolution [25].

Clearly, the new concepts of genomics and evolution must be integrated into systems biology, as the four major steps of current systems biology are based on gene theory, chemical law and evolutionary selection. If the genome theory is a departure from gene theory, and the two phases of cellular evolution replace stepwise evolution, the framework of systems biology should also be modified accordingly.

4.3. Heterogeneity is the Layer of Complexity Representing Evolutionary Potential

Taking this new framework into account changes previous perceptions of heterogeneity. Rather than simply “noise” that must be filtered out or disregarded, heterogeneity provides potential for evolution and represents a new layer of complexity [4, 97, 98]. In terms of complex tissues and organs, we have previously discussed the role of heterogeneity in performing complex functions while providing the robustness necessary to withstand the stress associated with those functions [90]. In a tumor cell population, heterogeneity provides cancer the opportunity to survive and thrive against various constraints (e.g. tissue, immune system) and environmental insults (e.g. chemotherapeutics). Due to the large presence and importance of heterogeneity, we must reconsider the development and application of artificial computational filters that act to eliminate noise, as by doing so, we generate artificial biological understandings. It should be pointed out that increased genomic heterogeneity can be detected from many normal tissue types [121, 122], and the de novo gene mutation represents a common phenomenon [123]. The high degree of multiple levels of genetic heterogeneity observed from various normal tissues strongly support the importance of fuzzy inheritance in somatic cells and the importance of heterogeneity in somatic cell adaptation [4]. Knowing this, one priority for systems biology as well as post-genomic studies is to seriously consider the issue of heterogeneity when constructing networks.

4.4. The Reductionist Approach is Limited for Understanding Evolution

There are different types of biological understandings. The molecular characterizations of a specific gene mutation or pathway, or a genetic circuit/network, represent a simpler task. However, for a more general evolutionary understanding, it is essential to integrate heterogeneity, the entire process, and outlier response (exceptions) under stress. Since single profile “snapshots” of samples do not provide

accurate understandings of the overall disease process, these do not aptly illustrate dynamic processes. Cancer, for example, represents an evolutionary process at multiple genetic levels (karyotype, gene, epigene), which cannot simply be explained or understood with end product profiling, nor correctly profiled by fixed markers. Further, the genome is not simply a string or bag of gene and non-coding sequence, as the genome topology plays a key role in defining and governing the genetic system. System emergence by genome reorganization (e.g. through genome chaos) is a common phenomenon in cancer, and as a result, gene roles may differ and even conflict from one genetic system (tumor cell) to the next. Thus, the reductionist’s approach does not apply to systems biology of cancer, as a role of one gene or network is not an accurate representation of the entire tumor cell population. Matters are also too complicated for the reductionist’s approach in most common complex diseases, as multi-level heterogeneity is also observed.

Without a crucial genome and evolutionary framework, there are many misconceptions. For example, somatic cellular evolution is assumed to be a stepwise progression where small changes can accumulate to become big changes; biomolecules can freely interact without genomic topological constraint as the affinity among molecules is the most deterministic factor; each gene’s information is precise, and the gap between genotype and phenotype is caused by environments (but the mechanism is explainable only by hand waving); genetic profiles of the germline should predict the phenotype despite that the somatic cell undergoes evolutionary selection for years and genomic variations are essential for cellular adaptation; identifying cancer pathways is essential even though when targeted, these pathways often are drastically altered; the higher the resolution, the better the understanding (without seriously testing this hypothesis in the context of disease research); the more data, the better (not realizing that when the concept is not correct, more data collection will reduce the value of research). The list goes on (more examples can be found in Heng 2015 [4]).

Interestingly, some recent systems biology analyses have confirmed several key conclusions from our “watching karyotypic evolution in action” experiments. For example, based on the fact that the same phenotype (i.e. cellular immortalization, drug resistance) can be achieved by different karyotypes (which unify different genetic profiles), we have promoted the evolutionary mechanism of cancer. Such a synthesis is supported by the conclusion that the same cellular phenotype can be produced by different system phenotypes (or system interaction patterns). A more exciting observation is that the pattern of switching between growth and survival is reflected by the relationship of CCAs (growth) and NCCAs (survival). This has also been confirmed by profiling the growth of *E.coli*, on the basis of metabolic network switching and evolutionary trade-off [124, 125]. In these experiments, it is clear that the heterogeneity of growth/survival represents an essential balancing act for *E.coli*, which is the key both for the long-term existence of the species and short-term adaptation at the individual level.

4.5. The Unique Role Molecular Cytogenetics will Play in Postgenomic Research

We call for putting systems-integrated cytogenetics into the driver’s seat to understand and tackle complex genetic

diseases including cancer [12]. This strategy will promote the correct usage of the wealth of information researchers have uncovered and reduce cherry picking of the data based on some specific limited hypotheses. Since systems biology is not only about parts characterization but characterizing the heterogeneous process, the “noise” should be included as the key feature of the investigated systems. This shift from reductionism to an evolutionary systematic approach as the priority of future research will be challenging; however, this approach will offer new solutions through holistic understandings.

First and foremost, the goal is to integrate karyotype (chromosome set)-coded system inheritance into systems biology. The illustration of how karyotype coding works will bring about a new way of thinking in systems biology. Clearly, genome chaos will change the karyotype coding, resulting in a new type of genetic network. More integrated molecular cytogenetic and genetic network studies are needed, in particular, by applying the concept of the emergent properties among agents (different genes) within the framework of an altered topological platform (re-organized locations within the nucleus). For example, we need to compare network rewiring mechanisms in a normal developmental process with those affected by genome chaos.

Following evolution at the genome (karyotype) level will also unify individual molecular mechanisms (holistic approach), identify phases of instability/stability, and provide a single-cell resolution understanding of population dynamics. Monitoring the phases of stability can potentially predict systems behavior, which is useful both for diagnosis and watching the trend of disease progressions and responses to treatments.

This emerging framework has shown promise in both cancer and Gulf War Illness (GWI) studies, confirming how genome instability serves as a common mechanism for many common and complex diseases. Increased genome-level heterogeneity has been observed at key transitional events in cancer [39, 126]. We recently examined the genome instability of GWI patients by comparing the degree of stochastic chromosomal aberrations from short-term lymphocyte cultures. Various karyotypic abnormalities were observed, and we demonstrated that genome instability was significantly elevated in GWI patients compared to controls [37, 97, 127]. Similar analysis was also performed in a Chronic Fatigue Syndrome (CFS) study, where again the overall level of genome instability was significantly higher in CFS patients than controls [37], (Ye *et al.*, manuscript in preparation). A general model has been recently proposed that links stress-induced genomic adaptation and increased disease potential as an evolutionary trade-off [90, 97, 128].

As we discussed in a recent perspective article, to achieve the goal of establishing systems-integrated cytogenetics, new cytogenetic methodologies are urgently needed [12]. These methods include: 1) technologies to integrate multiple levels of genetic alterations, and prioritize their importance in the context of stages of disease progression or evolutionary process. In particular, current -omics studies must include karyotype information, as when the “system inheritance” is altered, the meaning of “parts inheritance” often will be changed. The theoretical basis of this concept

has been illustrated by multiple level landscape models [35, 36]; 2) methodologies to measure the multiple types of heterogeneity including the complexity of altered karyotypes. Recently, we have reported many new types of chromosomal aberrations [37, 99], many of which represent the new mechanism of generating fuzzy inheritance at the genome level [4]. There is a strong link between karyotype complexity, tumorigenesis, drug resistance, and clinical prediction (including stem cell/degenerative disease research, reproductive medicine, aging, cancer, and other common and complex illnesses), as well as cellular adaptation [71, 129-137]. Unique cytogenetic/cytogenomic research needs to be combined with, not replaced by, Next Generation Sequencing (NGS) [99, 138]. Of equal importance, karyotype studies are important for organismal evolutionary studies as well, as the capability of passing the same or a very similar karyotype guarantees the identity of a species by controlling system inheritance [4, 85, 107, 108, 139]. Various chromosomal level methods must be developed to directly study the genome [38, 55, 140-142]. Hi-C technology, which comprehensively detects chromatin interactions in the nucleus [143], requires integration of karyotype level information; 3) quantitative methods to integrate measuring combined heterogeneity and complexity [144]; and 4) methods to measure multiple levels of fuzzy inheritance (from epigene and gene to genome). Fuzzy inheritance functions as the internal mechanism of heterogeneity. There is increased discussion regarding the relationship between stress-induced heterogeneity and the evolutionary potential for diseases [90, 91, 145, 146]. Since epigene/gene and genome alteration mainly contribute to the different phases of cellular evolution, characterizing/measuring fuzzy inheritance at different levels should play an important role in understanding the evolution of disease [4] as well as organismal evolution [147].

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No Animals/Humans were used for studies that are base of this research.

CONSENT FOR PUBLICATION

Not applicable.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

This article is part of a series of studies entitled ‘The mechanisms of somatic cell and organismal evolution’. This work was supported by the WSU Office of the Vice President for Research Bridge Funding Grant to H. H. Heng. Due to the focused scope of this review, we apologize to others in this diverse research area whose contributions should be also acknowledged.

REFERENCES

- [1] Venter, J.C.; Adams, M.D.; Myers, E.W.; Li, P.W.; Mural, R.J.; Sutton, G.G.; Smith, H.O.; Yandell, M.; Evans, C.A.; Holt, R.A.; Go-

- cayne, J.D.; Amanatides, P.; Ballew, R.M.; Huseon, D.H.; Wortman, J.R.; Zhang, Q.; Kodira, C.D.; Zheng, X.H.; Chen, L.; Skupski, M.; Subramanian, G.; Thomas, P.D.; Zhang, J.; Miklos, G.L.; Nelson, C.; Broder, S.; Clark, A.G.; Nadeau, J.; McKusick, V.A.; Zinder, N.; Levine, A.J.; Roberts, R.J.; Simon, M.; Slayman, C.; Hunkapiller, M.; Bolanos, R.; Delcher, A.; Dew, I.; Fasulo, D.; Flanigan, M.; Florea, L.; Halpern, A.; Hannenhalli, S.; Kravitz, S.; Levy, S.; Mobarry, C.; Reinert, K.; Remington, K.; Abu-Threideh, J.; Beasley, E.; Biddick, K.; Bonazzi, V.; Brandon, R.; Cargill, M.; Chandramouliswaran, I.; Charlab, R.; Chaturvedi, K.; Deng, Z.; Di Francesco, V.; Dunn, P.; Eilbeck, K.; Evangelista, C.; Gabrielian, A.E.; Gan, W.; Ge, W.; Gong, F.; Gu, Z.; Guan, P.; Heiman, T.J.; Higgins, M.E.; Ji, R.R.; Ke, Z.; Ketchum, K.A.; Lai, Z.; Lei, Y.; Li, Z.; Li, J.; Liang, Y.; Lin, X.; Lu, F.; Merkulov, G.V.; Milshina, N.; Moore, H.M.; Naik, A.K.; Narayan, V.A.; Neelam, B.; Nusskern, D.; Rusch, D.B.; Salzberg, S.; Shao, W.; Shue, B.; Sun, J.; Wang, Z.; Wang, A.; Wang, X.; Wang, J.; Wei, M.; Wides, R.; Xiao, C.; Yan, C.; Yao, A.; Ye, J.; Zhan, M.; Zhang, W.; Zhang, H.; Zhao, Q.; Zheng, L.; Zhong, F.; Zhong, W.; Zhu, S.; Zhao, S.; Gilbert, D.; Baumhueter, S.; Spier, G.; Carter, C.; Chavchik, A.; Woodage, T.; Ali, F.; An, H.; Awe, A.; Baldwin, D.; Baden, H.; Barnstead, M.; Barrow, I.; Beeson, K.; Busam, D.; Carver, A.; Center, A.; Cheng, M.L.; Curry, L.; Danaher, S.; Davenport, L.; Desilets, R.; Dietz, S.; Dodson, K.; Doup, L.; Ferreira, S.; Garg, N.; Gluecksmann, A.; Hart, B.; Haynes, J.; Haynes, C.; Heiner, C.; Hladun, S.; Hostin, D.; Houck, J.; Howland, T.; Ibegwam, C.; Johnson, J.; Kalush, F.; Kline, L.; Koduru, S.; Love, A.; Mann, F.; May, D.; McCawley, S.; McIntosh, T.; McMullen, I.; Moy, M.; Moy, L.; Murphy, B.; Nelson, K.; Pfannkoch, C.; Pratts, E.; Puri, V.; Qureshi, H.; Reardon, M.; Rodriguez, R.; Rogers, Y.H.; Romblad, D.; Ruhfel, B.; Scott, R.; Sitter, C.; Smallwood, M.; Stewart, E.; Strong, R.; Suh, E.; Thomas, R.; Tint, N.N.; Tse, S.; Vech, C.; Wang, G.; Wetter, J.; Williams, S.; Williams, M.; Windsor, S.; Winn-Deen, E.; Wolf, K.; Zaveri, J.; Zaveri, K.; Abril, J.F.; Guigó, R.; Campbell, M.J.; Sjolander, K.V.; Karlak, B.; Kejariwal, A.; Mi, H.; Lazareva, B.; Hatton, T.; Narechania, A.; Diemer, K.; Muruganujan, A.; Guo, N.; Sato, S.; Bafna, V.; Istrail, S.; Lippert, R.; Schwartz, R.; Walenz, B.; Yoosheph, S.; Allen, D.; Basu, A.; Baxendale, J.; Blick, L.; Caminha, M.; Carnes-Stine, J.; Caulk, P.; Chiang, Y.H.; Coyne, M.; Dahlke, C.; Mays, A.; Dombroski, M.; Donnelly, M.; Ely, D.; Esparham, S.; Fosler, C.; Gire, H.; Glanowski, S.; Glasser, K.; Glodek, A.; Gorokhov, M.; Graham, K.; Gropman, B.; Harris, M.; Heil, J.; Henderson, S.; Hoover, J.; Jennings, D.; Jordan, C.; Jordan, J.; Kasha, J.; Kagan, L.; Kraft, C.; Levitsky, A.; Lewis, M.; Liu, X.; Lopez, J.; Ma, D.; Majoros, W.; McDaniel, J.; Murphy, S.; Newman, M.; Nguyen, T.; Nguyen, N.; Nodell, M.; Pan, S.; Peck, J.; Peterson, M.; Rowe, W.; Sanders, R.; Scott, J.; Simpson, M.; Smith, T.; Sprague, A.; Stockwell, T.; Turner, R.; Venter, E.; Wang, M.; Wen, M.; Wu, D.; Wu, M.; Xia, A.; Zandieh, A.; Zhu, X. The sequence of the human genome. *Science* **2001**, *291*(5507), 1304-1351. Available from: clinchem.aaccjnl.org/content/61/9/1207
- [2] Heng, H.H. The genome-centric concept: Resynthesis of evolutionary theory. *Bioessays*, **2009**, *31*(5), 512-525.
- [3] Heng, H.H.; Liu, G.; Stevens, J.B.; Bremer, S.W.; Ye, K.J.; Abdallah, B.Y.; Horne, S.D.; Ye, C.J. Decoding the genome beyond sequencing: The new phase of genomic research. *Genomics*, **2011**, *98*(4), 242-252.
- [4] Heng, H.H. *Debating Cancer: The Paradox in Cancer Research*, World Scientific Publishing: Singapore/New Jersey, **2015**.
- [5] Ideker, T.; Galitski, T.; Hood, L. A new approach to decoding life: Systems biology. *Annu. Rev. Genomics Hum. Genet.*, **2001**, *2*, 343-372. Available from: <http://www.annualreviews.org/doi/abs/10.1146/annurev.genom.2.1.343>
- [6] Kitano, H. Systems biology: A brief overview. *Science*, **2002**, *295*(5560), 1662-1664. Available from: science.sciencemag.org/content/295/5560/1662.full
- [7] Butcher, E.C.; Berg, E.L.; Kunkel, E.J. Systems biology in drug discovery. *Nat. Biotechnol.*, **2004**, *22*(10), 1253-1259.
- [8] Abu-Asab, M.S.; Abu-Asab, N.; Loffredo, C.A.; Clarke, R.; Amri, H. Identifying early events of gene expression in breast cancer with systems biology phylogenetics. *Cytogenet. Genome Res.*, **2013**, *139*(3), 206-214.
- [9] Werner, H.M.; Mills, G.B.; Ram, P.T. Cancer Systems Biology: A peek into the future of patient care? *Nat. Rev. Clin. Oncol.*, **2014**, *11*(3), 167-176.
- [10] Palsson, B.O. *Systems Biology: Constraint-based Reconstruction and Analysis*, Cambridge University Press: Cambridge, **2015**.
- [11] Pecina-Slaus, N.; Pecina, M. Only one health, and so many omics. *Cancer Cell Int.*, **2015**, *15*, 64. Available from: <https://cancer.cbiomedcentral.com/articles/10.1186/s12935-015-0212-2>
- [12] Heng, H.H.; Regan, S. A systems biology perspective on molecular cytogenetics. *Curr. Bioinform.*, **2016**, *12*(1), 4-10.
- [13] Pearson, H. What is gene? *Nature*, **2006**, *441*(7029), 399-401. Available from: <https://www.nature.com/articles/441398a>
- [14] Fox, K.E.; Harel, D. Beyond the gene. *PLoS One*, **2007**, *2*(11), e1231. Available from: journals.plos.org/plosone/article?id=10.1371/journal.pone.0001231
- [15] Wilkins, A.S. For the biotechnology industry, the penny drops (at last): Genes are not autonomous agents but function within networks! *Bioessays*, **2007**, *29*(12), 1179-1181.
- [16] Heng, H.H. The gene-centric concept: A new liability? *Bioessays*, **2008**, *30*, 196-197.
- [17] Check, Hayden. E. Human genome at ten: Life is complicated. *Nature*, **2010**, *464*(7289), 664-667. Available from: www.genomicaybioeconomia.org/images/6-2.pdf
- [18] Chakravarti, A. Genomics is not enough. *Science*, **2010**, *334*(6052), 15. Available from: science.sciencemag.org/content/334/6052/15
- [19] Ball, P. DNA: Celebrate the unknowns. *Nature*, **2013**, *496*(7446), 419-420. Available from: www.nature.com/articles/496419a
- [20] Heng, H.H. Cancer genome sequencing: The challenges ahead. *Bioessays*, **2007**, *29*(8), 783-794.
- [21] Lander, E.S. Initial impact of the sequencing of the human genome. *Nature*, **2011**, *470*(7333), 187-197. Available from: <https://www.nature.com/articles/nature09792>
- [22] Noble, D. Physiology is rocking the foundations of evolutionary biology. *Exp. Physiol.*, **2013**, *98*(8), 1235-1243.
- [23] Joyner, M.J.; Prendergast, F.G. Chasing Mendel: Five questions for personalized medicine. *J. Physiol.*, **2014**, *592*(11), 2381-2388.
- [24] Noble, D. Evolution beyond neo-Darwinism: A new conceptual framework. *J. Exp. Biol.*, **2015**, *218*, 1273. Available from: jeb.biologists.org/content/218/1/7
- [25] Abdallah, B.Y.; Horne, S.D.; Stevens, J.B.; Liu, G.; Ying, A.Y.; Vanderhyden, B.; Krawetz, S.A.; Gorelick, R.; Heng, H.H. Single cell heterogeneity: Why unstable genomes are incompatible with average profiles. *Cell Cycle*, **2013**, *12*(23), 3640-3649.
- [26] Weinberg, R.A. Coming full circle from endless complexity to simplicity and back again. *Cell*, **2014**, *157*(1), 267-271.
- [27] Rancati, G.; Pavelka, N.; Fleharty, B.; Noll, A.; Trimble, R.; Walton, K.; Perera, A.; Staehling-Hampton, K.; Seidel, C.W.; Li, R. Aneuploidy underlies rapid adaptive evolution of yeast cells deprived of a conserved cytokinesis motor. *Cell*, **2008**, *135*, 879-893.
- [28] Duesberg, P.; Iacobuzio-Donahue, C.; Brosnan, J.A.; McCormack, A.; Mandrioli, D.; Chen, L. Origin of metastases: Subspecies of cancers generated by intrinsic karyotypic variations. *Cell Cycle*, **2012**, *11*, 1151-1166. Available from: www.tandfonline.com/doi/abs/10.4161/cc.11.6.19580
- [29] Lanctot, C.; Cheutin, T.; Cremer, M.; Cavalli, G.; Cremer, T. Dynamic genome architecture in the nuclear space: Regulation of gene expression in three dimensions. *Nat. Rev. Genet.*, **2007**, *8*(2), 104-115.
- [30] Martincorena, I.; Roshan, A.; Gerstung, M.; Ellis, P.; Van Loo, P.; McLaren, S.; Wedge, D.C.; Fullam, A.; Alexandrov, L.B.; Tubio, J.M.; Stebbings, L.; Menzies, A.; Widaa, S.; Stratton, M.R.; Jones, P.H.; Campbell, P.J. Tumor evolution. High burden and pervasive positive selection of somatic mutations in normal human skin. *Science*, **2015**, *348*(6237), 880-886. Available from: <http://science.sciencemag.org/content/348/6237/880>
- [31] Elsasser, W.M. Outline of a theory of cellular heterogeneity. *Proc. Natl. Acad. Sci. U.S.A.*, **1984**, *81*(16), 5126-5129.
- [32] Heppner, G.H. Tumor heterogeneity. *Cancer Res*, **1984**, *44*(6), 2259-2265.
- [33] Rubin, H. Early origin and pervasiveness of cellular heterogeneity in some malignant transformations. *Proc. Natl. Acad. Sci. U.S.A.*, **1984**, *81*(16), 5121-5125.
- [34] Heng, H.H.; Bremer, S.W.; Stevens, J.B.; Ye, K.J.; Liu, G.; Ye, C.J. Genetic and epigenetic heterogeneity in cancer: A genome-centric perspective. *J. Cell Physiol.*, **2009**, *220*(3), 538-547.
- [35] Heng, H.H.; Stevens, J.B.; Bremer, S.W.; Liu, G.; Abdallah, B.Y.; Ye, C.J. Evolutionary mechanisms and diversity in cancer. *Adv.*

- Cancer Res.*, **2011**, *112*, 217-253. Available from: <http://linkinghub.elsevier.com/retrieve/pii/B9780123876881000089>
- [36] Huang, S. Genetic and non-genetic instability in tumor progression: Link between the fitness landscape and the epigenetic landscape of cancer cells. *Cancer Metastasis Rev.*, **2013**, *32*(3-4), 423-448.
- [37] Heng, H.H.; Liu, G.; Stevens, J.B.; Abdallah, B.Y.; Horne, S.D.; Ye, K.J.; Bremer, S.W.; Chowdhury, S.K.; Ye, C.J. Karyotype heterogeneity and unclassified chromosomal abnormalities. *Cytogenet. Genome Res.*, **2013**, *139*(3), 144-157.
- [38] Heng, H.H.; Stevens, J.B.; Liu, G.; Bremer, S.W.; Ye, C.J. Imaging genome abnormalities in cancer research. *Cell Chromos.*, **2004**, *3*(1), 1. Available from: <https://cellandchromosome.biomedcentral.com/articles/10.1186/1475-9268-3-1>
- [39] Heng, H.H.; Stevens, J.B.; Liu, G.; Bremer, S.W.; Ye, K.J.; Reddy, P.V.; Wu, G.S.; Wang, Y.A.; Tainsky, M.A.; Ye, C.J. Stochastic cancer progression driven by non-clonal chromosome aberrations. *J. Cell Physiol.*, **2006**, *208*(2), 461-472.
- [40] Horne, S.D.; Wexler, M.; Stevens, J.B.; Heng, H.H. Insights on processes of evolutionary tumor growth. *Atlas Genet. Cytogenet. Oncol. Haematol.*, **2016**, *4*. Available from: <http://atlasgeneticsoncology.org/Deep/EvolTumorGrowthID20142.html>
- [41] Heng, H.H.; Bremer, S.W.; Stevens, J.; Ye, K.J.; Miller, F.; Liu, G.; Ye, C.J. Cancer progression by non-clonal chromosome aberrations. *J. Cell Biochem.*, **2006**, *98*(6), 1424-1435.
- [42] Heng, H.H.; Liu, G.; Bremer, S.; Ye, K.J.; Stevens, J.; Ye, C.J. Clonal and non-clonal chromosome aberrations and genome variation and aberration. *Genome*, **2006**, *49*(3), 195-204.
- [43] Horne, S.D.; Ye, C.J.; Abdallah, B.Y.; Liu, G.; Heng, H.H. Cancer genome evolution. *Translation. Can. Res.*, **2015**, *4*(3), 303-313.
- [44] Iourov, I.Y.; Vorsanova, S.G.; Yurov, Y.B. *In silico* molecular cytogenetics: A bioinformatic approach to prioritization of candidate genes and copy number variations for basic and clinical genome research. *Mol. Cytogenet.*, **2014**, *7*(1), 98. Available from: <https://molecularcytogenetics.biomedcentral.com/articles/10.1186/s13039-014-0098-z>
- [45] Rowley, J.D. The critical role of chromosome translocations in human leukemias. *Annu. Rev. Genet.*, **1998**, *32*, 495-519. Available from: www.annualreviews.org/doi/pdf/10.1146/annurev.genet.32.1.495
- [46] Worton, R.G.; Thompson, M.W. Genetics of Duchenne muscular dystrophy. *Annu. Rev. Genet.*, **1998**, *22*, 601-629. Available from: <http://www.annualreviews.org/doi/10.1146/annurev.ge.22.120188.003125>
- [47] Kunkel, L.M. 2004 William Allan Award address. Cloning of the DMD gene. *Am. J. Hum. Genet.*, **2005**, *76*(2), 205-214.
- [48] Consortium TEPKD. The polycystic kidney disease 1 gene encodes a 14 kb transcript and lies within a duplicated region on chromosome 16. The European polycystic kidney disease consortium. *Cell*, **1994**, *77*(6), 881-894.
- [49] Rabbitts, T.H. Chromosomal translocations in human cancer. *Nature*, **1994**, *372*(6502), 143-149. Available from: www.nature.com/articles/372143a0
- [50] Horne, S.D.; Stevens, J.B.; Abdallah, B.Y.; Liu, G.; Bremer, S.W.; Ye, C.J.; Heng, H.H. Why imatinib remains an exception of cancer research. *J. Cell Physiol.*, **2013**, *228*(4), 665-670.
- [51] Lichter, P.; Ledbetter, S.A.; Ledbetter, D.H.; Ward, D.C. Fluorescence *in situ* hybridization with Alu and L1 polymerase chain reaction probes for rapid characterization of human chromosomes in hybrid cell lines. *Proc. Natl. Acad. Sci. U.S.A.*, **1990**, *87*(17), 6634-6638.
- [52] Fan, Y.S.; Davis, L.M.; Shows, T.B. Mapping small DNA sequences by fluorescence *in situ* hybridization directly on banded metaphase chromosomes. *Proc. Natl. Acad. Sci. U.S.A.*, **1990**, *87*(16), 6223-6227.
- [53] Heng, H.H.; Tsui, L.C. Modes of DAPI banding and simultaneous *in situ* hybridization. *Chromosoma*, **1993**, *102*(5), 325-332.
- [54] Kunz, J.; Scherer, S.W.; Klawitz, I.; Soder, S.; Du, Y.Z.; Speich, N.; Kalf-Suske, M.; Heng, H.H.; Tsui, L.C.; Grzeschik, K.H. Regional localization of 725 human chromosome 7-specific yeast artificial chromosome clones. *Genomics*, **1994**, *22*(2), 439-448.
- [55] Heng, H.H.; Spyropoulos, B.; Moens, P.B. FISH technology in chromosome and genome research. *Bioessays*, **1997**, *19*(1), 75-84.
- [56] Heng, H.H.; Squire, J.; Tsui, L.C. High-resolution mapping of mammalian genes by *in situ* hybridization to free chromatin. *Proc. Natl. Acad. Sci. U.S.A.*, **1992**, *89*(20), 9509-9513.
- [57] Heng, H.H.; Tsui, L.C. High resolution free chromatin/DNA fiber fluorescent *in situ* hybridization. *J. Chromatogr. A*, **1998**, *806*, 219-229. Available from: europepmc.org/abstract/med/9639891
- [58] Heng, H.H.; Tsui, L.C.; Moens, P.B. Organization of heterologous DNA inserts on the mouse meiotic chromosome core. *Chromosoma*, **1994**, *103*(6), 401-407.
- [59] Heng, H.H.; Chamberlain, J.W.; Shi, X.M.; Spyropoulos, B.; Tsui, L.C.; Moens, P.B. Regulation of meiotic chromatin loop size by chromosomal position. *Proc. Natl. Acad. Sci. U.S.A.*, **1996**, *93*(7), 2795-2800.
- [60] Moens, P.B.; Heddle, J.A.; Spyropoulos, B.; Heng, H.H. Identical megabase transgenes on mouse chromosomes 3 and 4 do not promote ectopic pairing or synapsis at meiosis. *Genome*, **1997**, *40*(5), 770-773.
- [61] Haaf, T.; Ward, D.C. Structural analysis of alpha-satellite DNA and centromere proteins using extended chromatin and chromosomes. *Hum. Mol. Genet.*, **1994**, *3*(5), 697-709.
- [62] Heng, H.H.; Spyropoulos, B.; Moens, P.B. DNA-protein *in situ* covisualization for chromosome analysis. *Meth. Mol. Biol.*, **2000**, *123*, 15-27. Available from: <https://link.springer.com/protocol/10.1385%2F1-59259-677-0%3A15>
- [63] Goldman, A.S.; Hulten, M.A. Meiotic analysis by FISH of a human male 46,XY,t(15;20)(q11.2;q11.2) translocation heterozygote: Quadrivalent configuration, orientation and first meiotic segregation. *Chromosoma*, **1993**, *102*(2), 102-111.
- [64] Bayani, J.; Squire, J.A. Advances in the detection of chromosomal aberrations using spectral karyotyping. *Clin. Genet.*, **2001**, *59*(2), 65-73.
- [65] Heng, H.H. Bio-complexity: Challenging reductionism In: *Handbook on Systems and Complexity in Health*; Sturmberg, J.P., Martin, C.M., Eds.; Springer: New York; **2013**; pp. 193-208.
- [66] Redon, R.; Ishikawa, S.; Fitch, K.R.; Feuk, L.; Perry, G.H.; Andrews, T.D.; Fiegler, H.; Shaper, M.H.; Carson, A.R.; Chen, W.; Cho, E.K.; Dallaire, S.; Freeman, J.L.; Gonzalez, J.R.; Gratacos, M.; Huang, J.; Kalaitzopoulos, D.; Komura, D.; MacDonald, J.R.; Marshall, C.R.; Mei, R.; Montgomery, L.; Nishimura, K.; Okamura, K.; Shen, F.; Somerville, M.J.; Tchinda, J.; Valsesia, A.; Woodwark, C.; Yang, F.; Zhang, J.; Zerjal, T.; Armengol, L.; Conrad, D.F.; Estivill, X.; Tyler-Smith, C.; Carter, N.P.; Aburatani, H.; Lee, C.; Jones, K.W.; Scherer, S.W.; Hurler, M.E. Global variation in copy number in the human genome. *Nature*, **2006**, *444*, 444-454. Available from: <https://www.nature.com/articles/nature05329>
- [67] Feuk, L.; Carson, A.R.; Scherer, S.W. Structural variation in the human genome. *Nat. Rev. Genet.*, **2006**, *7*(2), 85-97.
- [68] Ishkanian, A.S.; Malloff, C.A.; Watson, S.K.; DeLeeuw, R.J.; Chi, B.; Coe, B.P.; Snijders, A.; Albertson, D.G.; Pinkel, D.; Marra, M.A.; Ling, V.; MacAulay, C.; Lam, W.L. A tiling resolution DNA microarray with complete coverage of the human genome. *Nat. Genet.*, **2004**, *36*(3), 299-303.
- [69] Kawamata, N.; Ogawa, S.; Zimmermann, M.; Kato, M.; Sanada, M.; Hemminki, K.; Yamamoto, G.; Nannya, Y.; Koehler, R.; Flohr, T.; Miller, C.W.; Harbort, J.; Ludwig, W.D.; Stanulla, M.; Schrappe, M.; Bartram, C.R.; Koefler, H.P. Molecular allelokaryotyping of pediatric acute lymphoblastic leukemias by high-resolution single nucleotide polymorphism oligonucleotide genomic microarray. *Blood*, **2008**, *111*(2), 776-784.
- [70] Eichler, E.E.; Flint, J.; Gibson, G.; Kong, A.; Leal, S.M.; Moore, J.H.; Nadeau, J.H. Missing heritability and strategies for finding the underlying causes of complex disease. *Nat. Rev. Genet.*, **2010**, *11*(6), 446-450.
- [71] Heng, H.H. Missing heritability and stochastic genome alterations. *Nat. Rev. Genet.*, **2010**, *11*(6), 813. Available from: <https://www.nature.com/articles/nrg2809-c3>
- [72] Cremer, T.; Cremer, C. Rise, fall and resurrection of chromosome territories: A historical perspective. Part II. Fall and resurrection of chromosome territories during the 1950s to 1980s. Part III. Chromosome territories and the functional nuclear architecture: experiments and models from the 1990s to the present. *Eur. J. Histochem.*, **2006**, *50*(4), 223-272.
- [73] Walter, J.; Joffe, B.; Bolzer, A.; Albiez, H.; Benedetti, P.A.; Muller, S.; Speicher, M.R.; Cremer, T.; Cremer, M.; Solovei, I. Towards many colors in FISH on 3D-preserved interphase nuclei. *Cytogenet. Genome Res.*, **2006**, *114*(3-4), 367-378.

- [74] Iborra, F.; Cook, P.R.; Jackson, D.A. Applying microscopy to the analysis of nuclear structure and function. *Methods*, **2003**, *29*(2), 131-141.
- [75] Zink, D.; Sadoni, N.; Stelzer, E. Visualizing chromatin and chromosomes in living cells. *Methods*, **2003**, *29*(1), 42-50.
- [76] Heng, H.H.; Krawetz, S.A.; Lu, W.; Bremer, S.; Liu, G.; Ye, C.J. Re-defining the chromatin loop domain. *Cytogenet. Cell Genet.*, **2001**, *93*(3-4), 155-161.
- [77] Bode, J.; Goetze, S.; Heng, H.; Krawetz, S.A.; Benham, C. From DNA structure to gene expression: Mediators of nuclear compartmentalization and dynamics. *Chromosome Res.*, **2003**, *11*(5), 435-445.
- [78] Heng, H.H.; Goetze, S.; Ye, C.J.; Liu, G.; Stevens, J.B.; Bremer, S.W.; Wykes, S.M.; Bode, J.; Krawetz, S.A. Chromatin loops are selectively anchored using scaffold/matrix-attachment regions. *J. Cell Sci.*, **2004**, *117*(Pt 7), 999-1008.
- [79] Stevens, J.B.; Heng, H.H. Differentiating Chromosome Fragmentation and Premature Chromosome Condensation In: *Human Interphase Chromosomes: Biomedical Aspects*. Yurov, Y.B., Vorsanova, S.G., Iourov, I.Y., Eds.; Springer: New York; **2013**; pp. 85-105.
- [80] Yang, F.; Alkalaeva, E.Z.; Perelman, P.L.; Pardini, A.T.; Harrison, W.R.; O'Brien, P.C.; Fu, B.; Graphodatsky, A.S.; Ferguson-Smith, M.A.; Robinson, T.J. Reciprocal chromosome painting among human, aardvark, and elephant (superorder Afrotheria) reveals the likely eutherian ancestral karyotype. *Proc. Natl. Acad. Sci. U.S.A.*, **2003**, *100*(3), 1062-1066.
- [81] Froenicke, L.; Caldes, M.G.; Graphodatsky, A.; Muller, S.; Lyons, L.A.; Robinson, T.J.; Volleth, M.; Yang, F.; Wienberg, J. Are molecular cytogenetics and bioinformatics suggesting diverging models of ancestral mammalian genomes? *Genome Res.*, **2006**, *16*(3), 306-310.
- [82] Graphodatsky, A.S.; Trifonov, V.A.; Stanyon, R. The genome diversity and karyotype evolution of mammals. *Mol. Cytogenet.*, **2011**, *4*, 22. Available from: <https://molecularcytogenetics.biomedcentral.com/articles/10.1186/1755-8166-4-22>
- [83] Rocchi, M.; Archidiacono, N.; Stanyon, R. Ancestral genomes reconstruction: An integrated, multi-disciplinary approach is needed. *Genome Res.*, **2006**, *16*, 1441-1444. Available from: <http://genome.cshlp.org/content/16/12/1441.full>
- [84] Ferguson-Smith, M.A. History and evolution of cytogenetics. *Mol. Cytogenet.*, **2015**, *8*, 19. Available from: <https://molecularcytogenetics.biomedcentral.com/articles/10.1186/s13039-015-0125-8>
- [85] Ye, C.J.; Liu, G.; Bremer, S.W.; Heng, H.H. The dynamics of cancer chromosomes and genomes. *Cytogenet. Genome Res.*, **2007**, *118*, 237-246. Available from: <https://www.karger.com/Article/PDF/108306>
- [86] Rasnick, D. *The Chromosomal Imbalance Theory of Cancer: The Autocatalyzed Progression of Aneuploidy is Carcinogenesis*. CRC Press, **2010**.
- [87] Pavelka, N.; Rancati, G.; Zhu, J.; Bradford, W.D.; Saraf, A.; Florens, L.; Sanderson, B.W.; Hattem, G.L.; Li, R. Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. *Nature*, **2010**, *468*(7321), 321-325. Available from: <https://www.nature.com/articles/nature09529>
- [88] Stevens, J.B.; Horne, S.D.; Abdallah, B.Y.; Ye, C.J.; Heng, H.H. Chromosomal instability and transcriptome dynamics in cancer. *Cancer Metastasis Rev.*, **2013**, *32*(3-4), 391-402.
- [89] Stevens, J.B.; Liu, G.; Abdallah, B.Y.; Horne, S.D.; Ye, K.J.; Bremer, S.W.; Ye, C.J.; Krawetz, S.A.; Heng, H.H. Unstable genomes elevate transcriptome dynamics. *Int. J. Cancer*, **2014**, *134*(9), 2074-2087.
- [90] Horne, S.D.; Chowdhury, S.K.; Heng, H.H. Stress, genomic adaptation, and the evolutionary trade-off. *Front. Genetics*, **2014**, *5*, 1-6. Available from: <https://www.frontiersin.org/articles/10.3389/fgene.2014.00092/full>
- [91] Horne, S.D.; Pollick, S.A.; Heng, H.H. Evolutionary mechanism unifies the hallmarks of cancer. *Int. J. Can.*, **2015**, *136*(9), 2012-2021.
- [92] Horne, S.D.; Ye, C.J.; Heng, H.H. Chromosomal instability (CIN) in cancer. *eLS*, **2015**, 1-9. Available from: <http://www.els.net/WileyCDA/ElsArticle/refId-a0006069.html>
- [93] Duesberg, P.; Li, R.; Sachs, R.; Fabarius, A.; Upender, M.B.; Hehlmann, R. Cancer drug resistance: The central role of the karyotype. *Drug Resist. Updat.*, **2007**, *10*(1-2), 51-58.
- [94] Heng, H.H. Karyotypic chaos, a form of non-clonal chromosome aberrations, plays a key role for cancer progression and drug resistance In: *FASEB Summer Research Conferences: Nuclear Structure and Cancer Vermont Academy*; Saxtons River, Vermont, **2007**.
- [95] Stepanenko, A.A.; Kavsan, V.M. Karyotypically distinct U251, U373, and SNB19 glioma cell lines are of the same origin but have different drug treatment sensitivities. *Gene*, **2014**, *540*(2), 263-265.
- [96] Heng, H.H.; Bremer, S.W.; Stevens, J.B.; Horne, S.D.; Liu, G.; Abdallah, B.Y.; Ye, K.J.; Ye, C.J. Chromosomal instability (CIN): What it is and why it is crucial to cancer evolution. *Can. Metast. Rev.*, **2013**, *32*(3-4), 325-340.
- [97] Heng, H.H.; Horne, S.D.; Stevens, J.B.; Abdallah, B.Y.; Liu, G.; Chowdhury, S.K.; Bremer, S.W.; Zhang, K.; Ye, C.J. Heterogeneity mediated system complexity: The ultimate challenge for studying common and complex diseases In: *The Value of Systems and Complexity Sciences for Healthcare*; Sturmberg, J.P., Ed.; Springer International Publishing; **2016**; pp. 107-120.
- [98] Heng, H.H.; Regan, S.; Ye, C.J. Genotype, environment, and evolutionary mechanism of diseases. *Environ. Dis.*, **2016**, *1*(1), 14-23.
- [99] Heng, H.H.; Regan, S.M.; Liu, G.; Ye, C.J. Why it is crucial to analyze non clonal chromosome aberrations or NCCAs? *Mol. Cytogenet.*, **2016**, *9*, 15. Available from: <https://molecularcytogenetics.biomedcentral.com/articles/10.1186/s13039-016-0223-2>
- [100] Hulten, M.A.; Jonasson, J.; Iwarsson, E.; Uppal, P.; Vorsanova, S.G.; Yurov, Y.B.; Iourov, I.Y. Trisomy 21 mosaicism: We may all have a touch of Down syndrome. *Cytogenet. Genome Res.*, **2013**, *139*, 189-192. Available from: <https://content.karger.com/Article/Abstract/346028>
- [101] Iourov, I.Y.; Vorsanova, S.G.; Yurov, Y.B. Chromosomal mosaicism goes global. *Mol. Cytogenet.*, **2008**, *1*, 26. Available from: <https://molecularcytogenetics.biomedcentral.com/articles/10.1186/1755-8166-1-26>
- [102] Liehr, T.; Claussen, U.; Starke, H. Small supernumerary marker chromosomes (sSMC) in humans. *Cytogenet. Genome Res.*, **2004**, *107*(1-2), 55-67.
- [103] Rommens, J.M.; Iannuzzi, M.C.; Kerem, B.; Drumm, M.L.; Melmer, G.; Dean, M.; Rozmahel, R.; Cole, J.L.; Kennedy, D.; Hidaka, N. Identification of the cystic fibrosis gene: Chromosome walking and jumping. *Science*, **1989**, *245*(4922), 1059-1065. Available from: [www.cell.com/trends/genetics/fulltext/0168-9525\(89\)90154-6](http://www.cell.com/trends/genetics/fulltext/0168-9525(89)90154-6)
- [104] Pearson, H. Human genetics: One gene, twenty years. *Nature*, **2009**, *460*(7252), 164-169. Available from: <https://www.nature.com/news/2009/090708/full/460164a.html>
- [105] Vaux, D.L. In defense of the somatic mutation theory of cancer. *Bioessays*, **2011**, *33*(5), 341-343.
- [106] Garraway, L.A.; Lander, E.S. Lessons from the cancer genome. *Cell*, **2013**, *153*(1), 17-37.
- [107] Heng, H.H. Elimination of altered karyotypes by sexual reproduction preserves species identity. *Genome*, **2007**, *50*(5), 517-524.
- [108] Gorelick, R.; Heng, H.H. Sex reduces genetic variation: A multidisciplinary review. *Evolution*, **2011**, *65*(4), 1088-1098.
- [109] Wilkins, A.S.; Holliday, R. The evolution of meiosis from mitosis. *Genetics*, **2009**, *181*(1), 3-12.
- [110] Stevens, J.B.; Abdallah, B.Y.; Liu, G.; Horne, S.D.; Bremer, S.W.; Ye, K.J.; Huang, J.Y.; Kurkinen, M.; Ye, C.J.; Heng, H.H. Heterogeneity of cell death. *Cytogenet. Genome Res.*, **2013**, *139*(3), 164-173.
- [111] Stevens, J.B.; Abdallah, B.Y.; Liu, G.; Ye, C.J.; Horne, S.D.; Wang, G.; Savasan, S.; Shekhar, M.; Krawetz, S.A.; Huttemann, M.; Tainsky, M.A.; Wu, G.S.; Xie, Y.; Zhang, K.; Heng, H.H. Diverse system stresses: Common mechanisms of chromosome fragmentation. *Cell Death Dis.*, **2011**, *2*, e178. Available from: <https://www.nature.com/articles/cddis201160>
- [112] Kussell, E.; Leibler, S. Phenotypic diversity, population growth, and information in fluctuating environments. *Science*, **2005**, *309*(5743), 2075-2078. Available from: <https://www.mimuw.edu.pl/~miekisz/leiblerfluct.pdf>
- [113] Acar, M.; Mettetal, J.T.; van Oudenaarden, A. Stochastic switching as a survival strategy in fluctuating environments. *Nat. Genet.*, **2008**, *40*(4), 471-475.
- [114] Begley, C.G.; Ellis, L.M. Drug development: Raise standards for

- preclinical cancer research. *Nature*, **2012**, *483*(5073), 531-533. Available from: <https://www.nature.com/articles/483531a>
- [115] Heng, H.H. The conflict between complex systems and reductionism. *JAMA*, **2008**, *300*(13), 1580-1581.
- [116] Abdallah, B.Y.; Horne, S.D.; Kurkinen, M.; Stevens, J.B.; Liu, G.; Ye, C.J.; Barbat, J.; Bremer, S.W.; Heng, H.H. Ovarian cancer evolution through stochastic genome alterations: Defining the genomic role in ovarian cancer. *Syst. Biol. Reprod. Med.*, **2014**, *60*(1), 2-13.
- [117] Navin, N.; Kendall, J.; Troge, J.; Andrews, P.; Rodgers, L.; McIndoo, J.; Cook, K.; Stepansky, A.; Levy, D.; Esposito, D.; Muthuswamy, L.; Krasnitz, A.; McCombie, W.R.; Hicks, J.; Wigler, M. Tumour evolution inferred by single-cell sequencing. *Nature*, **2011**, *472*(7341), 90-94. Available from: <https://www.nature.com/articles/nature09807>
- [118] Klein, C.A. Selection and adaptation during metastatic cancer progression. *Nature*, **2013**, *501*(7467), 365-372. Available from: <https://www.nature.com/nature/journal/v501/n7467/full/nature12628.html>
- [119] Wang, Y.; Waters, J.; Leung, M.L.; Unruh, A.; Roh, W.; Shi, X.; Chen, K.; Scheet, P.; Vattathil, S.; Liang, H.; Multani, A.; Zhang, H.; Zhao, R.; Michor, F.; Meric-Bernstam, F.; Navin, N.E. Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature*, **2014**, *512*(7513), 155-160. Available from: <https://www.nature.com/articles/nature13600>
- [120] Sottoriva, A.; Kang, H.; Ma, Z.; Graham, T.A.; Salomon, M.P.; Zhao, J.; Marjoram, P.; Siegmund, K.; Press, M.F.; Shibata, D.; Curtis, C. A Big Bang model of human colorectal tumor growth. *Nat. Genet.*, **2015**, *47*(3), 209-216.
- [121] Biesterfeld, S.; Gerres, K.; Fischer-Wein, G.; Bocking, A. Polyploidy in non-neoplastic tissues. *J. Clin. Pathol.*, **1994**, *47*(1), 38-42.
- [122] Bocking, A. Comparability of tumor-cytogenetics and -DNA-cytometry. *Mol. Cytogenet.*, **2015**, *8*(1), 28-32.
- [123] Lodato, M.A.; Woodworth, M.B.; Lee, S.; Evrony, G.D.; Mehta, B.K.; Karger, A.; Chittenden, T.W.; D'Gama, A.M.; Cai, X.; Luquette, L.J.; Lee, E.; Park, P.J.; Walsh, C.A. Somatic mutation in single human neurons tracks developmental and transcriptional history. *Science*, **2015**, *350*(6256), 94-98. Available from: <http://science.sciencemag.org/content/350/6256/94>
- [124] Isalan, M.; Lemerle, C.; Michalodimitrakis, K.; Horn, C.; Beltrao, P.; Raineri, E.; Garriga-Canut, M.; Serrano, L. Evolvability and hierarchy in rewired bacterial gene networks. *Nature*, **2008**, *452*(7189), 840-845. Available from: <https://www.nature.com/articles/nature06847>
- [125] Guzman, G.I.; Utrilla, J.; Nurk, S.; Brunk, E.; Monk, J.M.; Ebrahim, A.; Palsson, B.O.; Feist, A.M. Model-driven discovery of underground metabolic functions in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.*, **2015**, *112*, 929-934. Available from: <https://www.frontiersin.org/articles/10.3389/fmicb.2015.00958/full>
- [126] Ye, C.J.; Stevens, J.B.; Liu, G.; Bremer, S.W.; Jaiswal, A.S.; Ye, K.J.; Lin, M.F.; Lawrenson, L.; Lancaster, W.D.; Kurkinen, M.; Liao, J.D.; Gairola, C.G.; Shekhar, M.P.; Narayan, S.; Miller, F.R.; Heng, H.H. Genome based cell population heterogeneity promotes tumorigenicity: The evolutionary mechanism of cancer. *J. Cell Physiol.*, **2009**, *219*(2), 288-300.
- [127] Liu, G.; Ye, C.J.; Chowdhury, S.K.; Abdallah, B.Y.; Horne, S.D.; Nichols, D.; Heng, H.H. Detecting chromosome condensation defects in gulf war illness patients. *Curr. Genomics*, **2018**, *19*(3) (in press, the same issue press).
- [128] Heng, H.H. Challenges and new strategies for gulf war illness research. *Environ. Dis.*, **2016**, *1*(4), 118-125.
- [129] Telerman, A.; Amson, R. The molecular programme of tumour reversion: The steps beyond malignant transformation. *Nat. Rev. Cancer*, **2009**, *9*(3), 206-216.
- [130] Van Echten-Arends, J.; Mastenbroek, S.; Sikkema-Raddatz, B.; Korevaar, J.C.; Heineman, M.J.; van der Veen, F. Chromosomal mosaicism in human preimplantation embryos: A systematic review. *Hum. Reprod. Update*, **2011**, *17*(1), 620-627.
- [131] Nguyen, H.T.; Geens, M.; Spits, C. Genetic and epigenetic instability in human pluripotent stem cells. *Hum. Reprod. Update*, **2013**, *19*(4), 187-205.
- [132] Pellestor, F.; Gatinois, V.; Puechberty, J.; Geneviève, D.; Lefort, G. Chromothripsis: Potential origin in gametogenesis and preimplantation cell divisions. A review. *Fertil. Steril.*, **2014**, *102*(6), 1785-1796.
- [133] Zhang, S.; Mercado-Urbe, I.; Sun, B.; Kuang, J.; Liu, J. Generation of cancer-stem-like cells through the formation of polyploid giant cancer cells. *Oncogene*, **2014**, *33*(1), 116-28.
- [134] Bloomfield, M.; Duesberg, P. Karyotype alteration generates the neoplastic phenotypes of SV40-infected human and rodent cells. *Mol. Cytogenet.*, **2015**, *8*, 79.
- [135] Spittel, H.; Kubek, F.; Kreskowski, K.; Ziegler, M.; Klein, E.; Hamid, A.B.; Kosyakova, N.; Radhakrishnan, G.; Junge, A.; Kozłowski, P.; Schulze, B.; Martin, T.; Huhle, D.; Mehnert, K.; Rodríguez, L.; Ergun, M.A.; Sarri, C.; Militar, M.; Stipoljev, F.; Tittelbach, H.; Vasheghani, F.; de Bello Cioffi, M.; Hussein, S.S.; Fan, X.; Volleth, M.; Liehr, T. Mitotic stability of small supernumerary marker chromosomes: A study based on 93 immortalized cell lines. *Cytogenet. Genome Res.*, **2014**, *142*(3), 151-60.
- [136] Andriani, G.A.; Vijg, J.; Montagna, C. Mechanisms and consequences of aneuploidy and chromosome instability in the aging brain. *Mech. Ageing Dev.*, **2016**, *161*(Pt A), 19-36
- [137] Niederwieser, C.; Nicolet, D.; Carroll, A.J.; Kolitz, J.E.; Powell, B.L.; Kohlschmidt, J.; Stone, R.M.; Byrd, J.C.; Mrózek, K.; Bloomfield, C.D. Chromosome abnormalities at onset of complete remission are associated with worse outcome in patients with acute myeloid leukemia and an abnormal karyotype at diagnosis: CALGB 8461 (Alliance). *Haematologica*, **2016**, *101*(12), 1516-1523.
- [138] Bakker, B.; Taudt, A.; Belderbos, M.E.; Porubsky, D.; Spierings, D.C.; de Jong, T.V.; Halsema, N.; Kazemier, H.G.; Hoekstra-Wakker, K.; Bradley, A.; de Bont, E.S.; van den Berg, A.; Guryev, V.; Lansdorp, P.M.; Colomé-Tatché, M.; Foijer, F. Single-cell sequencing reveals karyotype heterogeneity in murine and human malignancies. *Gen. Biol.*, **2016**, *17*(1), 115. Available from: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-016-0971-7>
- [139] Horne, S.D.; Abdallah, B.Y.; Stevens, J.B.; Liu, G.; Ye, K.J.; Bremer, S.W.; Heng, H.H. Genome constraint through sexual reproduction: Application of 4D-Genomics in reproductive biology. *Syst. Biol. Reprod. Med.*, **2013**, *59*(3), 124-130.
- [140] Giorgi, D.; Farina, A.; Grosso, V.; Gennaro, A.; Ceoloni, C.; Lucretti, S. FISHIS: Fluorescence *in situ* hybridization in suspension and chromosome flow sorting made easy. *PLoS One*, **2013**, *8*(2), e57994. Available from: journals.plos.org/plosone/article?id=10.1371/journal.pone.0057994
- [141] Iourov, I.V.; Vorsanova, S.G.; Yurov, Y.B. Interphase FISH for Detection of Chromosomal Mosaicism In: *Fluorescence in situ hybridization*; Liehr, T., Ed.; Springer; **2016**; pp. 361-372.
- [142] Li, P.; Cui, C. A broader view of cancer cytogenetics: From nuclear aberrations to cytogenomic abnormalities. *J. Mol. Genet. Med.*, **2016**, *10*, e108. Available from: <https://www.omicsonline.org/open-access/a-broader-view-of-cancer-cytogenetics-from-nuclear-aberrations-to-cytogenomic-abnormalities-1747-0862-1000E108.php?aid=68303>
- [143] Belton, J.M.; McCord, R.P.; Gibcus, J.H.; Naumova, N.; Zhan, Y.; Dekker, J. Hi-C: A comprehensive technique to capture the conformation of genomes. *Methods*, **2012**, *58*(3), 268-276.
- [144] Heng, H.H. Heterogeneity-mediated cellular adaptation and its trade-off: Searching for the general principles of diseases. *J. Eva. Clin. Pract.*, **2017**, *23*(1), 233-237.
- [145] Stepanenko, A.A.; Dmitrenko, V.V. HEK293 in cell biology and cancer research: Phenotype, karyotype, tumorigenicity, and stress-induced genome-phenotype evolution. *Gene*, **2015**, *569*(9), 182-190.
- [146] Valind, A.; Gisselsson, D. Reply to Heng: Inborn aneuploidy and chromosomal instability. *Proc. Natl. Acad. Sci. U.S.A.*, **2014**, *111*(11), E973. Available from: <http://www.pnas.org/content/111/11/E973>
- [147] McClintock, B. The significance of responses of the genome to challenge. *Science*, **1984**, *226*(4676), 792-801. Available from: science.sciencemag.org/content/226/4676/792