Published in final edited form as:

Int J Cancer. 2014 May 1; 134(9): 2074–2087. doi:10.1002/ijc.28531.

Unstable genomes elevate transcriptome dynamics

Joshua B. Stevens¹, Guo Liu¹, Batoul Y. Abdallah¹, Steven D. Horne¹, Karen J. Ye¹, Steven W. Bremer¹, Christine J. Ye², Stephen A. Krawetz^{1,3}, and Henry H. Heng^{1,4,5}

¹Center for Molecular Medicine and Genetics, Wayne State University School of Medicine, Detroit, MI

²Department of Internal Medicine, Wayne State University School of Medicine, Detroit, MI

³Department of Obstetrics and Gynecology, Wayne State University School of Medicine, Detroit, MI

⁴Karmanos Cancer Institute, Detroit, MI

⁵Department of Pathology, Wayne State University School of Medicine, Detroit, MI

Abstract

The challenge of identifying common expression signatures in cancer is well known, however the reason behind this is largely unclear. Traditionally variation in expression signatures has been attributed to technological problems, however recent evidence suggests that chromosome instability (CIN) and resultant karyotypic heterogeneity may be a large contributing factor. Using a well-defined model of immortalization, we systematically compared the pattern of genome alteration and expression dynamics during somatic evolution. Co-measurement of global gene expression and karyotypic alteration throughout the immortalization process reveals that karyotype changes influence gene expression as major structural and numerical karyotypic alterations result in large gene expression deviation. Replicate samples from stages with stable genomes are more similar to each other than are replicate samples with karyotypic heterogeneity. Karyotypic and gene expression change during immortalization is dynamic as each stage of progression has a unique expression pattern. This was further verified by comparing global expression in two replicates grown in one flask with known karyotypes. Replicates with higher karyotypic instability were found to be less similar than replicates with stable karyotypes. This data illustrates the karyotype, transcriptome, and transcriptome determined pathways are in constant flux during somatic cellular evolution (particularly during the macroevolutionary phase) and this flux is an inextricable feature of CIN and essential for cancer formation. The findings presented here underscore the importance of understanding the evolutionary process of cancer in order to design improved treatment modalities.

Keywords

transcriptome dynamics; non-clonal chromosome aberrations; NCCAs; clonal chromosome aberrations; CCAs; cancer evolution; genome chaos; karyotypic heterogeneity

Various high throughput technologies have been applied to cancer research in hopes of identifying recurrent, shared signatures of cancer. Differing sets of prognostic genes have been identified at both the expression and DNA sequence level, however there has been little overlap between studies.¹⁻³ Many reasons for this disconnection have been proposed such as technological immaturity, insufficient sample sizes, sampling errors and statistical methods.^{1,2} While these factors are challenging, they have largely been addressed by technological advancement, but the disconnection remains. As evidence, expression profiles of randomly selected 100 gene lists have at least as much diagnostic utility in breast cancer as gene sets identified as significant predictors of breast cancer outcome in previous studies.⁴

According to the gene mutation theory, cancer evolution is a stepwise process where accumulation of common genetic change(s) leads to a more cancerous phenotype.⁵⁻⁷ Current recommendations for reporting karyotype change focus only on clonal change despite the inability to identify commonly shared aberrations in most cancer.⁸ Recent studies however show that cancer is a disease of stochastic somatic evolution driven primarily by stochastic chromosome instability (CIN).^{9,10} Genome evolution during cancer progression occurs in two iterative cycles of punctuated and stepwise chromosomal change, respectively.¹⁰⁻¹³ Studies of chromosome evolution during cancer progression led to the elucidation of the evolutionary mechanism of cancer which is equal to the sum of all potential molecular mechanisms.¹⁴ The evolutionary mechanism of cancer explains why most individual gene expression markers provide little predictive value with regard to cancer outcome, as somatic cell selection acts to select a “winning” karyotype (genome defined system) rather than parts of the system (such as individual genes).^{12,13,15,16} This concept is consistent with findings from cancer genome sequencing projects that fail to identify commonly shared mutations for most tumors.^{9,17,18} The fact that nearly every gene, pathway or network marker can contribute to cancer progression under certain conditions challenges the idea of applying these pathways/markers to cancer therapy and therefore their clinical value.^{17,19,20}

It is well established that gene mutations can affect the transcriptome and resulting networks, however the impacts of punctuated chromosomal change that occurs during the macroevolutionary phase of cancer evolution and genomic heterogeneity on the transcriptome is unknown. Genome theory states that the genome context (gene content plus genomic topology) or more simply the karyotype defines the inheritance of the system. Recent analyses support this viewpoint by illustrating that karyotypic change alters cellular networks.^{15,16,21-23} It is thus necessary to study how karyotypic change influences global gene expression throughout both phases of cancer evolution, especially using a well-defined system with clonal and non clonal cell populations.^{10,17,22,24,25} Previously clonal changes in chromosomal copy number have been linked to changes in gene expression.²⁶ In many cases a basic correlation between amplification of gene copy number with increases in expression of those genes has been found, but this correlation does not always hold true. For instance 62% of highly amplified genes have been shown to have slightly to extremely elevated expression levels, while 38% have no expression changes or suppressed expression.²⁷ Thus

positive correlations between copy number change and alteration in gene expression have been the focus while negative links are often overlooked.

Here, spectral karyotyping (SKY) and transcriptome expression were co-monitored throughout the process of immortalization and drug treatment, using a model of spontaneous immortalization of fibroblasts from a patient with Li-Fraumeni syndrome.^{10,28} SKY was used to monitor genome level change as SKY provides needed population resolution that array and sequencing based strategies of monitoring genomic change miss. It has been proposed that a lack of population based resolution may be a significant contributor to the discrepancy between the identification and application of cancer gene sets.^{15,29} In addition, karyotype change is a better predictor of cancer progression than the monitoring of mutations.^{9,30} Furthermore, genome level change provides cells with a unique ability to overcome environmental hardships more effectively than gene mutations.

Cells in the above mentioned immortalization system exhibited various levels of karyotypic heterogeneity as they progressed from the macroevolutionary phase of cancer to the microevolutionary phase of cancer allowing for the comparison of expression patterns and their diversity with genome alteration heterogeneity to better understand the dynamics of the entire system. Our systematic comparison through an entire run of the somatic evolutionary process leading to immortalization reveals distinctive pattern(s) of transcriptome profiles during different phases of somatic cell evolution. This data underscores the importance of the karyotype, as the pattern of gene expression changes is dependent on the karyotypic context and heterogeneity. Thus gene expression profiles provide limited information on how specific pathways contribute to cancer evolution, especially during periods when NCCAs dominate. However, when used in conjunction with appropriate monitoring of karyotypic patterns, transcriptome profiles give additional information about overall system dynamics and the resultant evolutionary potential of a population.

Material and Methods

Cell culture

Culture of MDAH-041 cells was performed as previously reported.^{10,28,31}

Spectral karyotyping

Spectral karyotyping was performed as described.^{10,14}

RNA isolation

RNA was isolated with Trizol (Invitrogen) according to the manufacturer's recommendations.

Microarray hybridization

Biotinylated cRNA were prepared according to the manufacturer's protocol (Whole-Genome Gene Expression with Intelli-Hyb Seal System Manual, 2006, Illumina) and hybridized to Illumina WG-8 v2 arrays (Chip IDs: 1787466094 & 1714478097). For comparison of replicates collected from the same flask, biotinylated cRNA was hybridized to two Illumina

human v2 WG-6 arrays. Sentrix Expression BeadChip, staining and washing were carried out as described (Illumina, Wallingford). Scanning procedures were carried out according to manufacturer's recommendations (Illumina, Wallingford).

Microarray analysis

Normalization—Raw intensities from arrays were Z-score normalized to the mean of each replicate.³² Briefly Z-scores were computed by subtracting the mean intensity of all measured probes in the array from the intensity of a given probe and dividing that by the standard deviation of all measured intensities based on the formula:

$$Z_{score} = \frac{(\text{Intensity}_{\text{Probel}} - \text{Mean intensity}_{\text{All probes from replicate samples}})}{\text{Std Dev}_{\text{All probes from replicate samples}}}$$

Expression was directly compared by calculating Z ratios using Diane 6.0 software running on the JMP platform.³³ For comparisons across passages (either passage 7 vs. other passages or nearest neighbor comparisons), previously computed Z-scores were averaged from each replicate. Z ratios were calculated by subtracting the “control” Z-score from the “experimental” Z-score. The results were then divided by the standard deviation of all of the calculated Z-score differences, according to the formula:

$$Z_{ratio} = \frac{[(Z_{score_{Gene1avg}})_{Exp} - (Z_{score_{Gene1avg}})_{Ctrl}]}{\text{Std Dev}_{\text{All Z-score differences}}}$$

In these analyses Z-ratios of ± 1.96 were considered significant ($p < 0.05$) although previous empirical observations have suggested significance ± 1.50 maybe significant.³²

Microarrays were also analyzed using GeneSpring GX (Agilent Technologies). In this analysis, expression values were normalized using a percentile shift. Raw signals were thresholded to 1.0 with a percentile target of 75. Values were then baselined to the median of all samples. All entities that were flagged as present were retained. Retained values were analyzed using ANOVA with post hoc Tukey's HSD and Storey curve fitting for multiple testing corrections³⁴. Significant transcripts with changes in expression > 1.5 fold were subjected to pathway analyses.

Pathway analyses

Gene ontology analysis was performed using Genespring GX version 12 using transcripts whose expression was found to be significantly changed (corrected $p < 0.05$) by ANOVA analysis at greater than 1.5-fold change. Pathways reported were found to be significant at $p < 0.1$. NLP networks were constructed from the Natural Language Processing (NLP) algorithm in Genespring. NLP networks are presented in “neato” format with unattached nodes removed.

Data access

Microarray data is available in the GEO database (accession numbers GSE46931 <http://1.usa.gov/YHBwfb> and GSE46932 <http://1.usa.gov/12yjoUC>).

Results

I. Karyotype alteration is associated with gene expression change

A. Pattern of expression changes in concert with the karyotype—The spontaneous immortalization model used here reveals two distinct phases of somatic evolution, the stepwise phase, where most cells share a similar karyotype, and the punctuated phase where the majority of cells have unique karyotypes.¹⁰ The link between non-clonal karyotypic change and cancer macroevolution raises the possibility that cell populations with high levels of karyotypic heterogeneity may also have high levels of expression heterogeneity.²⁹ To address this issue, karyotypic alteration and transcriptome change was monitored during spontaneous immortalization. Pairwise transcriptome comparisons using either a Z-test ratio or ANOVA with post-hoc pair-wise comparisons were performed. Two post-hoc comparisons were made, pairwise nearest neighbor comparisons (*e.g.*, Pd7 *vs.* Pd17, Pd17 *vs.* Pd25, Pd25 *vs.* Pd 54 and Pd54 *vs.* Pd109) and comparison of all stages to Pd7, as Pd7 is the closest sample to normal, and therefore is similar to a comparison of tumor tissue to non-transformed tissue. Overall in the ANOVA analysis 2,345 and 2,669 transcripts were found to be differentially expressed in pairwise comparisons and Pd7 compared to all other samples respectively (Supporting Information Tables 1a and 1b). Calculation of Z-test ratios identified 936 differentially expressed transcripts in Pd7 *versus* all other passages and 1102 in nearest neighbor comparisons (Supporting Information Table 2a and 2b). A major change in gene expression is evident following Pd7 (Fig. 1a).

B. Decreased karyotypic stability influences expression variance in replicates

—We next determined the effects of the clonal and non-clonal phases of somatic evolution on the transcriptome (Fig. 1a). The overall pattern of expression is similar to what is observed at the karyotypic level, with large deviations in expression when genome replacement takes place. For example the magnitude of deviation in expression is largest in the shift from the stable (stepwise) phase of somatic evolution to the unstable (punctuated) phase, and during the unstable phase (Fig. 1b). To further address the relationship between genome and expression level heterogeneity, principal component analysis (PCA) was performed to compare similarities in expression profiles of replicates of samples. As expected, replicates from an Ambion reference RNA group closely to each other as do the Pd7 samples. Pd17 and Pd 25 replicates (both represent unstable populations) do not group closely, while replicates from stable stages such as Pd7 and Pd109 do (Fig. 1c). Treatment of Pd109 with doxorubicin which induces genomic instability results in a large distance between the two replicates compared with global expression in untreated Pd109 cells. These results suggest that replicates from stages with clonal karyotypes have a more similar expression profile than replicates from stages of progression with elevated NCCA frequency. This data shows that high levels of genomic heterogeneity result in gene expression changes that depend on karyotypic heterogeneity within a given sample, similar

to acute myeloid leukemia (AML) which is noted for heterogeneous karyotype change and heterogeneous protein expression.³⁵

II. Somatic evolution is a stochastic process

A. Karyotype data—Chromosomal change is endemic throughout immortalization. Even early on at Pd7 only 11% of cells contained a normal karyotype (Fig. 2a). In all passages following Pd7 no karyotypically normal cells are detectable. Aneuploidy is evident at Pd7, jumping from on average under 60 chromosomes per cell to over 70 at Pd17 followed by a small decrease at Pd25 and an increase to ~84 chromosomes per cell at Pd109. (Fig. 2b; Table 1). Pd25 however shows the largest variation in chromosomal number ranging from 25 chromosomes to 226. Structural chromosomal changes follow a similar pattern to numerical changes with total structural change peaking prior to immortalization at Pd25 (average of 8.87 per cell structural aberrations including NCCAs and CCAs). Following immortalization chromosomal copy number is also altered, however much of this aneuploidy is shared within the population. At Pd109 shared aneuploidy is largely lost again, in agreement with structural data where an increase in NCCAs is observed (Supporting Information Figs. 1a–1e). Overall, this data demonstrates that structural and numerical chromosomal change occurs in two phases during somatic evolution.

B. Correlating karyotype and expression data—Average Z-score values for each chromosome were calculated at each passage and regression analysis was applied to these values and the average copy number of each chromosome at each measured stage. We were unable to find a significant correlation between chromosomal copy number and overall expression of genes located on that chromosome (Supporting Information Table 3). The largest R^2 value was only 0.09 ($p = 0.16$) at Pd54. All other passages had far lower R^2 and p -values suggesting no direct link between chromosome level copy number change and overall gene expression.

C. Pathways altered during immortalization—Somatic evolution is stochastic at the karyotypic and transcriptomic levels. Gene Ontology (GO) analysis of ANOVA results from nearest neighbor and all passages compared to Pd7 comparisons identified 49 total pathways that were involved in the immortalization process, however 22% of these pathways do not overlap the two comparisons (Supporting Information Table 4).

When all passages are compared to Pd7, 44% of the significant GO terms were both up- and down-regulated at least one time point in the entire course demonstrating that multiple pathways are in flux throughout somatic evolution.

A total of 101 and 149 pathways were found to be significantly altered ($p < 0.05$) by Wikipathway analysis of transcript expression values for all passages *versus* Pd7 and for pairwise comparison of neighboring passages, respectively. In nearest neighbor pairwise comparisons a portion (27% total, 16% up and 11% down) of these over-represented pathways are over-represented in only one stage comparison (e.g., only significantly up in Pd25 vs. Pd54) (Supporting Information Table 5a). 12% of identified pathways are significantly up at a given comparison and down at another. 58% of the significantly altered networks were both down and up at various stages. This data shows that a unique

combination of differing pathways is utilized during immortalization, and there is no overall central, critical pathway throughout the entire evolutionary process of carcinogenesis.

Networks significantly altered in Pd7 *versus* other passage comparisons followed a different pattern than pairwise comparisons with only 17% of altered networks showing both significant increase and decrease during the entire time course (Supporting Information Table 5b). 14% of the affected networks showed increase in expression in one comparison, and 16% showed this in two. Of the down regulated networks, the majority were shared in two comparisons (17%) or in all four comparisons (13%). Comparison of RNA expression changes between Pd7 affirms that the largest change to RNA expression occurs following Pd7 when karyotypic evolution accelerates. This data suggests that for the genes and networks whose expression is significantly altered when compared to Pd7, this change in expression is not static throughout somatic evolution. For example here the majority of significantly altered pathways are both up at some point(s) and down at other(s). This analysis indicates the dominant impact of genome level change on the adaptive genomic landscape of cancer.^{9,17}

GO analysis of nearest neighbor and Pd7 *versus* all other passages revealed significant alteration primarily related to metabolic change, but did not directly identify pathways linked to genomic stability. WikiPathway analysis on the other hand identified a number of significantly altered pathways related to genomic instability including DNA damage response, base excision repair, and homologous recombination. The majority of these pathway alterations occur prior to Pd54 when there are large changes in the karyotypes of the measured populations, suggesting disruptions in the ability to maintain genome stability play an important role in cancer evolution.

Previous studies of the immortalization process in Li-Fraumeni fibroblasts implicated significant alterations in interferon, cell cycle and cytoskeleton GO terms, however those terms were not found to be significant in this study.^{31,36} GO analysis resulted in a list of significant terms that included multiple terms with little informativeness such as “DNA” or “RNA,” so data were further analyzed using WikiPathways. In contrast to GO analysis WikiPathway analysis did identify a significant increase in “Regulation of Actin Skeleton” in Pd17 *versus* Pd7 and Pd25 *versus* Pd7, however, no significant alterations were identified following immortalization (Fig. 3a). Similarly, “G1 to S cell cycle control” was significantly down at Pd17 *versus* Pd7 as was “Cell Cycle” at Pd17 *versus* Pd7 and Pd109 *versus* Pd54 (Fig. 3a). Interferon pathway(s) alterations were more complex. In Pd7 *versus* other passage comparisons, both Type 1 and 2 interferon signaling are significantly increased at Pds 17 and 25, and are significantly decreased at Pds 54 and 109. Interestingly when nearest neighbor comparisons are made, interferon signaling is revealed to be much more dynamic. Both type I and II interferon signaling is significantly up regulated at Pd17 compared to Pd7, is significantly decreased at Pd25 compared to Pd17, and is again significantly decreased at Pd54 compared to Pd25 (Fig. 3a). The two pathways deviate from each other at Pd109 *versus* Pd54 with type II being significantly upregulated and type I signaling being significantly down regulated. This underscores the importance of “observing evolution in action” throughout carcinogenesis, as the genetic pathways involved in the process are in stochastic flux.¹⁷

Cancer progression has been thought to occur in a stepwise fashion where mutations and changes in expression of onco- and tumor suppressor-genes accumulate progressively as cells become more tumorigenic. Similar to expression of most genes, expression of oncogenes and tumor suppressor genes does not follow a pattern, as expression of genes of each type are both increased and decreased throughout immortalization (Figs. 3b and 3c). Overall according to gene based theories of cancer progression, throughout cancer progression there should be an overall decrease in expression of tumor suppressor genes and an increase in expression of oncogenes. The data presented here shows neither. This demonstrates that while specific tumor suppressor- and oncogenes may play a role in a given tumor at specific stages, the roles they play are limited to a specific tumor with a specific karyotype at a specific time point. These results closely mirror the recently hypothesized evolutionary mechanism of cancer, where diverse genetic and/or environmental factors contribute to cancer in specific cases but not in general.³⁷

Taken together, these results show that overall there is a large change in expression following Pd7 (Fig. 1a). This is underscored by Genespring NLP pathway analysis which can be used to build networks interaction structures based on a Natural Language Processing algorithm that mines published abstracts for interaction data. NLP demonstrates that once again the transcriptomic data largely follows karyotypic data as the largest network changes occur during periods of karyotypic instability, and demonstrates no recurrent pattern of pathway alteration (Figs. 4a–4d). Karyotypically the Pd7 population is closest to normal and despite achieving karyotypic stability following immortalization the karyotypes continually evolve to become more abnormal. Thus from stage to stage there is strong evidence for network switching in that the genetic networks that play an important role at each stage change as the karyotypes from which these networks are derived change. This may offer insight into recent results showing that stochastically selected gene signatures are just as effective at identifying cancers as “cancer gene signatures” as the one thing that all cancers share is a deviation from normal both at the karyotypic and gene expression levels.³⁸ These data mirrors recent data showing that karyotypically defined subtypes of AML utilize unique protein signaling networks.³⁹ The above data indicate that specific gene based pathways only play fleeting roles during somatic evolution as the gene networks functioning within cells are therefore highly dynamic.

III. Validation

Above, it was demonstrated that expression variation reaches its peak when genomic instability is highest. To validate this finding we used two approaches. First, destabilization of the genome by treatment of MDAH-041 cells at a karyotypically homogeneous stage (Pd109) with doxorubicin induces karyotypic instability resulting in a decreased correlation in expression in replicates (Fig. 1a). Second, to validate data linking karyotypic alteration to transcriptome change, sub-clones selected for karyotypic stability (stable or unstable) were isolated from various stages. Clones were grown in a single flask and mitotic cells were collected for SKY (Data not shown). Each flask was then divided into two halves and RNA was collected separately from the two halves of each flask as replicates. This approach has previously been used to assess technical reproducibility with the underlying premise that all

of the cells within the assayed population are the same. Here we employ the same strategy to determine the contribution of karyotypic variation to expression variation.

If genome context plays a large role in determining expression, expression profiles of replicates from karyotypically heterogeneous sub-clones (Unstable-1, Unstable-2 and Stable-1 Dox) should show a larger difference in expression than replicates from the stable sub clone. Indeed this is verified by PCA (Fig. 4e). Interestingly Tan-065 and CCL-110 cell lines were reported to be normal diploid cell lines, however PCA of RNA profiles of same flask replicates of these lines indicated a similar expression difference as same flask replicates of unstable MDAH-041 sub-clones. Subsequent SKY analysis revealed Tan-065 and CCL-110 cells used here to be karyotypically heterogeneous at a level similar to unstable MDAH-041 clones (Data not shown).

This data not only validates above observations but also indicates that monitoring karyotypic change results in transcriptional change, underscoring the importance of verifying the karyotypic makeup of cell lines before and after experiments are carried out. The link between genome heterogeneity and transcriptome heterogeneity presented here is significant as it demonstrates a mechanistic link between genome heterogeneity and heterogeneity of other features of the cellular system such as growth that have recently been described.⁴⁰ This relationship holds true regardless of whether the karyotypic heterogeneity occurs spontaneously during culture or is induced by drug treatment. Other stresses including changes in culture conditions and other environmental stresses would likely have similar effects as stress in general promotes karyotypic change, especially in cells that have bypassed the constraint of a healthy, unchanged genome. Cells with uncompromised genome constraint undergo cell death when exposed to sufficient stress.⁴¹ This underscores the importance of single cell analysis as cells in karyotypically homogeneous populations are likely to share similar expression profiles whereas expression profiles from individual cells from karyotypically diverse cell populations are likely to vary widely.

Discussion

I. Expression changes reflect karyotypic change

Organization of the genome architecture (genomic topology) is the highest level of genetic organization and defines the genetic system.^{15,16} Genome organizational change has been shown to be important in somatic evolution leading to cancer, other diseases and organismal evolution.^{42,43} Even single numerical and structural changes to a given genome can influence the expression of hundreds of genes.⁹ Stochastic genome change throughout immortalization and especially in the macroevolutionary phase prior to immortalization resulted in constant flux of the overall expression of any given pathway. This suggests that each run of evolution is unique in both the karyotype and in the gene networks that function during the process. The uniqueness of each individual run of evolution raises an important problem, in most cases of cancer, even within the same tumor type, a specific molecular mechanism only contributes to a minority of isolated cases (especially under experimental settings), limiting its relevance to the large number of potential mechanisms.^{14,17} While it is true that individual genes such as transcription factors can completely alter a transcriptome profile, if their expression is altered within the same genome (for instance in clonal

populations) a recurrent pattern of transcriptional change is readily identifiable, however the changing expression levels of the same transcription factor within a karyotypically heterogeneous population will result in multiple differing transcriptome profiles. Selection of a successful genome and its resulting functional pathways is dependent on multiple factors including environmental conditions, growth and death rates of the various cells in the population, and genomic stability of the population. In fact, multiple changes in genes, pathways and processes have been linked to immortalization including mutation altered signaling in the following pathways, INK4A (p16), INK4B (p15), p19/ARF, cell cycle pRB/p53, interferon-related, insulin growth factor-related, MAP kinase and oxidative stress pathway as well as changes to the cytoskeleton and aneuploidy.^{36,44–46} The plethora of pathways or changes linked to immortalization supports the evolutionary mechanism of cancer hypothesis, as regardless of which pathway(s) dominate in a specific experimental setting at a specific time, the common mechanism of cancer progression is the evolutionary selection of the genome.^{15,17} The dynamics of genome level and transcriptome changes have important clinical implications as their dynamic nature demonstrates how large scale change can occur over a short time scale. This is particularly important in designing cancer treatments as these treatments are often based on a set of markers that are found a period of time prior to treatment. Large scale chromosomal change occurring over only a few cellular generations could significantly alter the tumor landscape, reducing the efficacy of treatment.

A. Effect of copy number change on expression—Previous studies of aneuploidy have contrastingly shown correlation, anti-correlation and a lack of correlation to gene expression levels and copy number changes.^{22,25,26} Most studies that have successfully linked copy number variation to expression changes have lacked population level resolution as these studies utilize assays based on molecular averaging such as array based SNP or CGH platforms have successfully identified links between copy number change and changes in expression.^{26,47} Even in cases where a correlation is found, the correlation does not hold true for all genes.²⁷ Expression and copy number was measured in a panel of radiation hybrid cell lines and compared to data from exome sequencing of glioblastoma multiform tumors.²⁶ 30% of quantitative trait loci had an anti-correlation, and approximately 1/3 of the genome had a copy number-expression correlation in glioblastoma. 1/3 of these genes from glioblastoma did not overlap with the hybrid data suggesting again that the genome context is a major determinant of overall expression. Interestingly it was found that repression of QTL expression from the X chromosome was not due to Xist function but rather some alternative process underscoring the importance of the genome context in the function of cellular processes rather than a specific pathway per se.²⁶ In the present study no correlation was identifiable between chromosomes with increased or decreased copy numbers and changes in transcript expression.

The above lack of correlation may be due to either a true lack of correlation or it could be due to the karyotypic complexity of the cells and the high degree of karyotypic variation within the populations of cells. For example, even at stable stages such as Pd54, each cell contains multiple NCCAs and complex shared rearrangements. To resolve this, a combination of CGH to identify sub-chromosomal amplifications with SKY to monitor NCCAs may elucidate a correlation between transcript expression and copy number. It

should be pointed out those previous studies assessing correlations between expression and karyotype change lacked complex karyotypic change. Complex karyotypic change is often associated with high frequencies of NCCAs (for example Pd17 in Fig. 2), and in these cases methods that rely on average molecular profiles are problematic. Nevertheless, we have tried to link to transcriptomic dynamics to clonal chromosome aberrations without success. The lack of overlap shows that any given molecular mechanism can drive the evolution of cancer and that each run of evolution is unique. Single cell analyses are therefore essential when NCCA frequencies are high.

We compared our data with a previously reported expression signature of aneuploidy where 94 and 137 genes were shown to be up and down regulated respectively in aneuploid cells.⁴⁸ Of the 231 genes of this signature we found significant ($p < 0.05$) change in expression at least one time point compared to Pd7 in 17% of aneuploid signature genes (Supporting Information Table 6). However, significantly changed genes did not follow the previously published patterns of expression, as some genes of the aneuploid signature were previously found to be over-expressed while they were down regulated here and vice versa.

B. Expression patterns during stable and unstable phases—Although a specific correlation between transcripts and copy number was not found, an important correlation was identified between chromosomal and expression stability. Replicates from stages with stable, shared karyotypes have more similar expression profiles than replicates from unstable non-shared karyotypes stages of progression (Figs. 1a and 3a). This relationship is highlighted by experiments where the karyotypes of a single flask of cells was observed and then RNA from two halves of that flask was then harvested independently and subjected to expression profiling. This experimental setup has been used in the past to assess reproducibility of technological platforms. Here, we have used a mature technology that has been shown to be highly reproducible to monitor the effects of karyotypic heterogeneity on transcript expression, and have shown that cells with divergent karyotypes typically have divergent expression profiles. These results call for caution in interpreting results from assays and analyses that homogenize heterogeneous cellular populations. Homogenization of a largely karyotypically homogeneous population may truly reflect the biology of that population, but for genomically heterogeneous populations methods that average the population should be avoided in favor of single cell analyses, or at the very least be interpreted with care.

II. Gene networks may be influenced by karyotypic change and evolutionary processes

Gene products function by forming complex networks of interaction. Many networks have been identified, but they are not static entities, and novel pathways and features can be acquired by genome topology alteration.^{15,23,49} The data presented here, supports the model that chromosome change alters the functional network structure of the cell by creating new genomic topologies which lead to novel pathways and/or novel interactions (Fig. 5).^{15,16,50} Genome context defined network/pathway dynamics are demonstrated here as significantly altered pathways do not follow a specific pattern of alteration during cancer progression (Figs. 5b and 5c). In fact, most pathways are only significantly altered in only one pairwise comparison, and of the fraction that are altered in multiple comparisons, many are both up

and down regulated throughout the evolutionary process demonstrating that pathway contribution is not static, and suggesting that each round of somatic evolution leading to cancer is unique. Fundamentally, success of the evolutionary process especially within the punctuated phase is not dependent on a specific gene or pathway (parts), but is dependent on genomic heterogeneity of the population.¹⁵

Reproducibility is a big challenge for expression profiling of cancer. Major efforts have been focused on improving technological platforms and analytic methods however, one overlooked problem is the overall genome dynamics of the samples being profiled.^{1,2,10} Our data shows that reproducibility is lower from samples with high population diversity at the genome level than samples with shared structural and numerical change, suggesting that each unique genome results in a unique transcriptome. Thus much of the variability could be attributable not to the platform of analysis, but to the system itself as cells that do not share the same genome makeup do not share the same transcriptional profile. Over time, introduction of genome dynamics alters the system resulting in changes in levels of available proteins and functional networks, which is why chemotherapy works so much better in hematological malignancies than in solid tumors.⁵¹ The stability of the genome system also explains why the causality of specific gene mutations can be illustrated in certain cancers and in well-defined experimental models, but not in patient populations where consistent genome change occurs during somatic evolution, altering the interaction network structure continually, and reducing the impact of inhibiting a given pathway or gene.¹⁵

III. Implications

The data presented here cautions against simply focusing on specific genes as causative factors of cancer.²⁰ If there is a paucity of recurrent gene mutations in the vast majority of tumors, how can this be reconciled with the fact that specific oncogenes and tumor suppressor genes seem to cause cancer in specific cases and in animal models? This disparity can be reconciled by the evolutionary mechanism of cancer.^{17,19} Based on the evolutionary mechanism of cancer, the only recurrent feature of cancer(s) is a departure from normal. Tumors may share key phenotypic hallmarks but, these hallmarks emerge from a multitude of genome and gene changes.^{17,19} Departure from normalcy in cancer has been verified at all levels of organization including the genome/chromosome level, the sequence level and the gene expression level.¹⁵ First, at the genome/chromosome level clinical data has revealed that most tumors do not share karyotypic features and previous studies have demonstrated stochastic turnover of the karyotype during cancer evolution (Fig. 5).

Second, as previously mentioned, cancer genome sequencing projects have failed to identify recurrent mutations in most tumor types.¹⁵ Third, a recent report found that stochastically selected gene signatures were at least as effective at predicting breast cancer as previously identified signatures of breast cancer, demonstrating again that regardless of which list of genes is observed, that simply a divergence from normal is the signature of cancer.³⁸ Single cell whole genome sequencing has confirmed karyotypic heterogeneity exists commonly in tumors, in addition increased genome heterogeneity has been reported in both normal and cancerous tissues.^{11,14,52} The work presented here suggests that each cell with a unique

karyotype will have a unique transcriptome profile. Such karyotype mediated transcriptomic heterogeneity should be present in most tumors as most tumors have a high frequency of NCCAs.¹⁶ This further supports the concept that NCCAs provides significant functional diversity and evolutionary potential.¹⁴ This genome alteration mediated transcriptomic heterogeneity could further contribute to the non-genetic heterogeneity of a tumor.⁵³ In practice, however, non-genetic heterogeneity may be more apparent during the microevolutionary phase of cancer evolution as the large scale impact of genome alteration during the macroevolutionary phase likely over powers, reducing the significance of non-genetic effects.⁵³ Interestingly evidence of the functional importance of karyotypic heterogeneity has recently been shown where karyotypically diverse populations of cells have more heterogeneous growth rates compared to clonal populations.⁴⁰ Follow-up studies are urgently needed to jointly measure single cell transcriptomes and genomes in normal tissues and in tumors in vivo to better understand how genomic heterogeneity creates evolutionary potential in cancer.

If deviation from normal at multiple levels of organization is the signature of cancer, identifying the most definitive level of organization is necessary.^{16,18} The above data demonstrates that changes in genome structure result in changes to expression patterns, and those populations with multiple divergent karyotypes have more transcriptome variation. Thus monitoring change at the genome level gives insight to both karyotypic and expression heterogeneity. Furthermore, karyotype data is superior to sequence data, as changes to the karyotype promote adaptive evolution more strongly than sequence changes, karyotype data can identify structural changes of chromosomes, and karyotyping is more cost effective than whole genome sequencing.^{15,17} Even if the cost to sequence an entire genome at the single cell level drops significantly, sequencing 20 or more cells is required from each tumor in order to assess population heterogeneity while only one spectral karyotyping would be required for the same sample.

This work underscores the need for the development of robust platforms to study cellular heterogeneity at multiple levels of organization especially at the genome level.^{17,18} It also calls for refinements of current methods of analysis. For example including spiked in reference RNAs can allow for better normalization schemes.⁵⁴ The majority of tumors are heterogeneous at multiple levels including mutational profiles, gene expression signatures, and karyotypes. Understanding and correctly identifying this heterogeneity will lead to improved treatments that reduce occurrence of drug resistance.

Clinically the above results will have an impact as they demonstrate the importance of monitoring genome heterogeneity to gauge evolutionary potential.¹⁷ This is important for determining initial treatment and managing ongoing treatments. Consider two tumors of the same type. One of which is karyotypically homogeneous and one heterogeneous. The homogenous tumor could be predicted to be more responsive to targeted therapy than the heterogeneous tumor. Similarly, the presence of high levels of heterogeneity should lead to caution in basing treatment on markers from biopsies which have been performed some time prior to commencement of treatment. The presence of karyotypic heterogeneity in this case increases the probability that the tumor will have further evolved. Karyotypic heterogeneity is also significant for acquisition of therapeutic resistance. Karyotypically diverse

populations of cells contain diverse transcriptomes which increases the probability of one or more cells expressing proteins or pathways that would allow for escape from treatment pressure. Similarly a patient with a more stable tumor would be expected to respond more positively to targeted therapy.¹⁸ Measuring genome level changes during treatment also provides insight into whether a given treatment is reducing tumor mass while inducing genomic instability in the surviving cells, increasing the probability of a resistant survivor. Finally, this work may help to explain the discrepancy between positive experimental results in cell lines and animal models versus results in clinical settings. In this case experimental systems may be much more homogeneous on the genome level and the resultant transcript level, while in practice tumors that evolve slowly over years or decades are much more heterogeneous and thus more likely to escape selection pressures of given treatments.

Despite acknowledging the importance of genomic instability in cancer progression, it is often assumed that cancer cells contain identical genomes or that “driver” genomic change persists during cancer progression with a background of non-shared “passenger” change.^{15,16} This idea is based on findings in hematological malignancies such as chronic myelogenous leukemia where highly penetrant chromosomal change drives cancer progression. These highly penetrant clonal changes are rare in solid tumors where punctuated evolution frequently occurs.¹⁶ Genomic heterogeneity increases the probability that one or more cells may survive a given selection pressure by allowing each individual cell to utilize a unique set of pathways and in some cases create novel pathways.¹³ The pathways on which each individual cell uses may switch as the genome changes stochastically throughout cancer progression. Therefore understanding the effects of genomic heterogeneity on the networks within these cells will increase our ability to treat cancer. New strategies to profile both karyotypic and transcriptional heterogeneity in a systematic manner are thus required to better understand and treat cancer.

In conclusion, genome alteration mediated cancer evolution can be explained by elevated transcriptome dynamics caused by said genome alterations rather than recurrent, shared changes to specific pathways. This relationship or molecular mechanism is especially true in heterogeneous cancer cell populations that constantly change where each unique karyotype would have a unique transcriptome that would change constantly as the karyotype evolves during cancer evolution.^{13,14} Despite the importance of each individual pathway in a given cell at a given stage, no specific pathway(s) are responsible for the entire process of cancer evolution. Thus it is challenging to identify the overall pathway or genetic signatures for most cancers, as elevated transcriptome dynamics, rather than specific gene mutations, provide evolutionary potential. The data presented here demonstrates the important of monitoring on karyotypic alterations, especially when these alterations are heterogeneous, in order to better understand biological issues including cancer and complex diseases.^{55,56}

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This manuscript is part of a series of studies entitled “The mechanism of somatic and organismal evolution.” The authors thank Graham Johnson for assistance in data retrieval, Adrian Platts for valuable input, Dr. Michael Tainsky for providing the MDAH-041 cell line, Drs. Kevin Becker and Supriyo De for providing the DIANE software.

Grant sponsor: Komen Foundation, See DNA Inc., and the US Department of Defense; **Grant number:** GW093028; **Grant sponsor:** WSU CMMG post-doctoral fellowship (to J.B.S.)

References

1. Chan SK, Griffith OL, Tai IT, et al. Meta-analysis of colorectal cancer gene expression profiling studies identifies consistently reported candidate biomarkers. *Cancer Epidemiol Biomarkers Prev.* 2008; 17:543–52. [PubMed: 18349271]
2. Van Haren R, Feldman D, Sinha AA. Systematic comparison of nonmelanoma skin cancer microarray datasets reveals lack of consensus genes. *Br J Dermatol.* 2009; 161:1278–87. [PubMed: 19681882]
3. Evans JP, Meslin EM, Marteau TM, et al. Genomics. Deflating the genomic bubble. *Science.* 2011; 331:861–2. [PubMed: 21330519]
4. Venet D, Dumont JE, Detours V. Most random gene expression signatures are significantly associated with breast cancer outcome. *PLoS Comput Biol.* 2011; 7:e1002240. [PubMed: 22028643]
5. Nowell PC. The clonal evolution of tumor cell populations. *Science.* 1976; 194:23–8. [PubMed: 959840]
6. Vogelstein B, Kinzler KW. The multistep nature of cancer. *Trends Genet.* 1993; 9:138–41. [PubMed: 8516849]
7. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell.* 2000; 100:57–70. [PubMed: 10647931]
8. Shaffer, LG.; Slovak, ML.; Campbell, LJ. ISCN 2009 an international system for human cytogenetic nomenclature. Karger, Basel; Switzerland: 2009.
9. Heng HH, Bremer SW, Stevens JB, et al. Chromosomal Instability (CIN): what it is and why it is crucial to cancer evolution. *Cancer Met Rev.* 2013 Epub ahead of print.
10. Heng HH, Stevens JB, Liu G, et al. Stochastic cancer progression driven by non-clonal chromosome aberrations. *J Cell Physiol.* 2006; 208:461–72. [PubMed: 16688757]
11. Navin N, Kendall J, Troge J, et al. Tumour evolution inferred by single-cell sequencing. *Nature.* 2011; 472:90–4. [PubMed: 21399628]
12. Stevens JB, Abdallah BY, Liu G, et al. Heterogeneity of cell death. *Cytogenet Genome Res.* 2013; 139:164–73. [PubMed: 23548436]
13. Stevens JB, Abdallah BY, Horne SD, et al. Chromosomal instability and transcriptome dynamics in cancer. *Cancer Met Rev.* 2013 Epub ahead of print.
14. Ye CJ, Stevens JB, Liu G, et al. Genome based cell population heterogeneity promotes tumorigenicity: the evolutionary mechanism of cancer. *J Cell Physiol.* 2009; 219:288–300. [PubMed: 19115235]
15. Heng HH, Liu G, Stevens JB, et al. Decoding the genome beyond sequencing: the new phase of genomic research. *Genomics.* 2011; 98:242–52. [PubMed: 21640814]
16. Heng HH. The genome-centric concept: resynthesis of evolutionary theory. *Bioessays.* 2009; 31:512–25. [PubMed: 19334004]
17. Heng HH, Stevens JB, Bremer SW, et al. Evolutionary mechanisms and diversity in cancer. *Adv Cancer Res.* 2011; 112:217–53. [PubMed: 21925306]
18. Heng HH, Liu G, Stevens JB, et al. Genetic and epigenetic heterogeneity in cancer: the ultimate challenge for drug therapy. *Curr Drug Targets.* 2010; 11:1304–16. [PubMed: 20840073]
19. Heng HH, Stevens JB, Bremer SW, et al. The evolutionary mechanism of cancer. *J Cell Biochem.* 2010; 109:1072–84. [PubMed: 20213744]
20. Kaiser J. Cancer genetics with an edge. *Science.* 2012; 337:282–84. [PubMed: 22822123]

21. Zhu J, Pavelka N, Bradford WD, et al. Karyotypic determinants of chromosome instability in aneuploid budding yeast. *PLoS Genet.* 2012; 8:e1002719. [PubMed: 22615582]
22. Pavelka N, Rancati G, Zhu J, et al. Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. *Nature.* 2010; 468:321–25. [PubMed: 20962780]
23. Blount ZD, Barrick JE, Davidson CJ, et al. Genomic analysis of a key innovation in an experimental *Escherichia coli* population. *Nature.* 2012; 489:513–8. [PubMed: 22992527]
24. Gao C, Furge K, Koeman J, et al. Chromosome instability, chromosome transcriptome, and clonal evolution of tumor cell populations. *Proc Natl Acad Sci USA.* 2007; 104:8995–9000. [PubMed: 17517657]
25. Upender MB, Habermann JK, McShane LM, et al. Chromosome transfer induced aneuploidy results in complex dysregulation of the cellular transcriptome in immortalized and cancer cells. *Cancer Res.* 2004; 64:6941–9. [PubMed: 15466185]
26. Wang RT, Ahn S, Park CC, et al. Effects of genome-wide copy number variation on expression in mammalian cells. *BMC Genomics.* 2011; 12:562. [PubMed: 22085887]
27. Pollack JR, Sørlie T, Perou CM, et al. Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. *Proc Natl Acad Sci U S A.* 2002; 99:12963–68. [PubMed: 12297621]
28. Bischoff FZ, Yim SO, Pathak S, et al. Spontaneous abnormalities in normal fibroblasts from patients with Li-Fraumeni cancer syndrome: aneuploidy and immortalization. *Cancer Res.* 1990; 50:7979–84. [PubMed: 2253239]
29. Heng HH, Stevens JB, Lawrenson L, et al. Patterns of genome dynamics and cancer evolution. *Cell Oncol.* 2008; 30:513–4. [PubMed: 18936532]
30. Galipeau PC, Li X, Blount PL, et al. NSAIDs modulate CDKN2A, TP53, and DNA content risk for progression to esophageal adenocarcinoma. *PLoS Med.* 2007; 4:e67. [PubMed: 17326708]
31. Fridman AL, Tang L, Kulaeva OI, et al. Expression profiling identifies three pathways altered in cellular immortalization: interferon, cell cycle, and cytoskeleton. *J Gerontol A Biol Sci Med Sci.* 2006; 61:879–89. [PubMed: 16960018]
32. Cheadle C, Vawter MP, Freed WJ, et al. Analysis of microarray data using Z score transformation. *J Mol Diagn.* 2003; 5:73–81. [PubMed: 12707371]
33. Gleichmann M, Zhang Y, Wood WH III, et al. Molecular changes in brain aging and Alzheimer's disease are mirrored in experimentally silenced cortical neuron networks. *Neurobiol Aging.* 2012; 33:205 e1–18. [PubMed: 20947216]
34. Storey JD. A direct approach to false discovery rates. *J R Stat Soc B.* 2001; 64:479–98.
35. Kornblau SM, Tibes R, Qiu YH, et al. Functional proteomic profiling of AML predicts response and survival. *Blood.* 2009; 113:154–64. [PubMed: 18840713]
36. Fridman AL, Tainsky MA. Critical pathways in cellular senescence and immortalization revealed by gene expression profiling. *Oncogene.* 2008; 27:5975–87. [PubMed: 18711403]
37. Heng HH, Bremer SW, Stevens JB, et al. Genetic and epigenetic heterogeneity in cancer: a genome-centric perspective. *J Cell Phys.* 2009; 220:538–47.
38. Venet D, Dumont JE, Detours V. Most random gene expression signatures are significantly associated with breast cancer outcome. *PLoS Comput Biol.* 2011; 7:e1002240. [PubMed: 22028643]
39. York H, Kornblau SM, Qutub AA. Network analysis of reverse phase protein expression data: characterizing protein signatures in acute myeloid leukemia cytogenetic categories t(8;21) and inv(16). *Proteomics.* 2012; 12:2084–93. [PubMed: 22623292]
40. Abdallah BY, Horne SD, Stevens JB, et al. Single cell heterogeneity: why unstable genomes are incompatible with average profiles. *Cell Cycle.* 2013 Epub ahead of print.
41. Stevens JB, Abdallah B, Liu G, et al. Diverse system stresses: common mechanisms of chromosome fragmentation. *Cell Death Dis.* 2011; 2:e178–e86. [PubMed: 21716293]
42. Heng HH. Elimination of altered karyotypes by sexual reproduction preserves species identity. *Genome.* 2007; 50:517–24. [PubMed: 17612621]
43. Gorelick R, Heng HH. Sex reduces genetic variation: a multidisciplinary review. *Evolution.* 2011; 65:1088–98. [PubMed: 21091466]

44. Whibley C, Odell AF, Nedelko T, et al. Wild-type and Hupki (human p53 knock-in) murine embryonic fibroblasts: p53/ARF pathway disruption in spontaneous escape from senescence. *J Biol Chem*. 2010; 285:11326–35. [PubMed: 20118236]
45. Williams BR, Prabhu VR, Hunter KE, et al. Aneuploidy affects proliferation and spontaneous immortalization in mammalian cells. *Science*. 2008; 322:703–9. [PubMed: 18974345]
46. Yasaei H, Gilham E, Pickles JC, et al. Carcinogen-specific mutational and epigenetic alterations in INK4A, INK4B and p53 tumour-suppressor genes drive induced senescence bypass in normal diploid mammalian cells. *Oncogene*. 2013; 32:171–9. [PubMed: 22410783]
47. Kuijjer ML, Rydbeck H, Kresse SH, et al. Identification of osteosarcoma driver genes by integrative analysis of copy number and gene expression data. *Genes Chromosomes Cancer*. 2012; 51:696–706. [PubMed: 22454324]
48. Sheltzer JM, Torres EM, Dunham MJ, et al. Transcriptional consequences of aneuploidy. *Proc Natl Acad Sci U S A*. 2012; 109:12644–49. [PubMed: 22802626]
49. Rancati G, Pavelka N, Fleharty B, et al. Aneuploidy underlies rapid adaptive evolution of yeast cells deprived of a conserved cytokinesis motor. *Cell*. 2008; 135:879–93. [PubMed: 19041751]
50. Heng, HH. Biocomplexity: Challenging reductionism. In: Sturmberg, JP.; Martin, CC., editors. *Handbook on Systems and Complexity in Health*. New York: Springer; 2013. p. 193-208.
51. Horne SD, Stevens JB, Abdallah BY, et al. Why imatinib remains an exception of cancer research. *J Cell Physiol*. 2013; 228:665–70. [PubMed: 23018746]
52. Heng HH, Liu G, Stevens JB, et al. Karyotype heterogeneity and unclassified chromosomal abnormalities. *Cytogenet Genome Res*. 2013; 2:144–57. [PubMed: 23571381]
53. 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 Epub ahead of print.
54. Loven J, Orlando DA, Sigova AA, et al. Revisiting global gene expression analysis. *Cell*. 2012; 151:476–82. [PubMed: 23101621]
55. Heng HH. Missing heritability and stochastic genome alterations. *Nat Rev Genet*. 2010; 11:812–12.
56. Heng HH. Genomics: HeLa genome versus donor's genome. *Nature*. 2013; 501:167–67. [PubMed: 24025829]

Abbreviations

CCA	Clonal chromosomal aberration. Chromosomal aberration shared in two or more cells within a population
CIN	Chromosomal instability
NCCA	Non-clonal chromosome aberration. Chromosomal aberration only found in one cell within a population
QTL	Quantitative trait loci

What's new?

Karyotypic alteration is a driving force of cancer evolution, but its effects on the transcriptome are not well understood. In the present study, the pattern of genome alteration and expression dynamics during somatic evolution was interrogated using an immortalization model. The model revealed that major karyotypic alterations, such as structural or numerical changes, result in significant deviations in gene expression, suggesting that karyotype changes influence gene expression. The data further indicates that karyotype and transcriptome pathways are in constant flux during somatic cellular evolution and that this flux is essential for cancer formation.

Figure 1.

Passages with unstable karyotypes have unstable expression. (a) Schematic of the basic experimental design employed in this study. MDAH-041 cells were cultured and SKYed at indicated time points. At these time points RNA was harvested separately from replicate culture flasks and subjected to microarray analysis (b) Line graph of normalized intensity values throughout passages of each expressed transcript. Here each line represents each of the 22,185 transcripts profiled throughout the time course. Lines are colored in a continuous spectrum from blue (−4 Log10 normalized intensity) to red (4 Log10 normalized intensity) based on expression values at Pd7. A large change in expression patterns is evident between stages where extensive karyotypic change has occurred (Pd7→Pd17, Pd17→Pd 25, and Pd25→Pd54) where as the pattern is much more stable when little karyotypic change takes place (Pd54→Pd109). (c) Principal component analysis (PCA) of Z-score transformed data from samples from various stages of progression. PCA demonstrates that expression patterns of replicates from stable stages of tumor evolution Pd7 (red spheres) and Pd109 (orange spheres), and reference RNA replicates (purple spheres) are closely correlated as evidenced by their proximity while replicates from unstable stages Pd17 (green spheres) and Pd25 (blue spheres) are far apart indicating dissimilarity.

Figure 2.

(a) Karyograph of SKYed cells at Pd7. Each line represents the measured copy number of each noted chromosome in one cell. Aneuploidy is evident in the peaks of chromosomes 1–22 and X that show changes from the normal copy number of 2. Structural aberrations are graphed to the right of the normal chromosomes. Notice that the majority of chromosome translocations that take place are not shared. By Pd7 only 11% of cells contain a normal karyotype. High resolutions of karyographs from all measured stages are presented in Supplemental Figures 1a–e. (b) Average chromosome count per cell at measured passages. There was an overall trend of increasing chromosomal count throughout the process of immortalization, with a slight decrease in chromosome count from Pd17 to Pd25. This could be due to the fact that multiple outlying cells had less than 35 chromosomes. Error bars denote standard error. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Figure 3. Gene pathways that are involved in the immortalization process are in constant flux. Changes in interferon signaling, cell cycle and cytoskeleton have been implicated in spontaneous immortalization of Li-Fraumeni fibroblasts, here by Wikipathway analysis significant changes in these pathways are also found however they do not follow a specific pattern (a). Here significant changes in these pathways are shown in both Pd7 versus all (left) and nearest neighbor comparisons (right). Red boxes indicate significant ($p<0.1$) up-regulation in the passage listed first in each header row relative to the second passage in the header row, while blue indicates down-regulation, and yellow indicates both up- and down-regulation. Similarly, normalized expression levels of oncogenes (b) and tumor suppressor genes (c) do not show a distinct pattern over the immortalization process. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Figure 4.

(a–d) Natural Language Processing (NLP) pathways created in Genespring using genes found to be significantly expressed (corrected $p < 0.05$ with a minimum of 1.5 fold change) viewed in ‘neato’ layout with unconnected nodes removed in nearest neighbor comparisons at each stage. The networks and processes that are affected by expression change at each stage of progression indicating that genome change leads to network dynamics. High resolution images can be found in Supporting Information Figures 2a–d. This data demonstrates that there are unique processes and players (genes) interacting to create system function at each point of the time course and that these processes change as karyotype driven somatic evolution occurs. Each shape represents a gene or process that is significantly changed at each stage. Importantly the network of significant change is in constant flux especially when there are dramatic karyotypic changes from stage to stage i.e. during macro-evolution (a–c). In contrast networks do not show much change if the genome is relatively stable as occurs during micro-evolutionary phases (d). (e) Application of methods previously designed to assess technical reproducibility allows assessment of the effects of karyotypic heterogeneity reorganization on the transcriptome. Principle component analysis of expression profiles of MDAH-041 clones selected at various stages. Clones were profiled using SKY immediately prior to isolation of RNA for expression analysis. RNA for the replicates was harvested from the same flask of cells with half of the cells used for one

replicate and half used for the other. Similar experimental setups have been used in the past to assess technical reproducibility of microarray platforms; however in this instance expression heterogeneity resulting from karyotypic heterogeneity was assayed. PCA reveals that replicates from karyotypically homogenous populations (Stable-1) exhibit a more homogeneous expression profile than replicates with inherent or induced instability (Unstable-1 and Unstable-2, and Stable-1 Dox, respectively). For instance, karyotypically homogeneous cells from Stable-1 are more similar to each other than Unstable-1 or Stable-1 Dox. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Figure 5.

Model describing how genome instability during cancer progression affects transcriptome dynamics. Cancer evolution is driven by stochastic karyotypic alteration (a). Throughout cancer evolution chromosomes rearrange over time in two distinct phases of evolution, the punctuated and stepwise phases ((a) middle and right respectively). During this process interaction networks are altered as the genomic topology is altered (b) Methods that average the cellular population are deceptive as they conceal heterogeneity (c). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Table 1
Overview of karyotypic changes that occur during immortalization

Passage	N	Average chromosome count	Range	Total CCAs	Total NCCAs	Average CCAs per cell	Range	Average NCCAs per cell	Range
Pd7	37	59.06	35–156	10	31	1.33	0–5	0.84	0–7
Pd17	29	73.14	47–131	41	226	2.92	0–9	5.95	2–15
Pd25	48	69.22	25–226	34	101	4.35	0–8	2.1	0–7
Pd54	35	78.00	56–136	9	21	7.23	3–8	0.6	0–4
Pd109	20	84.50	48–123	11	25	6.8	4–9	1.25	0–3