1-1-2016

Fuzzy Unheritance: A Novel Form Of Somatic Cell Inheritance That Regulates Cell Population Heterogeneity

Batoul Abdallah
Wayne State University,

FUZZY INHERITANCE: A NOVEL FORM OF SOMATIC CELL INHERITANCE THAT REGULATES CELL POPULATION HETEROGENEITY

by

BATOUL Y. ABDALLAH

DISSERTATION

Submitted to the graduate school

of Wayne State University

Detroit, MI

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2016

MAJOR: MOLECULAR BIOLOGY AND GENETICS

Approved By:

_____

Advisor Date

_____

_____

_____

DEDICATION

This work is dedicated to:

Youssef and Leila Abdallah

Ahmad, Yahya, and Dalia Nasser

ACKNOWLEDGEMENTS

I would like to acknowledge the following people, for without them, none of this would have been possible. To my parents, Youssef and Leila Abdallah, there are no words that can express the gratitude I have for you both. You are not only my parents, but my role models and heroes. You have been a constant source of guidance and inspiration my entire life, and I hope that one day I can be one half the people you are. Thank for you everything.

To my husband, Ahmad Nasser: we met during my third year of graduate school, and here we are- four years, one marriage, and two beautiful children later. Without your love, understanding, patience, and unconditional support, I would not be writing this today. Thank you for always building me up and giving me the confidence I need to move forward.

I would like to express my utmost and sincere gratitude to my mentor and advisor, Dr. Henry Heng. Thank you for showing me that some simple questions can lead to loaded answers that fundamentally reshape scientific discourse. To say that I admire your adherence to principle and your constant search for truth would be an understatement. Your passion for science and ability to synthesize information from multiple disciplines have made me not just a better researcher, but a better individual overall.

I would also like to thank my dissertation committee, Dr. Stephen Krawetz, Dr. Francesca Luca, and Dr. Ghassan Saed for their insightful and hard questions, encouragement and support.

To my current and former colleagues in the lab: Dr. Joshua Stevens, Dr. Steve Horne, Guo Liu, Karen Ye, and Sarah Regan, thank you for your training and support. I would also like to thank the Center of Molecular Medicine and Genetics, faculty and staff, and especially Mrs. Suzanne Shaw, for all of their support.

To my sisters: Ghada Abdallah, Mona Fakih, Eman Abdallah, Fatima Abdallah, and Nahida Akkary (who is more of a sister than a niece), and brothers, Mahmoud Abdallah and Mahde Abdallah: you are all among the most remarkable and inspiring people in my life. Every single one of you has encouraged me in some form or another to be the best version of myself. I have the utmost admiration and respect for all of you.

To my best of friends: Salma Kobeissi, Fatima Musheinesh, Ammerah Saidi, Eftikhar Saleh, Jacqueline Saleh. You are among the best people that I know and I am truly blessed to call each of you my friend. You have become family. I have known all of you for as long as I can remember, thank you for your understanding, and constant encouragement over the years.

And last, but certainly not least, to my children, Yahya "Yayo" and Dalia, and my nieces and nephews, Raef Akkary, Maya Fakih, Leila Fakih, Muhammad Akkary, and Amal Fakih: you are all wonderful and adorable distractions. Thank you for filling my life with constant happiness and laughter.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Cancer is the second leading cause of death worldwide and in the United States (American Cancer Society 2016). In America, it is estimated that there will be approximately 600,000 cancer-related deaths and 1.6 million newly diagnosed patients in 2016 (AmericanCancerSociety 2016). Specifically, ovarian cancer is the leading gynecological malignancy in the United States, with an approximately 22,000 new diagnoses and 14,000 deaths every year. Ovarian cancer is a disease of highly complex origins. While many genetic/nongenetic risk factors have been associated with cancer, the process of initiation and progression, and specifically ovarian cancer, is not well understood.

A number of studies demonstrate that, similar to other cancers, multi-level heterogeneity is a key feature of ovarian cancer (Bayani et al. 2002, Bayani et al. 2008, Dubeau 2008, TCGA 2011). While much progress has been made towards understanding some aspects of ovarian cancer, progress has been slow as the 5-year survival rates have not changed significantly (Chan et al. 2008, Siegel et al. 2012). A number of mutated and differentially expressed genes have been identified, yet these genes are only found in a small proportion of the patient population, where additionally, a large degree of heterogeneity exists in the mutation profiles and gene signatures of different individuals (Sieben et al. 2004, Gevaert et al. 2008, Konstantinopoulos et al. 2008, Cooke et al. 2010, TCGA 2011, Vereczkey et al. 2011). Interestingly, as newer technologies and techniques allow for the very specific molecular characterization of ovarian cancer, the complexity of the disease only increases (Heng et al. 2011b). The identification of a common pattern of genetic change is highly sought after in all cancer studies. However, while aiming for this goal, the large degree of heterogeneity, specifically genome-level heterogeneity (a characteristic feature of most cancers) is often overlooked (Heng 2009).

The dismissal of multi-level heterogeneity is largely due to the conceptual and research methodologies that are used to characterize cancer cell populations. For example, popular methods, such as genome sequencing, RNA-seq, gene expression profiling, PCR and western blotting are powerful, but limited as they collect an average measurement across an entire population of cells, profiling only more dominant subpopulations and offering little or no information about outlier cells or side populations that may be more important for driving cancer evolution and promoting drug resistance. Furthermore, many of the aforementioned methods have been widely used to establish the current dominant conceptual framework of nearly all cancer-based research. In other words, most current cancer models depict the "average cancer cell." However, average-based measurements do not accurately assess the genetic/nongenetic complexity of cancer cell populations, as they mask variation and ignore heterogeneity and single outlier cells that may be more important for cancer evolution. Furthermore, as genomic clonality is assumed, genome-level heterogeneity and the role of outlier cells have not been properly incorporated into the conceptual framework of cancer evolution, as they are disregarded or not measured, due to the focus on clonal changes. This signifies a need to re-evaluate the current understanding of somatic cell and cancer evolution, and the role of outlier cells in promoting the cancer evolution process.

Recent large scale genome sequencing and other 'omic profiling studies have forcefully brought heterogeneity into the spotlight (Horne et al. 2015a, Heng et al. 2016). While heterogeneity is recognized as the key challenge in cancer research and cancer therapeutics, the mechanism of how it is generated and maintained is not well understood. According to conventional views of somatic cell genetic inheritance, the passing of inheritance is precise and errors that lead to genetic variation are low. However, this conceptual framework of inheritance greatly contradicts the data generated from large-scale sequencing studies that revealed a large

degree of heterogeneity, signifying a need to re-examine the process of inheritance in somatic cells. In this dissertation, a novel concept termed fuzzy inheritance is presented that provides a mechanism for the regulation and maintenance of cell population heterogeneity. Fuzzy inheritance is defined as a mechanism for cellular adaptation that regulates and generates cell population heterogeneity, where a given degree of heterogeneity is passed from mother cells to daughter cells. According to fuzzy inheritance, every somatic cell population exhibits a given degree of heterogeneity, and future cell populations can inherit that same degree of heterogeneity. Fuzzy inheritance is a novel concept, as it provides a mechanism for the regulation and maintenance of cell population heterogeneity, and especially cancer cell heterogeneity.

Brief introduction of clonal evolution in normal and cancer cells

Cancer is a disease of somatic cell evolution, as normal cells undergo a large degree of genetic or genomic change which lead to cancer. This process of genetic change has been compared to organismal evolution (Heng et al. 2006c, Merlo et al. 2006). Both organismal evolution and cancer evolution exhibit the basic principles of natural selection: genetic variation, competition, and inheritance (Merlo et al. 2006, Heng et al. 2010, Gillies et al. 2012). Many different conceptual models have been proposed to characterize cancer evolution. Of the multiple models that exist, clonal evolution has remained the most dominant model used by cancer researchers and clinicians.

The process of clonal expansion is relatively straightforward. Clonal evolution is the dominant form of somatic cell evolution for both normal cells and cancer cells. In normal cell populations, mitosis occurs with high fidelity, where a mother cell divides into two nearly identical daughter cells. During cell division, errors generated at low frequencies are accumulated as they are passed to future daughter cell populations. The average genomic mutation rate of normal

somatic cell division has been estimated to be $11.63x10^{-9}$ (Lynch 2009). Given this low degree of change, it would thus take a long time to accumulate a large degree of genetic change.

Clonal evolution of cancer was first proposed by Peter Nowell in 1976 (Nowell 1976). The process of mutation and selection are the same in cancer cells, however in cancer, genetic/nongenetic changes are reported to occur in specific oncogenes or tumor suppressor genes. Those changes that exhibit some phenotypic advantage consistent with cancer hallmarks will become dominant and are passed down (Hanahan and Weinberg 2011). Consequently, one or few cells will have accumulated enough genetic change and clonally expand into a tumor. The end product is a tumor composed of cells that theoretically can be traced to the founder clone. Recent studies have demonstrated that as few as 3 sequential driver mutations are necessary for tumor formation in some cancer types, like lung and colon adenomacarcinomas (Tomasetti et al. 2015). Previous studies have estimated a minimum of 6 or 7 mutations are necessary for tumor formation in other cancers (Nordling 1953).

Clonal evolution was confirmed in various experimental models. Of note, the Vogelstein group was the first to find a genetic basis for colorectal cancer (Fearon and Vogelstein 1990). This involved the accumulation, in a specific order, of mutations in various oncogenes and tumor suppressor genes, resulting in the formation of an adenoma, followed by a late adenoma, and finally a carcinoma. The adenoma-carcinoma sequence in colorectal cancers served as a general model for solid tumor formation and charged researchers to search for similar accumulated gene mutations in other cancer types.

An example of clonal evolution was observed in chronic mylogenous leukemia (CML), where a translocation between chromosomes 9 and 22 detected in a large number of patients led to the formation of a new gene known as BCR-ABL (Horne et al. 2013). Targeted therapy of this

particular gene fusion as an initial therapy for chronic phase patients has largely been successful, as 5-year survival is 89% (Druker et al. 2006). The success of CML combined with the work of the Vogelstein group prompted for the search for other highly penetrant gene mutations, clonal chromosomal aberrations and other genetic anomalies in nearly all other cancer types (Horne et al. 2013).

Challenges in clonal evolution theory

While cancers that are more stable, like CML, exhibit clonal aberrations, it has been rather difficult to identify clonal genetic/genomic aberrations in other cancer types, especially in solid tumors. Over 61,000 clonal chromosomal aberrations (CCAs) have been identified across nearly all cancer types (Duesberg and McCormack 2013) An even larger number of nonclonal genetic/genomic aberrations, particularly at the karyotype level, have been observed across nearly all cancer types, but these nonclonal changes remain unreported (Heng et al. 2006b, Heng et al. 2006c, Heng et al. 2013b, Heng 2016). A number of gene mutations or chromosome abnormalities have been identified, but their presence in the general patient population is extremely limited (Wood et al. 2007, Vogelstein et al. 2013). Therefore, the success enjoyed by the clonal cancers like CML remains an exception.

More recent initiatives to identify patterns of clonal evolution in cancer involve the genome sequencing of thousands of cancer genomes and multi-region sequencing of tumors (Collins and Barker 2007, TCGA 2011, Gerlinger et al. 2012). These studies overwhelmingly confirmed that heterogeneity is a dominant and key feature of nearly all cancer types and that precise, clonal change was minimal. This will be discussed in more detail in a later section in Chapter 1, entitled Multi-level heterogeneity in ovarian cancer.

Multi-region sequencing represents another experimental platform to detect clonal evolution. Two studies completed in clear cell renal carcinoma (Gerlinger et al. 2012, Gerlinger et al. 2014) and one in high-grade serous (HGS) ovarian cancer (Bashashati et al. 2013) performed multi-region exome sequencing of single tumors and/or metastatic sites to determine the genetic evolutionary history of each cancer. Phylogenetic analysis was completed to temporally identify truncal or conserved mutations. Interestingly, in both clear cell renal carcinoma studies, only one truncal gene mutation was detected in the VHL gene. No other gene mutations were identified as truncal, as most mutations were not detected across all or most regions. Most driver mutations were located in spatially separated subclones, and the number of drive mutations increased with every region sequenced. Similar results were observed in the high-grade serous ovarian cancer study, where each sample was genetically and spatially separated from others. With the exception of a mutation in TP53, most mutations were not shared across each sample. Furthermore, driver mutations in all studies were detected later in the cancers history as opposed to being early driving events. In both cancer types, an ancestral clone may have exhibited the one truncal or common mutation (VHL in CCRC; TP53 in HGS), however most other drivers were not detected in a sequential or linear fashion as theoretical clonal evolution would predict. Altogether these data demonstrate that neither linear clonal or branched clonal evolution does not occur in the clinical situation. The data further demonstrates that heterogeneity or noise is an overwhelmingly dominant feature of most cancer types. Most importantly, multi-region sequencing demonstrates that even precise changes are not passed over a number of cell divisions, as each tumor region was spatially disconnected from its neighboring region.

Conceptual limitations of clonal evolution

Fundamental conceptual gaps exist in the current study of cancer evolution. While clonal evolution is the dominant conceptual model used in cancer research, it poses significant challenges in furthering the understanding of the disease. Multi-level heterogeneity is a key feature of cancer evolution, and is widely acknowledged as the foremost challenge for current cancer research (Heppner 1984, Heng et al. 2009), however it is consistently unaccounted for in the most common conceptual models and experimental approaches. A second major challenge is the disregard of stochastic variation, particularly at the genome level. The main experimental framework over the last several decades has directed the field to search for the molecular magic bullet, a single or collection of mutations that can be targeted to cure cancer. However, these reductionist approaches that describe cancer evolution through a series of sequential and accumulated mutations ignore the overwhelming complexity that is a characteristic feature of cancer. Because cancer is a genome-defined systems disease, the tracing of individual gene mutations may not provide further insight into understanding the disease. These genes may be highly significant under defined laboratory conditions, but their role in natural settings are often more difficult to characterize due to the high level of genome heterogeneity. Natural settings are governed by a large degree of stochasticity, where it becomes difficult to detect cause and effect relationships for a particular gene in cancer (Capp 2005, Heng et al. 2011a, Heng et al. 2013a). Genome-level stochasticity is a dominant feature of cancer evolution, and it minimizes the importance of identifying single gene or linear changes, and it makes large scale genome change the driving force of cancer evolution (Heng et al. 2006c, Heng 2009, Heng et al. 2010, Heng et al. 2011a, Heng et al. 2013a, Liu et al. 2014). A shift in the conceptual framework that places heterogeneity at the forefront, i.e., genome-level heterogeneity, is crucial for furthering the current understanding of how heterogeneity drives cancer evolution and for understanding the cellular basis of heterogeneity. A greater understanding

of cancer evolution can be attained by monitoring the frequency of clonal and nonclonal change, particularly at the genome level (Heng et al. 2006c, Heng et al. 2011a, Heng et al. 2013a).

Technical limitations of tracing clonal evolution

There are also technical limitations in the validation of clonal evolution conceptual framework. Intra-tumor multi-level heterogeneity is a characteristic feature of most cancers (Heng et al. 2006a, Navin et al. 2011, Wang et al. 2014). However, routine technical methods used to study cancer evolution present significant limitations in the characterization and understanding of cell population dynamics, as they profile the average cell (Pelkmans 2012, Abdallah et al. 2013). As previously mentioned, the average cell does not accurately depict unstable cancer cells.

The use of the average also poses significant challenges to clinical application of cancer research, such as the use of chemotherapy. Average-based technologies and methods of analysis have been used to develop treatment that is specific to an oncogene or particular cancer phenotype, like fast proliferation. These treatments are tailored for a specific feature of cancer in a heterogeneous tumor mass, and will only eliminate specific subpopulations, leaving behind outlier cells or even forming drug resistant cells with altered genomes. Because these cells are so rare, they are not usually detected through average-based methods. Therefore the misrepresentation of cancer cell populations through the exclusion of heterogeneity challenges the validity and utility of most cancer research completed in heterogeneous cancer samples.

Multi-level heterogeneity in ovarian cancer

Like most other cancer types, ovarian cancer exhibits a high degree of multi-level heterogeneity. Ovarian cancer pathogenesis has been widely debated over the last several years due to its complex origins and diverse causes or risk factors (Dubeau 2008, Bast et al. 2009, Birrer 2010). Heterogeneity has been extensively documented at multiple genetic levels including the

gene, protein, epigenetic and genome levels (Bayani et al. 2002, Roschke et al. 2002, Khalique et al. 2007, Bayani et al. 2008). Multiple subtypes of ovarian cancers have been identified through the use of high-throughput technologies. Each subtype exhibits distinct morphological and molecular features (Kobel et al. 2008, Soslow 2008, Tothill et al. 2008).

Despite the stratification of ovarian cancers into these distinct subtypes, heterogeneity remains a challenge as genetic markers used to distinguish subtypes are heterogeneous in penetrance. For example, CA125 is a genetic marker commonly used for the ovarian cancer screening (Bast et al. 1998, Moss et al. 2005). However, its lack of specificity for ovarian cancer and its diverse penetration among ovarian cancer patients makes its use as a biomarker unreliable. Specific subtypes of ovarian cancer exhibit elevated CA125 expression, and its expression is not uniform across all patients. Additionally, elevated CA125 expression has also been observed in ovarian benign lesions, endometriosis and other non-malignant gynecological conditions. Finally, levels of CA125 expression have also been reported to rise during menstruation and pregnancy (Meden and Fattahi-Meibodi 1998, Kafali et al. 2004, Nossov et al. 2008, Ercan et al. 2012). Attempts have been made to combine CA125 expression with other genetic markers or diagnostic techniques. For example, CA125 expression has been combined with HE4 expression to improve the outcome of early screening in an approach known as ROMA. The combined approach has enjoyed some initial success, however there has been no significant improvement to early detection (Jacob et al. 2011, Van Gorp et al. 2011). While aconsiderable degree of overlap between healthy controls, benign masses, and epithelial and non-epithelial ovarian tumors exist, ROMA does not outperform HE4 alone and/or CA125 alone (Cho et al. 2015, Zhang et al. 2015).

Similar to other cancers, ovarian cancer is presented with challenges related to detecting common biomarkers. Genome sequencing and gene expression studies have amassed a large

amount of molecular data, however the clinical application of this data has been difficult to ascertain (Capp 2005, Wood et al. 2007, Heng et al. 2011b, Huang 2012, Roberts et al. 2012, Yates and Campbell 2012). While identified genes can be grouped into clinically relevant pathways, these pathways face the same challenges of limited overlap as the genes found within that pathway. Additionally, cancer cells exploit "pathway switching," where a cancer cell can switch to a different pathway if the dominant pathway has become nonfunctional. Due to the overwhelming karyotypic heterogeneity found in cancer cells, and especially ovarian cancer cells, there is unlimited potential for pathway switching (Stevens et al. 2013, Stevens et al. 2014). In summary: precise genetic changes in most cancers are a rare event; randomness or noise significantly outnumber those few changes that are precise; and most targets continuously change. The following sections present evidence to support these findings at every genetic level.

Genome-level heterogeneity in ovarian cancer

Cytogenetic and genome-level profiling has been completed in ovarian cancer in order to identify markers of genetic/genomic change that are related to ovarian cancer pathogenesis, disease progression, or drug resistance. Chromosome change is widespread in all subtypes of ovarian cancer, and especially high grade serous ovarian cancers (Bayani et al. 2002, Rao et al. 2002, Hoglund et al. 2003, Bayani et al. 2008). Many incidences of numerical chromosome instability have been reported, including regions in the following chromosomes: -3, -4, -6, -8, -11, -13, -15, -17, -18, -22, -X, +1, +2, +3, +6, +7, +8, +9, +12, and +20 (Pejovic et al. 1992, Taetle et al. 1999, Bayani et al. 2002, Jin et al. 2004). Many regions of structural breakpoints have also been identified, including regions in the following chromosomes: 1p, 1q, 2q, 3p, 3q, 5p, 5q, 6p, 6q, 7p, 7q, 8p, 8q, 11p, 11q, 12q, 13p, 13q, 17q, 19p, and 21p (Pejovic et al. 1992, Taetle et al. 1999, Bayani et al. 2002, Jin et al. 2004). Widespread chromosome change is also observed in ovarian

cancers that have mutations in TP53 and BRCA1/2, and in spontaneous ovarian cancers (Koul et al. 2000, Lawrenson 2010). In all of the studies analyzed, there has been no single or collection of chromosomal aberrations that is common to all ovarian cancer patients, regardless of the subtype analyzed.

While many widespread changes have been identified, these incidents include only those aberrations that are clonal. A significantly larger number of nonclonal structural and numerical changes remain unreported or ignored due to the intense focus of identifying clonal change. Despite the many studies devoted to identifying specific chromosome regions that are commonly deregulated in ovarian cancer, the only common feature that links all ovarian cancer is the large degree of genome-level heterogeneity. In other words, the data seems to indicate that there is significantly more noise than any precise chromosomal change.

Whole genome sequencing

The Cancer Genome Atlas (TCGA) is a large-scale cancer genome sequencing (and other omic) initiative that sought to identify all genetic/genomic variants involved in cancer evolution by sequencing a large enough sample to capture all or most mutated genes involved. Prior to the TCGA, a number of mutated genes had been identified but were present at low frequencies among patients. Exome sequencing was completed on a pilot study which included the sequencing of 316 high grade serous ovarian cancers (TCGA 2011). Overall, the results of the study confirmed the large degree of multi-level heterogeneity, especially gene mutation heterogeneity, observed in high grade serous ovarian cancer. Specifically, TP53 was mutated in over 90% of the samples. Other statistically significant mutations included: BRCA1, BRCA2, NF1, RB1, CSMD3, GABRA6, FAT3, and CDK 12 were found in 2%-9% of samples. A number of statistically significant mutations were identified, however the sequencing of a larger number of tumors may not bring

new insight into understanding ovarian cancer evolution due to the stochastic nature of cancer evolution (Heng 2007a). Against the initial expectations of many researchers, the TCGA did not identify previously reported gene mutations at high significance. For example, individuals with mutations in RAD51 have an increased susceptibility to developing cancer, yet it was only mutated in one single patient. Additionally, while they have not been previously linked to ovarian cancer, the TCGA study also identified novel genes CSMD3, GABRA6, and FAT3 related to ovarian cancer as highly significant. The rise of large-scale sequencing studies has significantly increased the number of statistically significant but clinically irrelevant gene mutations, or gene mutations that are so infrequent, they do not warrant clinical studies. Many of these genes are designated as passenger mutations because they do not seem to alter the fitness of the cancer and are thus ignored. Furthermore, many mutated genes identified from sequencing studies are members of clinically important pathways. More recent initiatives have been aimed at targeting these more common pathways. While this approach may seem reasonable, it does not address the fact that cancers are highly evolvable and readily undergo pathway switching during the evolutionary process (Stevens et al. 2013, Stevens et al. 2014). The number of and specific mutated genes will only increase due to pathway switching. Any further large scale sequencing initiatives in cancer will only result in the addition of more cancer genes. This observation was first made nearly one decade ago (Heng 2007a) and only recently has there been some discussion over the clinical utility of the sequencing approach (Kaiser 2012, Yates and Campbell 2012, Watson 2013).

Interestingly, the results obtained by the TCGA ovarian cancer sequencing project demonstrate that gene mutation profiles are highly heterogeneous and discontinuous between different subtypes. This indicates most ovarian cancer patients do not exhibit the same accumulated mutations as would be expected under stepwise, clonal evolution. Rather, the data

that revealed a large degree of heterogeneity suggests that cancer evolves through punctuated evolution. This observation indicates that, in contrary to stepwise clonal evolution, gene mutations found in early stages are commonly not detected at later stages (Heng 2007a, Heng et al. 2010). It is known that mutations in TP53 are highly penetrant in high grade serous ovarian cancers, but it is not readily detectable in other subtypes. Increased attention to the link between TP53 mutations and chromosome instability (CIN) is needed due to the integral role CIN plays in cancer evolution (Heng et al. 2013a). The relationship between TP53 and CIN is more likely a result of genome destabilization as opposed to oncogene-mediated molecular pathways.

Gene expression profiling

In addition to genome sequencing, whole genome microarray studies have been performed in order to understand aberrant gene expression patterns in ovarian cancer samples, specifically of the high grade serous subtype. Gene expression profiling is a commonly used high throughput method to identify aberrant gene expression for cancer evolution, resistance, and disease prognosis. Some initial studies revealed that ovarian cancer is characterized by heterogeneous gene expression patterns, as differentially expressed genes were widely patient-specific (Hough et al. 2000, Bayani et al. 2002, Jazaeri et al. 2002). Subsequent studies demonstrated a number of ovarian cancer subtypes characterized by specific molecular signatures (Welsh et al. 2001, Yousef et al. 2003, Bast et al. 2009). While these studies provide a wealth of information regarding ovarian cancer gene expression, the gene expression profiles varied widely across patient samples, suggesting that gene expression is patient and tumor specific.

For example, prognostic gene signatures are an application of gene expression profiling used to determine gene sets that confer patient prognosis. However, similar to genome sequencing, prognostic gene signatures also exhibit considerable heterogeneity. While some early studies may

have demonstrated some promising results, additional studies have demonstrated the limited ability of gene prognostic signatures to be validated by independent studies (Konstantinopoulos et al. 2008). More recent attempts to improve the prognostic capability of gene signatures involved the integration of multiple datasets to generate and validate signatures for a specific subtype. Separation by subtype was intended to eliminate any heterogeneity observed among gene expression. Despite this effort, considerable heterogeneity among different gene prognosis signatures was still evident. To demonstrate this, a comparison of four different gene prognostic signatures was completed on high grade serous ovarian cancers to identify any common genes (Bonome et al. 2008, Yoshihara et al. 2010, TCGA 2011, Verhaak et al. 2013). Comparison of the studies revealed low overlap of gene signatures, where the maximum overlap between any two studies was less than 30% (Figure 1). Gene overlap between three studies fell drastically to 3%-8%. Overlap between all studies revealed no common genes. Interestingly, two studies (TCGA, Veerhak) used the same datasets to generate their prognostic signatures. The TCGA used 285 of the 489 samples available to generate a 193 gene signature, while the Veerhak study used the full 489 samples to generate a 100 gene signature. While the two studies used much of the same samples, only 34 genes overlapped between the two studies. This likely occurred because new statistically significant aberrantly expressed genes are introduced with the addition of a single gene expression profile, which suggests that the gene expression behind the prognostic signatures are patient specific. A similar finding was observed when comparing different datasets of miRNA signatures in ovarian cancer (Wan et al. 2014). Interestingly in a breast cancer study, Venet et al produced random gene expression signatures and found over 90% of the randomly generated signatures (over 100 genes/signature) can discriminate between good or bad prognosis more efficiently than the gene signatures obtained from breast cancer patients (Venet et al. 2011).

Epigenetic heterogeneity in ovarian cancer

Many recent efforts have intensely focused on epigenetic dynamics in ovarian cancer, including methylation profiling, miRNA, shRNA, and histone modifications, with the purpose of identifying specific ovarian cancer markers (TCGA 2011, Almendro et al. 2013, Tang et al. 2015, Bai et al. 2016). Specifically, methylation profiling has been an intense area of research given the new technologies that allow for the whole genome detection of methylation sites at high resolution. Global changes in DNA methylation is a hallmark feature of most cancers, including ovarian cancer. Both hypermethylation and hypomethylation have been observed in ovarian cancer (Widschwendter et al. 2004). Despite the intense research in whole genome methylation profiling in cancer, the results of these studies are similar to genome sequencing studies, as there seems to be little common loci among patients. Therefore, similar to genome sequencing studies, the large degree of methylation heterogeneity is the common factor that links methylation profiling of tumor samples. For example, the TCGA ovarian cancer pilot completed methylation profiling on DNA promoter regions for the 489 high grade serous tumors using the Illumina Beadchip platform, in addition to genome sequencing and gene expression profiling. Compared to fallopian tube controls, 168 genes were hypermethylated. The following genes AMT CCL21, and SPARCL1 were silenced in most of the 489 tumors. Also silenced in a substantial number of tumors was RAB25, which ironically has been previously shown to be overexpressed and amplified (Cheng et al. 2004). Additionally, BRCA1 was epigenetically silenced in 56 of 489 tumors. Comparison of the 168 gene list generated by the TCGA to other methylation profiling studies of high grade serous ovarian cancer using the Illumina platform found that only 2 genes in common, IGF1 and MFAP4 (Yoon et al. 2010). In addition, the TCGA identified SPARCL1 as hypermethylated, while the Yoon study identified its paralog SPARC as hypermethylated. Clearly the identification of

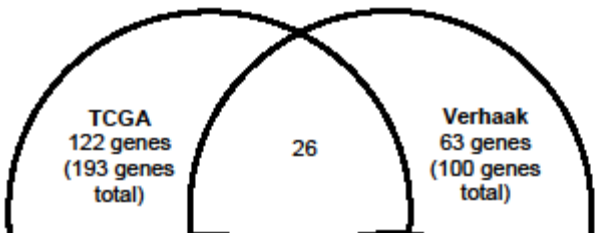
common aberrant methylated loci faces the same challenges as identifying common gene mutations.

Integrated data profiles

Regardless of the molecular genetic level being analyzed, statistically significant but lowly penetrant deregulated loci are prevalent and pose significant challenges for cancer therapy. It has become increasingly apparent that emphasis on a single or handful of genes at a single genetic level will not be sufficient for improved cancer treatment. In order to develop more effective treatment strategies, efforts aimed at integrating information from multiple genetic levels, including gene mutations, copy number alterations, mRNA expression, microRNA profiles, and DNA methylation, have been completed in a number of different cancer types, including ovarian cancer (Mankoo et al. 2011, Weinstein Jn Fau - Collisson et al. 2013, Zhang et al. 2013, Kim et al. 2014).

The 2011 TCGA ovarian cancer pilot was the first study in ovarian cancer to integrate genetic information from multiple genetic levels in order to find commonly deregulated pathways (TCGA 2011). The study found at least one member in the following pathways to be deregulated: RB signaling (67%), PI3K/RAS signaling (45%), NOTCH signaling (22%), homologous recombination (51%), and FOXM1 signaling (84%). Therefore, the integration of multiple genetic levels does result in the occurrence of more commonly deregulated pathways among a larger patient population. While this seems like a promising approach, the integration of data profiles may face some significant challenges, as it does not incorporate stochastic genome level change that is a characteristic feature of cancer.

Figure 1: Limited gene overlap among various prognostic gene signatures



Four studies were completed in HGS ovarian cancer that generated prognostic gene signatures (Bonome et al. 2008, Yoshihara et al. 2010, TCGA 2011, Verhaak et al. 2013). Studies were compared against each other to determine the number of common genes. Comparison any between two studies greatly reduced the number of common genes, where the number of common genes ranged between 3-37. When comparing any three studies, the number of common genes dropped, ranging from 0-7 genes. No common genes were found in all four studies. *Not pictured: 4 genes were found in common between the TCGA and Yoshihara studies. **ND=not determined.

The TCGA ovarian cancer data has been used by many groups in order to create their own lists of genetic/nongenetic features that detect prognosis or overall survival (Mankoo et al. 2011) or to identify novel subtypes that have some prognostic or survival value (Zhang et al. 2013, Kim et al. 2014). Mankoo et al used a multivariate Cox lasso model and an algorithm predicting median time-to-event to develop two signatures: a 156 feature signature that predicts prognosis, and a 182 feature signature that predicts overall survival. The progression free signature (PFS) was able to significantly discriminate between low-, intermediate-, and high-risk groups. Furthermore, the 156 features of the PFS signature were found to be members of 23 pathways of statistical significance, including Phospholipase c signaling and MAPK signaling. Eighty-five of the 156 features listed in the PFS signature were mRNA entities. Interestingly, when compared to previous gene

expression profiling signatures (Bonome et al. 2008, Tothill et al. 2008, Yoshihara et al. 2010, TCGA 2011, Verhaak et al. 2013), only five genes overlapped with the gene expression signature of the Veerhak study (CHIT1, SDF2L1, PHKA1, CITED2, and SLC37A4), two genes overlapped with the Yoshihara gene expression signature (ID4 and SLC2A11), and two genes were found in common with the original TCGA gene expression signature (ID4 and SLC2A11). Clearly, the problem of limited overlap also exists when comparing integrated data to previous studies, even when the studies use the same datasets. This further highlights that stochasticity and noise are inherently dominant features of cancer, and regardless of how many datasets being processed, common change will always be significantly outnumbered noise.

Zhang et al developed an adaptive clustering algorithm followed by an unsupervised super k-means clustering algorithm to stratify mRNA profiles, miRNA profiles, DNA methylation profiles and copy number profiles into 7 novel subtypes of ovarian cancer with significantly different median survival times (Zhang et al. 2013). The subtypes could be distinguished by 36 features, although these features were not made publically available. Of particular importance, subtype 2, designated as a bad prognosis subtype, was enriched for many cancer-related functions such as cell adhesion and angiogenesis, and various EMT and stem cell-related genes. Subtype 2 was also validated by other external datasets, while no mention was made of the validation for the other 6 subtypes. The clustering algorithm identified many molecular features for each subtype, some that were unique to a specific subtype and many other features that overlapped across a number of subtypes. For example, subtype 6 demonstrated an overall good prognosis and exhibited a deletion in chromosome regions 19q13.2-19q13.43 and amplifications in chromosome regions 12p13.33-p11.22 and 20p13-q13.2. Interestingly, subtypes 4 and 5 also exhibited the same amplifications, but both were designated as subtypes with bad prognosis. This suggests that the

clinical and/or prognostic value of specific genes or chromosome regions is not definitive and can change as cancer evolves or with the addition of new data. Integrating the data of additional patients may change the specific molecular features of some subtypes, or alter the subtype classification altogether. Regardless of the volume of data analyzed, the stochastic evolvability of cancers suggests that the integration of data will see limited results.

A final study that utilized integrated data completed by Kim et al generated a classification system that used semi supervised learning to examine the relationship among patients, assuming that patients with more similar data profiles had similar clinical outcomes (Kim et al. 2014). While no signature of features was provided, the semi supervised learning classification system was able to differentiate between short-term vs long-term survival, early stage vs late stage, and low grade vs high grade in ovarian cancer, and short-term vs long-term survival and initial diagnosis vs tumor recurrence in gliobastoma multiforme. Similar to many other studies, integration of genomic profiles demonstrated that while patients may have similar genetic profiles, they have drastically different survival outcomes. Furthermore, integrated data did not significantly increase area under the curve (AUC) values, as compared with individual genetic levels. For example, the ovarian cancer gene expression AUC value that predicted survival was 0.7651, while the AUC value for integrated data was 0.7867. Similarly, the AUC value differentiating early and stage was 0.8767 for copy number alterations alone while the AUC value for the integrated data was 0.8932. In this case, the integrated profile did not significantly add more value than individual genetic signatures.

Taken together, it seems that there are some significant challenges and limitations related to the current efforts integrating data. First, integrated profiles represent a single snapshot of the entire cancer evolutionary process. During punctuated macro-cellular evolution where stochastic genome replacement is the driving force, the molecular profiles will dramatically change due to

pathway switching that gives cancer cells endless avenues to pursue for their continued survival.(Heng et al. 2006c, Stevens et al. 2014) Therefore, this suggests that the clinical relevance of integrated data profiles is limited (Heng 2016, Horne S.D. 2016). This further underscores the fact that despite the large number of data, patients, or information being processed, stochasticity is a dominant feature in cancer and genetic profiles will change throughout the evolutionary process

Second, while each study used the same TCGA ovarian cancer data set, there seems to be limited overlap in feature lists with previous datasets. For example, the mRNA entities of the Mankoo et al PFS 156-feature signature exhibited little overlap with previous gene expression prognostic signatures. The identification of common pathways represents a new strategy to overcome the issue of limited overlap. However, the broad selection of an entire pathway or cellular process is vague or over-generalized and may not provide specific targetable action for treatment purposes. For example, MAPK signaling, TGF-beta signaling, and WNT signaling, among others, were identified in the Zhang et al study as common pathways or networks. The TCGA study also identified a number of pathways that were previously listed. The next logical step after identifying a common pathway is selecting a specific target. The question of which target to choose becomes especially difficult, as specific targets are not common to a larger patient population. The strategy is circular and faces the same challenges as identifying common genes. Furthermore, most of the identified pathways are cellular pathways actively used by normal cells and are not specific to any cancer type. The selection of a pathway that is integral to both normal and cancer cells may result in more negative consequences.

New discoveries in cancer evolution

The large degree of multi-level heterogeneity makes the characterization of all cancers, and specifically ovarian cancers, especially difficult. Of the many types of heterogeneities, genome-level heterogeneity has the greatest impact on cell population dynamics (Heng et al. 2006a, Ye et al. 2009, Heng et al. 2013a, Liu et al. 2014, Stevens et al. 2014). The overwhelming degree of heterogeneity observed in cancer, and in particular ovarian cancer, cannot be explained by Darwinian evolution or models of clonal expansion (Heng et al. 2013a, Ling et al. 2015). Therefore, a systematic reevaluation of cancer evolution is necessary to understand how the large degree of multi-level heterogeneity is maintained. The Genome Theory incorporates heterogeneity as the essential feature to explain what drives organismal and somatic cell evolution (Heng et al. 2006a, Heng et al. 2006c, Heng et al. 2010, Heng et al. 2011a, Heng et al. 2013a). According to the Genome Theory, a cellular system is defined by the genome, and large-scale genome-level change leads to the creation of new genome systems, as opposed to the accumulation of individual gene mutations. Somatic cell evolution occurs through two distinct phases, which will be described in detail in a later section in Chapter 1.

Another recent development in the theory of the evolution of cancer involves the discovery of genome chaos (Liu et al. 2014). Genome chaos is a rapid mechanism of genome re-organization where, after a stress event, chromosomes become fragmented and rejoin to form chaotic genomes. The process of fragmenting and rejoining continues until more stable karyotypes are established. Genome chaos represents a highly significant mechanism of punctuated evolution during crisis, as it allows for the rapid formation of new genomes. Subtypes of genome chaos, including chromothripsis or chromoplexy, have been identified by a number of sequencing studies (Stephens et al. 2011, Baca et al. 2013, Horne et al. 2014).

System inheritance

A crucial component of Genome Theory involves understanding the utmost importance and function of the genome. According to the genome theory, the genome is the highest and fundamental level of genetic organization, and genome-defined system inheritance represents the genetic blueprint of a biological system (Heng et al. 2006a, Heng et al. 2006c, Heng et al. 2011a, Heng et al. 2011b, Heng et al. 2013a). Several pieces of evidence support this claim. First, genes represent the parts of a biological system (for example, a specific protein). The genome, represented by a given karyotype, represents the whole system and provides a 3-dimensional structure that defines the genetic interactions and other emergent biological properties of the cellular system (Heng 2009). While the gene sequence is responsible for producing the tools used in the system, and tools can be modified, they are reliant on a given 3-dimensional context defined by the genome that provides the basis for genetic networks. This is called the genome context. Furthermore, different layers of genetic information are unique in their methods of handling genetic information. DNA is responsible for the inheritance of the parts, while the genome is responsible for system inheritance. Second, large-scale genome change results in a newly defined system. Translocations and aneuploidy have been shown to drastically alter the genetic networks and phenotypic properties of a cell population (Harewood et al. 2010, Stevens et al. 2014). Third, sexual reproduction ensures the preservation of a specific genome while asexual reproduction actually promotes genomic diversity (Heng 2007b, Gorelick and Heng 2010). Embryos with chromosome abnormalities almost always result in termination. Those embryos that do exhibit chromosome change are either sterile (for example, a mule), do not survive to a reproductive age. Fourth, speciation is associated with genome change (Heng et al. 2013a). While different species have many similarities in gene sequences, the genomes are unique. Finally, genome replacement

is the force that drives cancer evolution. Different stages of cancer progression are driven by large scale genome changes. This will be described in more detail next.

The two evolutionary phases of cancer evolution

The Genome Theory was established using an *in vitro* model of cellular immortalization in spontaneously transformed human fibroblasts (Heng et al. 2006c). Li Fraumeni cells were kept in continuous culture conditions, where genomic analysis was conducted at different evolutionary timepoints. At the onset of the experiment, cells were largely genomically stable, or karyotypically clonal. As the cells were kept in continuous culture conditions, internal cellular stress had accumulated and led to the genomic destabilization of a few cells, leading to increased cell population heterogeneity at the genome level, reflected predominantly as karyotype change that is nonclonal. Throughout the duration of the experiment, two forms of chromosomal aberrations were identified, clonal and nonclonal, and appeared in two distinct evolutionary phases. The first phase, called the macrocellular evolutionary punctuated phase, is characterized by a large degree of karyotype-level heterogeneity. Individual cells were karyotypically unique and exhibited nonclonal chromosomal aberrations (NCCAs). Spectral karyotyping of subsequent passage doublings of serially passaged cells demonstrated that cells evolved through punctuated evolution, as the same specific karyotypes were not inherited, and old genomes were replaced with new genomes with each cell division. In other words, there was no stepwise accumulation of genetic change. A selection event occurred at a crisis point, where cells that are genomically stable and exhibit clonal chromosomal aberrations (CCAs) were selected and evolved more slowly. This marked the beginning of the second evolutionary phase. In the micro-evolutionary phase, or stepwise phase, cell population heterogeneity was low at the genome level, reflected by cells that are karyotypically clonal. Evolutionary change was driven by gene-level change, which was

gradual and accumulated over time, and could be traced. The two phases continually cycle, thus reconciling punctuated macro-cellular evolution with stepwise, gradual micro-cellular evolution. The cycling of the two phases is significant because of its implications to cancer evolution and drug resistance. Because karyotypes are not inherited during the punctuated phase and genome replacement is dominant, it is nearly impossible to predict which specific karyotype will become dominant in the stepwise phase. Furthermore, developing chemotherapy for a specific NCCA may not be the best approach for cancer treatment because each tumor represents a single trial of evolution. The punctuated pattern of cancer has been confirmed by multiple sequencing studies (Navin et al. 2011, Baca et al. 2013, Wang et al. 2014).

The general mechanism of Genome Theory is comprised of all individual molecular mechanisms. The key feature that drives evolutionary change is genome replacement, or genome change. Genome change is triggered by accumulated genomic stress. In general, stress that triggers genome change can be endogenous or exogenous, that results in a change in the overall genome status. Factors that trigger stress include gene mutations, deregulated epigenetic change, chemotherapy, changes in temperature, and even continuous cell culture conditions (Heng et al. 2006, Ye et al. 2009, Stevens et al. 2011). Genome-level alterations are a characteristic feature of nearly all ovarian cancers, which includes nonclonal karyotypic change, a crucial component of the evolutionary process of ovarian cancer. Many factors have been linked to ovarian cancer, which include gene mutations (TP53 is mutated in most high grade serous ovarian cancers), chromosome instability, mutator phenotype, epigenetic dysfunction, and oxidative stress. TP53 has been found to be mutated in approximately 93% of high grade serous ovarian cancers (TCGA 2011, Kanchi et al. 2014). Despite its highly mutated status, a mutation in TP53 is not required for primary tumor

growth. Rather, it is more likely that mutated TP53 dynamically works with other stressors to promote genomic instability and tumor growth.

NCCA:CCA cycles and genome instability

Genome (chromosome) instability is a hallmark feature of cancer cells that can be generally defined as a process that involves chromosomal breakage, rearrangement and duplications (Horne 2015c). Additionally, it can also be defined as the acquisition of mutations related to the stability of the genome. Genome instability can occur through a variety of mechanisms: mutations in genes that work to maintain the stability of the genome (for example, TP53 or ATM); drug treatment or exposure to other stressors which cause chromosome breaks and aneuploidy; changes to cell culture conditions, among others (Heng et al. 2006c, Negrini et al. 2010). While many distinct molecular pathways can be deregulated to result in genome instability, they all share one common feature: elevated NCCA frequency at the genome level (Ye et al. 2009, Stevens et al. 2011). In the unstable phase, cells struggle for survival, and so most cells with stochastically generated NCCAs will be eliminated and replaced by cells with new NCCAs. Eventually a cell with an NCCA that is more stable will present itself with a molecular pathway that has a strong growth advantage and will become dominant, thus beginning the stable microcellular Darwinian phase where NCCA frequency is drastically reduced and CCAs become more dominant. Due to their high frequency in unstable cell populations, regardless of the specific molecular mechanism, NCCAs are thus reflective of overall genome instability and can be used to quantitatively measure the instability of a cell population. In cancer, multiple rounds of NCCA:CCA cycles occur, each coinciding with different stages of cancer evolution. NCCAs are conventionally regarded as genomic background. NCCAs have also been linked to transcriptome heterogeneity, growth heterogeneity, drug resistance, and cancer evolution. In contrast CCAs are markers of stability (Heng et al. 2006a,

Heng et al. 2006b, Heng et al. 2006c). NCCAs therefore reflect the status of the genome system and thus can be used as a measure of genome/chromosome instability to monitor cancer evolution. Finally, monitoring NCCA frequency is not for the identification of a specific chromosome aberration, but rather to trace them along with CCAs to understand general tumor dynamics.

Genome instability linked to transcriptome dynamics

Genome instability has significant effects on population dynamics. Specifically, genome instability is reported to impact the transcriptional dynamics of a cell population (Harewood et al. 2010, Lawrenson 2010, Creekmore et al. 2011, Stevens et al. 2013, Stevens et al. 2014). A time-course analysis was completed that investigated the relationship between genome heterogeneity and transcriptome dynamics (Stevens et al. 2014). Spontaneously transformed Li Fraumeni fibroblasts were kept in continuous culture, where spectral karyotyping and gene expression analysis was concurrently completed at different passage doublings that exhibited different degrees of genomic instability. Transcriptome dynamics were found to be heterogeneous in passages with high genomic instability with elevated NCCA frequency. Entire gene expression profiles drastically changed from one passage to the next. In contrast, the transcriptome dynamics of cell populations that were more stable were also more stable, as gene expression profiles largely remained constant and unchanged in successive passage doublings. These data suggest that karyotype-level changes result in global transcriptome change. It also illustrates how, throughout cancer evolution and most significantly punctuated evolution, the constant replacement of new genomes indicates global transcriptome dynamics are always changing. Altogether, this highlights the fluidity of the "cancer genome," as there are no static chromosome changes or transcriptional networks that are permanent when the genome is unstable.

The relationship between genome heterogeneity and transcriptome heterogeneity was further confirmed in two independent studies that used spontaneously transformed mouse ovarian epithelial cells (Lawrenson 2010, Creekmore et al. 2011) In a time-course analysis, gene expression was analyzed at different passages that exhibited elevated genome-level heterogeneity. The global transcription profiles of both studies shared some findings: nearly 600 genes were differentially expressed, and gene expression profiles changed at each passage. Researchers were not able to identify a group of genes that consistently changed throughout each time-.point. Furthermore, only a handful of genes were similarly expressed in more than one stage. This data suggested that genome heterogeneity results in elevated transcriptome dynamics. Analysis of GO categories found no significant overlap between the two studies. This indicates that, not only does genome heterogeneity lead to unstable transcriptome dynamics, but each different experiment represents one evolutionary run, where the transcriptome is unique. The fluctuating transcriptome dynamics in each stage are dependent on the degree of genome change, as well as selective conditions for somatic evolution. Similar to the previous study, if two consecutive time-points were stable and did not change their karyotypes, then the overall transcriptomes would not drastically change either. The findings in these mouse studies are similar to gene expression profiles of cancer patients, as the transcriptomes are patient specific and combining a large number of patients would yield little, if any, common genes.

Inheritance

Genetic inheritance can be defined as the transmission of genetic characters from parents to offspring. The current conceptual framework for genetic inheritance originates from the famous experiments conducted by Gregor Mendel in *Pisum sativum,* the common pea plant. By analyzing thousands of crosses between true-breeding hybrid pea plants for various traits, Mendel made

several conclusions regarding processes of genetic inheritance, which later came to be known as the laws of dominance, segregation, and independent assortment. While his work was initially dismissed when published in 1866, geneticists in the early twentieth century rediscovered Mendel's work and conducted many experiments in various model systems that supported his findings. These works eventually led to the modern synthesis that reconciled Darwinian evolution with Mendelian genetics, and established the current understanding of genetic inheritance and evolutionary biology, which provided the conceptual framework for most genetics research still used today.

Mendelian/classical inheritance/particulate inheritance

According to the conventional framework, genes are the distinct units that define inheritance. Genes are comprised of two alleles located at the same locus on each sister chromosome, which makeup an organisms genotype for a given trait. An individual receives one allele from each parent. Some alleles are dominant while others can be recessive. It is widely understood that a dominant phenotype emerges from when one or both copies of that dominant allele are present, as a dominant allele can mask the phenotype of the recessive allele. Two copies of the recessive allele are necessary for that particular recessive phenotype to emerge. Genes located on different chromosomes are said to be in linkage equilibrium, as they are inherited independently of each other. Genes in linkage disequilibrium are located on the same chromosome and have a given likelihood to be inherited together; the likelihood increases the closer genes are located on the chromosome. While only few traits are single-gene traits, most genes are polygenic and/or multifactorial, in that multiple genetic loci are associated with a single trait and the emergent phenotype is dependent on the environmental interaction with those loci.

Another important feature of inheritance is the manner by which genes evolve. A number of alleles exist for a particular gene at a given frequency in a population. According to the conventional framework, a single allele confers only a small advantage or disadvantage. Natural selection acts on genes and selects those advantageous alleles, and over time advantageous alleles accumulate while disadvantageous alleles are discarded. Other mechanisms by which allele frequencies can change are mutation, genetic drift and migration or gene flow. Speciation occurs when enough advantageous alleles have been accumulated.

According to the above, genetic inheritance may seem somewhat straightforward., However there are a number of factors that makes the pattern of genetic inheritance highly complex. While Mendel's work determined that some traits are discrete and follow a simple dominant/recessive autosomal inheritance pattern, most traits are polygenic, and many genes exhibit continuous phenotypes. For example, various traits of the common pea plant exhibited discrete phenotypes, e.g., peas were either round or wrinkled. Human height is a trait that exhibits a continuous phenotype where the inheritance patterns are more complex. Approximately 80% of human height inheritance is attributed to additive genetic factors (Visscher et al. 2006, Visscher 2008, Wood et al. 2014). Many studies have attempted to identify specific loci associated with human height. The approach has been straightforward: genome sequencing and SNP analysis for a large sample of individuals to find common loci associated with a particular height. Detecting common loci responsible for height has been challenging due to the small effect sizes of identified loci, thus requiring a larger and larger samples (Visscher 2008). Furthermore, in many cases a particular SNP may not be within a particular gene, but in a gene regulatory region. Thousands of variants have been identified from the sampling of over 250,000 individuals, all of which explain nearly 60% of human height variation. It is estimated that thousands more will be detected (Wood

et al. 2014). Adding an additional layer of complexity, phenotypes are heavily dependent on the environmental interaction with genes. This becomes especially complicated when exposure to a particular environment does not always produce the same phenotype, or when different environmental factors interact with different loci to produce the same phenotype. This complex and increasingly confusing pattern of inheritance is a common occurrence in cancer cell inheritance.

Genetic variation in sexual and asexual species

Genetic inheritance for organisms that undergo sexual reproduction is different than inheritance in asexually reproducing species. For sexual reproduction, inheritance begins with reproduction and the formation of gametes, sperm (male) or egg (female). Gametes are haploid cells that result from two meiotic cell divisions in the reproductive organ of a given species. Fertilization occurs when one sperm and one egg fuse to form a fertilized embryo. In most cases, each gamete comes from different organisms, although some organisms can self-fertilize. Gametes must exhibit one copy of each chromosome for the normal embryonic development and reproductive success of the zygote. In most cases, embryos with chromosomal abnormalities are spontaneously aborted. Therefore, the genome for any sexually reproducing individual is one half of each parent, and both parents and reproductively successful offspring must have the same genome system (for example 46 chromosomes for human parent and child). At the genome level, inheritance must be precise in sexual reproduction.

Genetic inheritance is simpler in asexual reproduction. Genetic material is duplicated during the cell cycle in S phase with high precision, although some mistakes may occur at a very low frequency. The 4N cell then undergoes mitosis, and one pair of sister chromosomes segregate to opposite ends of the cell. The end result after cytokinesis is two cells with identical genomes,

more or less. Both cells that are normal (for example, 2N=46) and abnormal, in that they exhibit a chromosomal abnormality, (for example, 2N≠46) undergo the same reproductive cell division process. A normal cell that is genomically stable will pass down a copy of 46 chromosomes to its daughter cell. In other words, a normal cell will make a normal cell. At the same time, an abnormal cell will pass that same abnormal genome to its daughter cell. In an abnormal cell with a de-novo chromosome change, that specific abnormality is passed to its daughter cell. In this instance, it seems that genome inheritance need not be precise for the reproductive success of a cell with an altered genome.

In asexually reproducing species, while the same genome must be passed from parent to offspring, there is still considerable genetic variation DNA-sequence level. Genetic variation can arise from genetic recombination, de novo mutations, random mating, and genetic drift. The mixing of two genomes at the sequence level leads to diversity of traits, ex: hair color, eye color, height, etc, all while the same genome is preserved. According to the current framework, the degree of genetic variation in asexually reproducing species is reportedly less because genetic inheritance is precise. Genetic variation in asexual species originates from de novo mutations and other genetic changes that occur during cell division at low frequencies. Genetic variation can also occur at the genome level, in the form of chromosome change. As stated earlier, the conventional framework dictates that the precise changes are inherited by the next generation of cells. Therefore, in a population of cells arising from one single founder clone, all cells should be identical with some minor variations. Taken together, it is largely believed that genetic variation is minimal in asexual species and significantly higher in organisms that undergo sexual reproduction.

Key challenges to the current paradigm

According to the above, inheritance in somatic cells is precise. A cell will pass an identical genome to its daughter cell. However, the data generated from our cytogenomic studies, as well as the cancer genome sequencing studies reveal an overwhelmingly large degree of multi-level heterogeneity that are not consistent with mechanisms of classical inheritance in somatic cells. The multi-level heterogeneity observed exceeds the heterogeneity expected by Darwinian evolution models (Ling et al. 2015). The data indicates that inheritance processes are not precise for all somatic cells. The following support this: First, NCCAs are frequently observed in somatic cells, especially in populations that are unstable. Second, punctuated macrocellular evolution is well documented in unstable cancer cells as the dominant form of evolution (Heng et al. 2011a, Navin et al. 2011, Baca et al. 2013). Unstable cells inherit new karyotypes after cell division, which directly conflict with precise, classical Darwinian models (Heng et al. 2006c, Voet et al. 2013). Third, genome chaos represents a rapid form of genome restructuring where new and complex karyotypes are inherited at every cell division (Heng et al. 2006c, Liu et al. 2014). Finally, even normal somatic cell populations exhibit substantial genome heterogeneity in the form of somatic mosaicism, again conflicting with the precise inheritance framework (Pack et al. 2005, Iourov et al. 2008, O'Huallachain et al. 2012). This leads to the following questions: By what mechanism are altered karyotypes passed? Currently there are no mechanisms to explain how altered karyotypes are passed from a mother cell to its daughter cell. Given that the passing of altered karyotypes creates cell population heterogeneity in the form of NCCAs, could this mechanism also explain how heterogeneity is created and maintained? Are there specific cells or genetic loci that are tasked with generating heterogeneity, or is this a general population-based mechanism?

Role of the "insignificant" outlier cell

Outlier cells represent an underappreciated feature of somatic cell populations, and especially cancer cell populations. Outlier cells are typically ignored from most analyses, as the primary focus has been placed on identifying genetic drivers that are highly penetrant in a cancer cell population. However, it has been difficult to identify highly penetrant drivers, as a significant majority of genetic/genomic changes found in cancer cells are rare. In a sense, cells that exhibit rare, nonclonal genomic change are outliers, making nearly every unstable cancer cell an outlier. Finally, studies have demonstrated new and important functions for outlier cells. For example, Chang et al (Chang et al. 2008) found that outlier cells can reconstitute cell population heterogeneity of gene expression and cell differentiation. Taken together, this suggests a fundamental role for outlier cells that goes beyond the insignificant bystander, where outlier cells may be the crucial feature that maintains cell population heterogeneity.

Somatic cells inherit heterogeneity

The next logical question to be asked is whether there is a general mechanism that facilitates the passing of altered karyotypes. In unstable cell populations, heterogeneity is generated through the constant formation of NCCAs during punctuated macro-cellular evolution. Here, a daughter cell inherits an altered genome (system inheritance) from its mother cell. By definition, a daughter cell must inherit something from its mother cell. The function of inheritance is to preserve the same genome system. If the same system inheritance is not inherited, then what is inherited? It must be that heterogeneity, or the evolutionary potential to generate altered genome systems, is inherited in the next cellular generation.

This synthesis has led to the following hypothesis:

Unstable variant/outlier cells drive cancer evolution and fuzzy inheritance represents a novel mechanism of passing heterogeneity within cell populations.

The hypothesis was tested and developed through the following chapters:

Chapter 1: Establish Single cell model and methodology

Chapter 2: Demonstrate that outlier cells are dominant

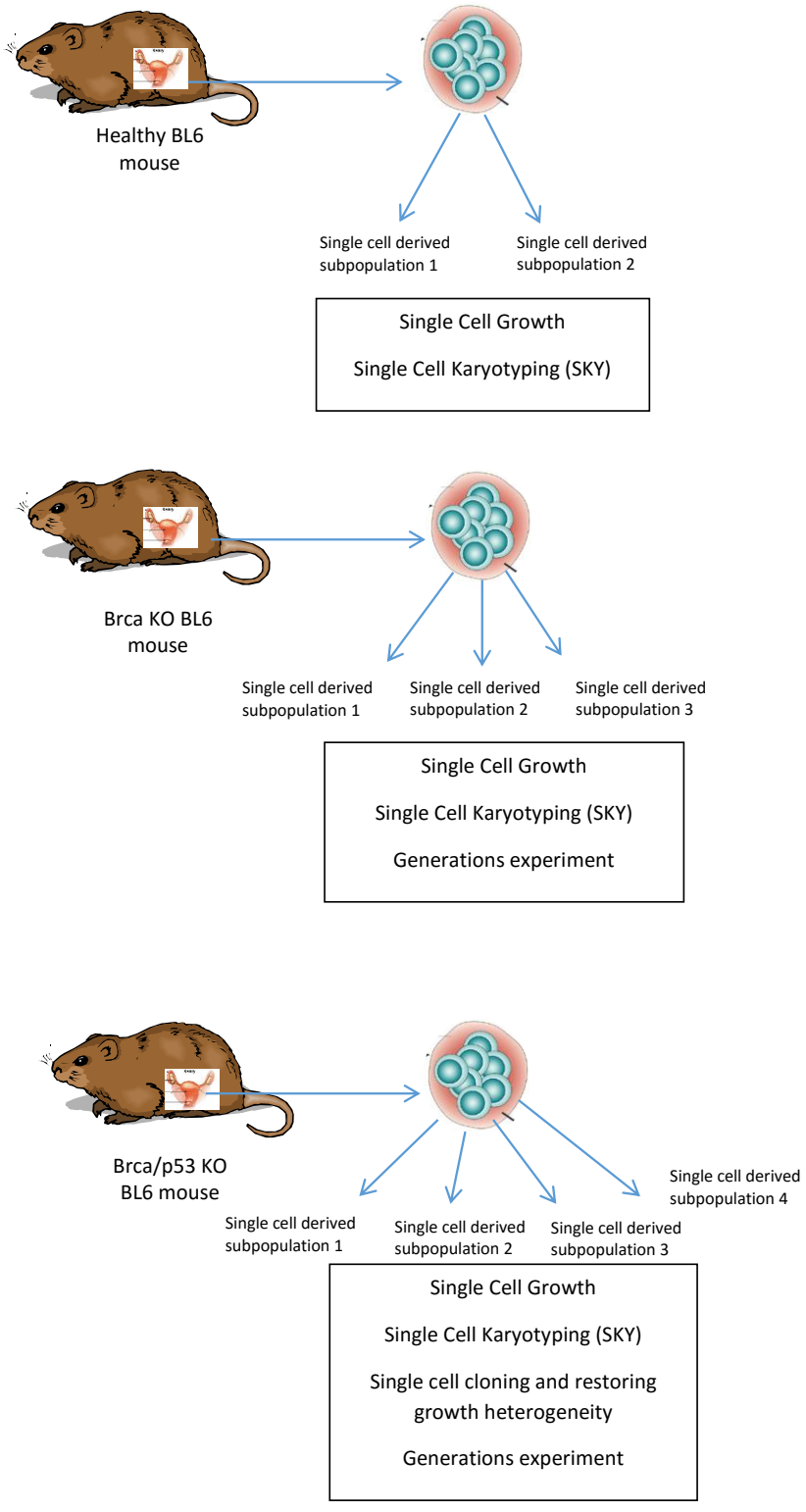
Chapter 3: Characterize mechanism of fuzzy inheritance

To identify if a new type of inheritance exists is tricky, as most research methods are average-based and not sensitive to detect fuzzy inheritance. First, it is necessary to determine if a mother cell population can pass heterogeneity to its daughter cell population (Figure 3A). Next, it is necessary to determine the impact of karyotype heterogeneity on other heritable features of the cell. To do this, karyotype heterogeneity needs to be linked with heterogeneity of other traits, like cell growth (Figure 3B). Next, it is necessary to determine that an isolated single cell can reconstitute cell population heterogeneity of karyotype and cell growth (Figure 3C). Next, it is necessary to determine that genome instability is directly related to the degree of heterogeneity that is passed (Figure 3D). Finally, the long term evolutionary benefit of fuzzy inheritance is demonstrated. While heterogeneity may decrease the overall fitness of the cell population (for example due to cell death, slow growth), the variability of genomes significantly increase the likelihood of survival under stressful conditions (Figure 3E). To conclude, …?

To complete these experiments, a single cell model has been developed from *ex-vivo* mouse ovarian surface epithelial cells originating from C57/BL6 mice that were wild type or conditionally activated for Brca1 and/or p53 (Figure 2). Single cell spectral karyotyping and single cell growth

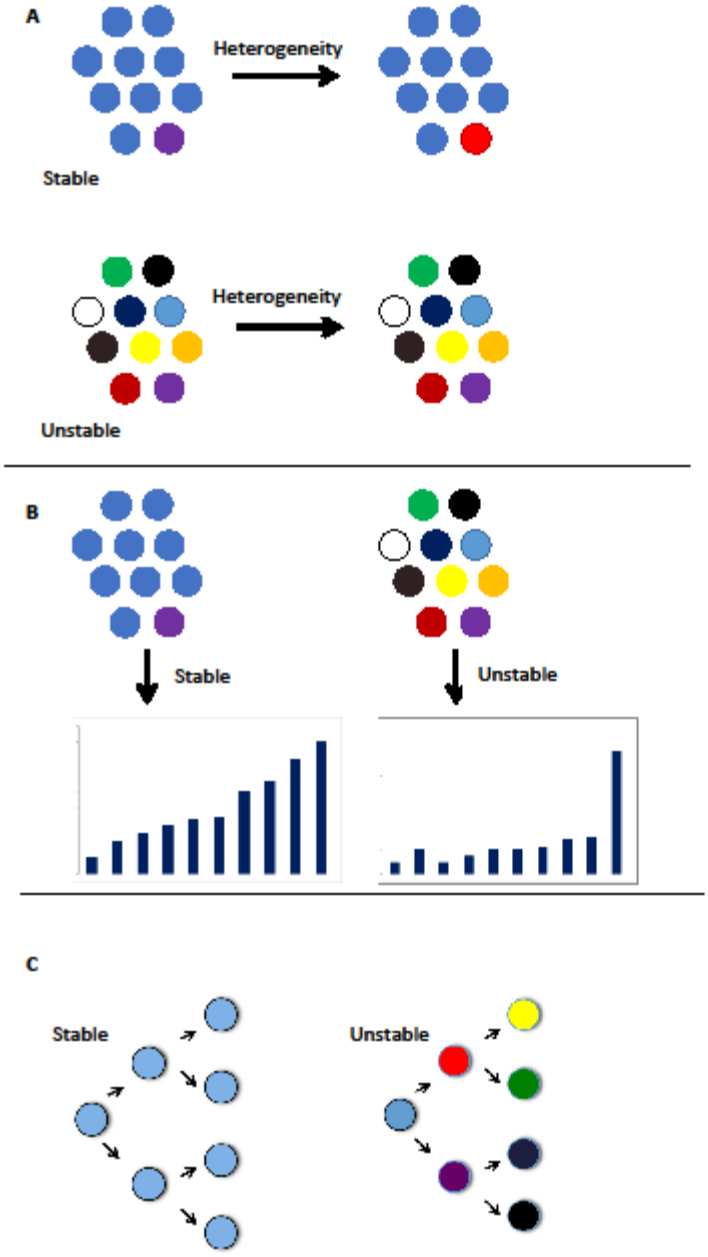
experiments were performed to test the pattern of inherited traits between mother cells and daughter cells.

Figure 2: Mouse ovarian surface epithelial cell-derived single cell model



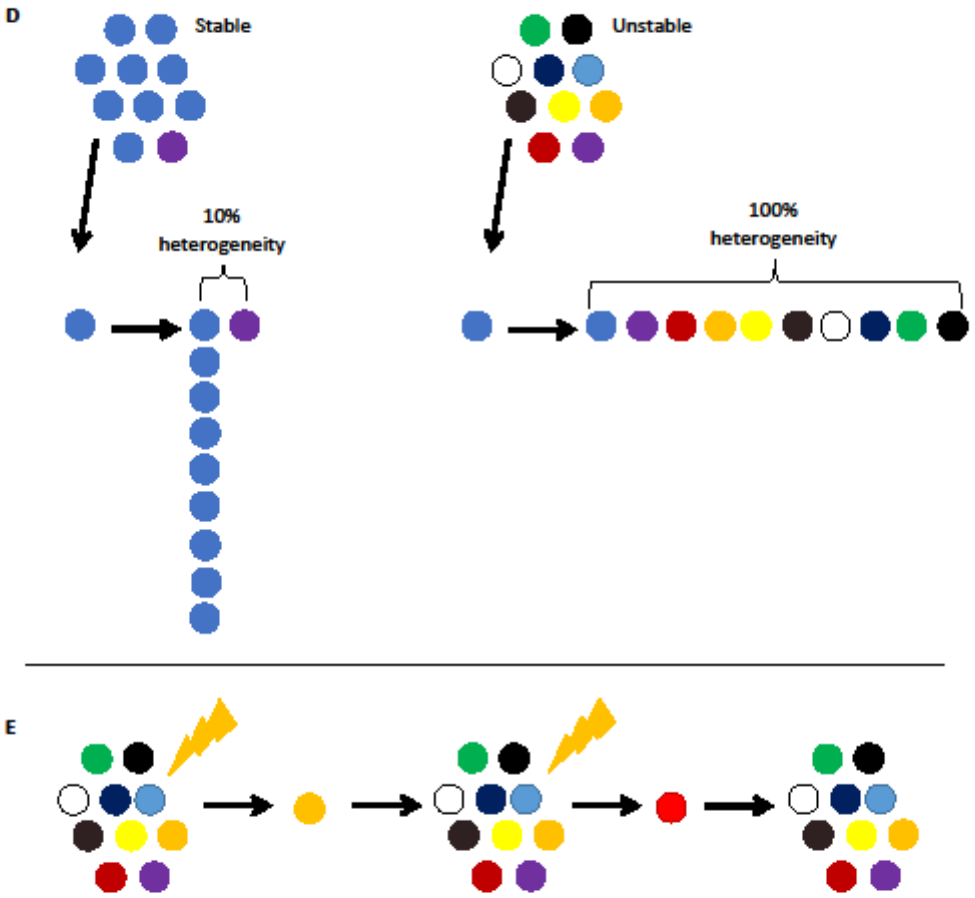
Single cells were isolated from resected ovaries originating from wild type and ovarian specific conditionally inactivated Brca1 and Brca1/p53 mice. Single cells were developed and kept in continuous culture.

Figure 3: Diagrams demonstrating the mechanism of fuzzy inheritance



The mechanism of fuzzy inheritance is illustrated in the above diagram. Each circle with a different color represents a cell with a unique genome. A) A low degree of genome-level heterogeneity is present in stable cell populations, while in unstable cells, genome heterogeneity is high. Stable cells pass stability, while unstable cells pass instability to future cell populations. B) Genome heterogeneity is linked to the heterogeneity of other traits, like cell growth. Stable cell populations exhibit more uniform and homogeneous growth while unstable cells exhibit heterogeneous and bimodal growth. C) A single stable cell will pass the same system inheritance (genome) to its daughter cells, while an unstable cell will pass altered system inheritance.

Figure 3, continued



D) A given degree of heterogeneity exists in each cell population, and that same degree of heterogeneity will be passed to future cell generations. For example, a stable cell population where 10% of the cells exhibit elevated frequency will pass the same degree of heterogeneity (10%) to its daughter cell populations. In unstable cells where heterogeneity is 100%, daughter cell populations will also exhibit 100% heterogeneity. E) Fuzzy inheritance increases the likelihood of cell population survival while under stress by providing the cell population diversity and evolvability necessary for survival.

CHAPTER 2: A SINLGE CELL GENOME-BASED MODEL OF OVARIAN CANCER EVOLUTION

In order to fully assess the mechanism of inheriting altered karyotypes, it is necessary to establish an experimental model that meets the following criteria: 1) the ability to monitor evolution in action under control and stress conditions; 2) the ability to detect system changes, or genome-level change; 3) the ability to utilize single cell experimental methods; 4) the ability to characterize multiple cell lines with different degrees of genome instability; and 5) the ability to measure the effects of genome heterogeneity on other heritable features of the system.

A number of *in vivo* and *in vitro* models have been previously developed for the study of ovarian cancer evolution, including transgenic mouse models and syngeneic mouse models (Vanderhyden et al. 2003, Fong and Kakar 2009), but not the maintenance of cellular heterogeneity. Both transgenic and syngeneic mouse models represent widely used research tools intended to mimic neoplastic transformation in humans that allow for the characterization of cancer evolution, and especially early genetic changes. A particular emphasis has been placed on early genetic events in ovarian cancer research, as most patients are diagnosed at late stages. The use of either type of model, as is, would not be suitable for the execution of this project. However, these models can be adapted to meet the criteria of the project. The following section describes the adaptation of both models for the execution of this project.

In vitro cellular model of fuzzy inheritance

An *in vitro* cellular model of cancer cell evolution is the most suitable model for characterizing the mechanism of fuzzy inheritance, as it meets all of the aforementioned criteria. An *in vitro* model is best for monitoring evolution in action, as cells can be isolated at any passage to monitor single cell dynamics, including genome-level changes and growth, at the single cell level. Furthermore, multiple cell lines can easily be incorporated through an *in vitro* model.

Previous studies investigated the process of spontaneous transformation and immortalization in wild type mouse ovarian surface epithelial cells (Lawrenson 2010). Wild type MOSE cells were kept under continuous cell culture conditions for a given period of time. Cells at earlier passages were genomically stable, while cells at middle or late passages exhibited increased genome heterogeneity. Because cell lines with different degrees of genome instability were desired for this project, early passage cells and late passage cells were obtained. Unfortunately, early passage cells from multiple early passages that were more genomically stable were not robust for cellular passaging or for single cell isolation, as cells reached doubling capacity and became senescent. Obtaining genomically unstable cells was much less difficult. Cells at later passages, after one year of continuous cell culture, and cells with that had been conditionally inactivated for Brca1 and Brca1/p53 at approximately 5 cellular passages, exhibited a large degree of karyotype heterogeneity and were robust in cell culture and single cell isolation, fulfilling the requirement for genomically unstable cells.

Because there was difficulty in obtaining genomically stable cells that are robust for cellular passaging and single cell isolation from wild type mice, MOSE cells isolated from the ovaries of conditionally inactivated Brca1 and Brca1/p53 were used. However even at early passage cells (passage 8), both of these cell types had exhibited a large degree of genome heterogeneity (data shown in Chapter 3). It became apparent that stable cell lines originating from MOSE cells that were robust for continuous cell culture and single cell isolation could not be obtained. Therefore, another cell line known for its karyotype stability and robustness in cell culture was needed. HCT116 cells fulfilled this requirement, as they exhibit long-term karyotype stability in cell culture and can grow well in cellular passaging under control and stress conditions (Knutsen et al. 2010).

Single cell isolation

Serial dilutions were performed on three different types of ex-vivo cells originating from the ovarian surface epithelium of C57-BL6 mice: spontaneously transformed wild type cells that have been in continuous culture for one year and spontaneously transformed, conditionally inactivated Brca1/p53 cells that have been in continuous culture for 70 days.

Serial dilutions were completed in 96-well plates. Two plates were used for the wild type cells, and four plates were used for Brca1/p53 -/- cells. As shown in Figure 4, 1,000 cells were pipetted into well A1. Cells were first diluted by a factor of 2 down column 1, and then across the plate using a multichannel pippettor. After dilutions, wells with single cells were identified the next day and verified by an independent investigator. After dilutions, wells with single cells were identified the next day and verified by an independent investigator. According to calculations, single cells were expected in the wells marked with a star. The following single cells were successfully isolated and cultivated into subpopulations:

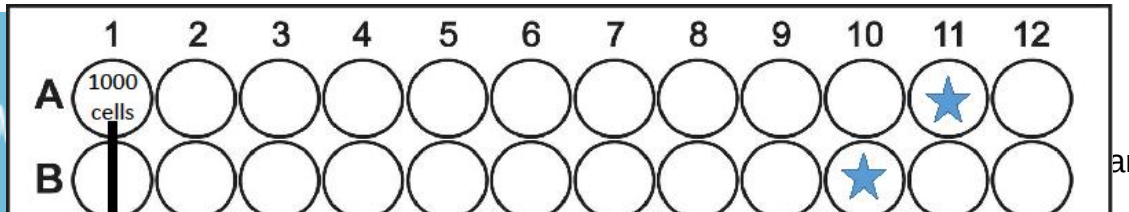
Wild type: WT_Sub1; WT_Sub2

Brca1 -/-: Br_Sub1; Br_Sub2; Br_Sub3

Brca1/p53 -/-: Sub1; Sub2; Sub3; Sub4

Each single-cell derived subpopulation was kept in continuous culture conditions. When cells reached confluency, all cells were trypsinized and transferred into 6-well plates, and then to T-25 and T-75 cell culture plates.

Figure 4: Serial dilutions for single cell isolation



Serial dilutions performed in ovarian surface epithelial cells resected from C57/BL6 mice that were wild type, and ovarian specific conditionally inactivated for Brca1 and Brca1/p53. 1,000 cells were pipetted into well A1. Cells were diluted by a factor of 2 down column 1 and then across each row. Single cells were expected in wells marked with a star.

Measuring population-level growth

The average doubling time and the number of doublings in the first passage are found in Table 1. All single-cell derived subpopulations underwent approximately 18-20 doublings in the first passage. Population doubling rates were measured for 30-60 passages for each single cell derived subpopulation. 100,000-500,000 cells were plated for each passage. Cells were

trypsinized and passaged at 80% confluency. Replicates of Sub1 and WT_Sub1 were also kept in parallel, continuous culture beginning at passages 8 and 7, respectively.

Measuring single-cell growth

Single cell growth was monitored for Sub1, Sub2 and WT_Sub1 cell populations. In order to determine the degree of single cell growth heterogeneity in unstable cell populations, cells were plated in a single T-25 culture flask, single cells were identified on Day 1 and their growth into colonies were monitored over the course of one week (Figure 5). Initial single cell growth experiments determined that WT_Sub1 cells were not ideal cells for measuring single cell growth. Cells were mobile as they would easily detach from the bottom of the flask and travel to a different location on the bottom of the flask, making single cell growth profiling difficult to monitor (data not shown). Sub1 and Sub2 cells, however, remained stationary throughout the duration of the experiment, allowing for identification of single cells and faithful monitoring of their growth.

Sub1 cells were used to optimize experimental parameters. Initial cell populations of 5,000 cells, 2,500 cells, 1,000 cells, 400 cells and 250 cells were plated in gridded T-25 culture flasks. Flasks seeded with 5,000 cells, 2,500 cells, and 1,000 cells were too concentrated for the monitoring of single cell growth heterogeneity. Single cells could be identified in flasks seeded with 400 and 250 cells. However flasks seeded with 250 cells were too sparsely plated, making single cell identification difficult. Single cells identified for the growth heterogeneity assay must be spaced at least 200μm apart so as to prevent colonies growing into each other.

Single cell growth heterogeneity was monitored over the course of one week. Images were taken of colony sizes almost daily. The endpoint chosen was 6 days. Six experiments were randomly

Table 1: Doubling time information for single cell-derived subpopulations

	Number of doublings until first passage	Average doubling time (hours)	Standard Deviation
WT_Sub1	20	26	10
WT_Sub2	19	28	12
WT_Sub1 replicate	--	22	10
Br_Sub1	Not determined	39	50
Br_Sub2	22	33	56
Br_Sub3	18	27	16
Sub1	20	42	35
Sub2	20	33	15
Sub3	18	28	16
Sub4	18	35	19
Sub1 replicate	--	23	10

Doubling information is provided for each isolated single cell-derived subpopulation. The first column calculates the number of doublings from single cell isolation until the first passage. The second column calculates the average doubling time in hours for each single cell derived subpopulation. The following number of doubling times were averaged for each subpopulation: wild type subpopulations, n=53 passages; Brca1 knockout, n=24 passages; Brca1/p53 subpopulations, n=38 passages.

chosen and the coefficient of variation (CV) values for daily growth were calculated on days 4, 5, and 6. There was no significant difference in the CV values between the chosen endpoints **(**Figure 6). By 5 or 6 days, at least one colony would have entered the exponential phase, with colony sizes exceeding 500 cells. Colony sizes of more than 500 cells had an approximate surface area of 200µm or more. Because single cells initially identified on Day 1 were at least

Figure 5: Monitoring of *in-situ* single cell growth

Single cells were identified on Day 1 and their growth was monitored daily for up to 7 days.

200µm apart, 6 days was chosen as an endpoint to prevent colonies from growing into each other.

Measuring growth heterogeneity

Images of single cell derived colonies were taken almost daily and were quantified manually using ImageJ software. The number of cells per colony on day 6 was used to measure growth heterogeneity for most experiments. Growth heterogeneity was measured by calculating the CV, which is obtained by dividing the standard deviation by the mean. Several studies have used the CV to quantify different types of heterogeneity, including growth heterogeneity (Keren et al. 2015).

Power calculation

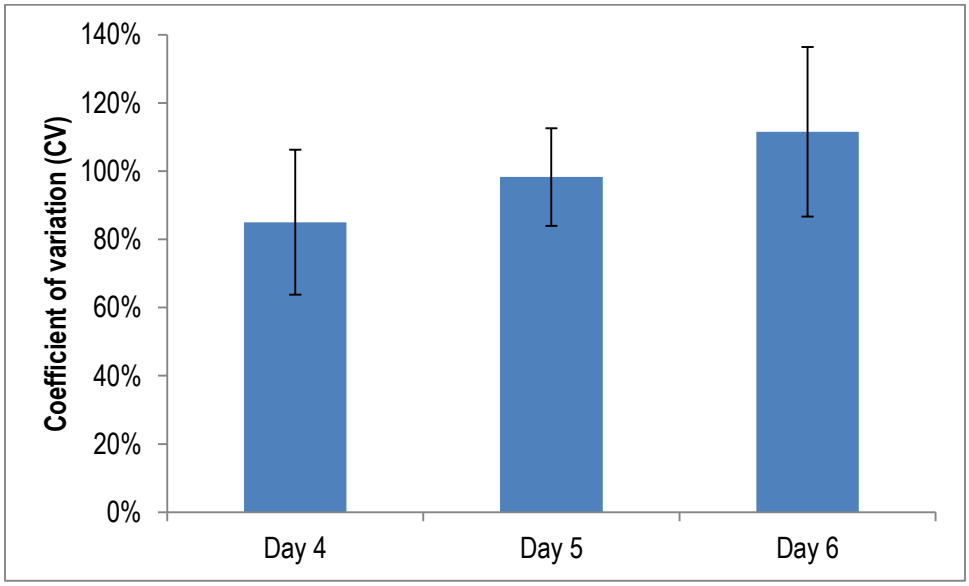
An alternate method of power calculation was completed in order to obtain a sample size that will give an accurate assessment of the growth heterogeneity of a cell population. Most power calculations and other forms of hypothesis testing use the statistical average as the point estimate. The single cell experiments completed measures not only growth, but variation. While variation is a component of power calculations, it is not the end measurement. Therefore, an alternate method of determining an appropriate sample size was developed in order to account for variation. Increased sample size was plotted by the standard deviation of single cell growth (Figure 7a). As the sample size increased by an increment of 5 cells, the variation in the standard deviation decreased and eventually leveled off at approximately 25-30 cells. Therefore, a sample size of 25-30 cells were used. Multiple single cell *in situ* growth experiments were completed at the appropriate sample size. Each experiment demonstrated a wide degree of growth heterogeneity, as single cells exhibited bi-modal or even multi-modal growth (Figure 7b).

Elimination of confounding factors

Some factors may influence the rate of growth of cells growing on a monolayer dish. One such factor is contact inhibition, which occurs after cells growing on a monolayer dish come into physical contact, and cease growing. In order to determine that contact inhibition was not a factor in the growth heterogeneity of colonies, in situ single cell growth experiments were performed on unstable Sub1 cells and day 6 colony sizes were quantified. Colonies were then grouped based on their type of growth: clustered growth and non-clustered growth to determine if contact inhibition was a factor in growth heterogeneity (Figure 8a). It was found that contact inhibition was not a factor in overall cell growth, as there was no significant difference in the growth between colonies where cells were tightly clustered, and those colonies where cells were more dispersed (Figure 8b).

The spatial proximity between two separate colonies may also be another factor in the growth heterogeneity of single cell derived colonies, where two cells in closer proximity may have similar growth rates compared to cells that are more distant. In order to determine

Figure 6: Growth heterogeneity is similar across different endpoints

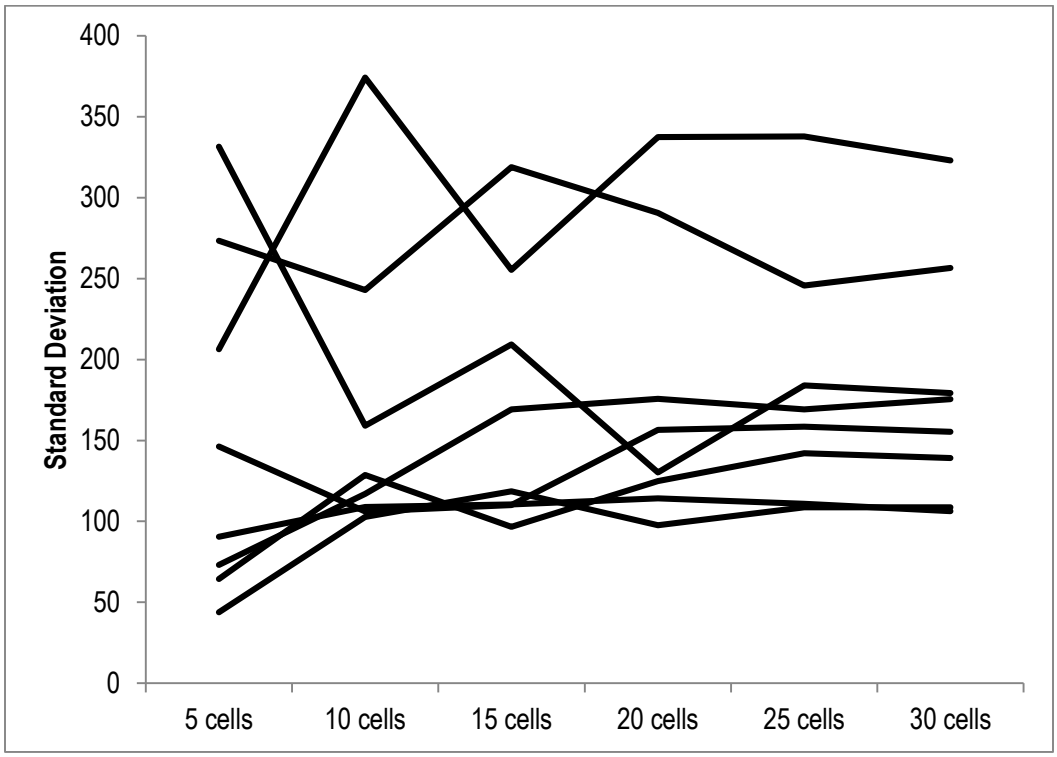


Single cell growth measured at different endpoints exhibited similar heterogeneous growth distributions. Single cell growth was measured at days 4, 5 and 6 (n=6 experiments). The coefficient of variation was calculated and compared. There was no significant difference in the CV among the various endpoints (students t-test, p-value ≥0.1).

whether spatial proximity played a factor in growth heterogeneity of cells, 50 cell pairs that grew between 200 µm- 250 µm were compared to cell pairs with greater than 250 µm difference. Results

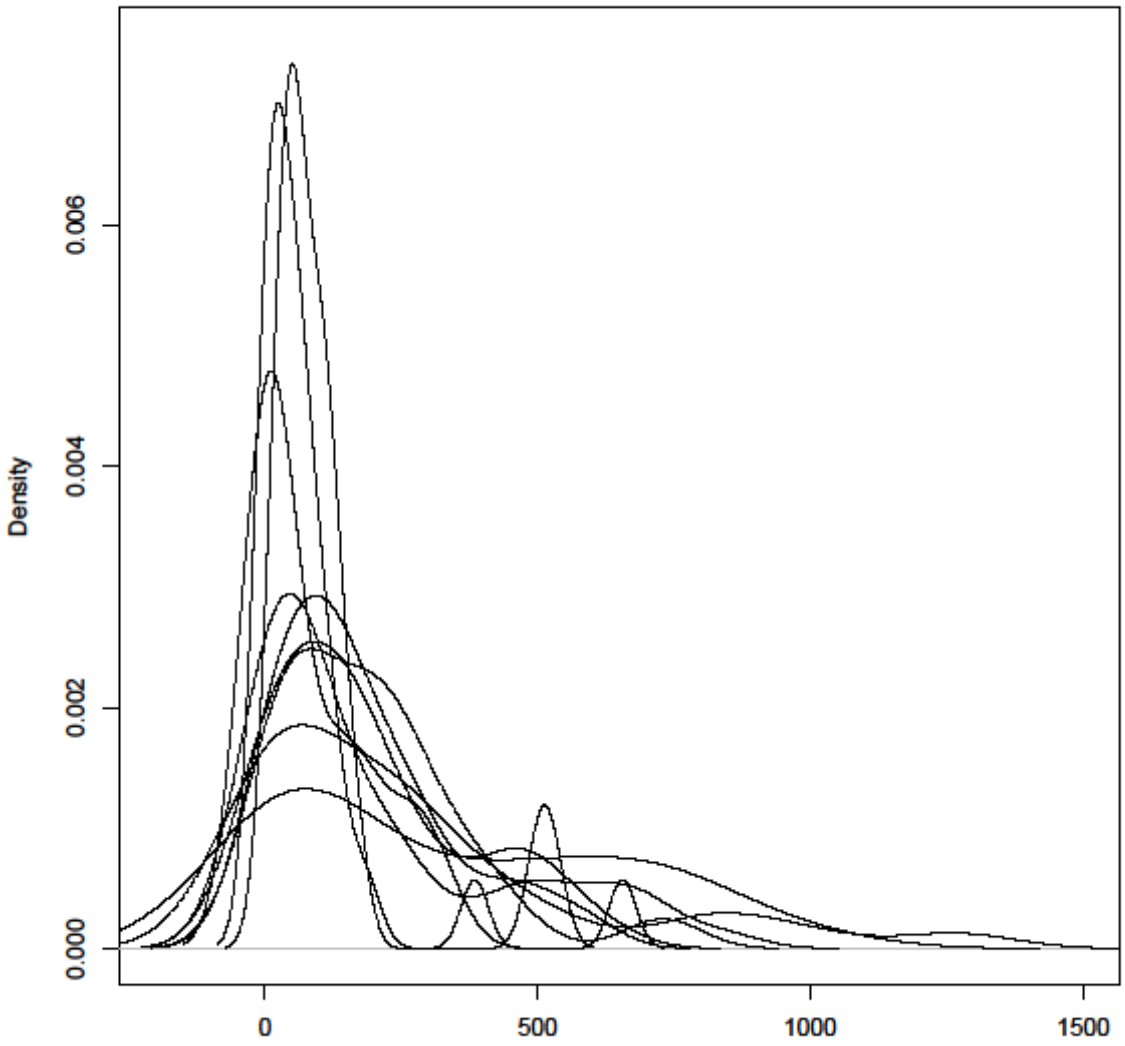
demonstrate that there was no significant difference in colony size between colonies in close proximity compared to colonies that were more distant (Figure 8c).

Figure 7a: Alternate method for determining sample size for *in situ* single cell growth experiments



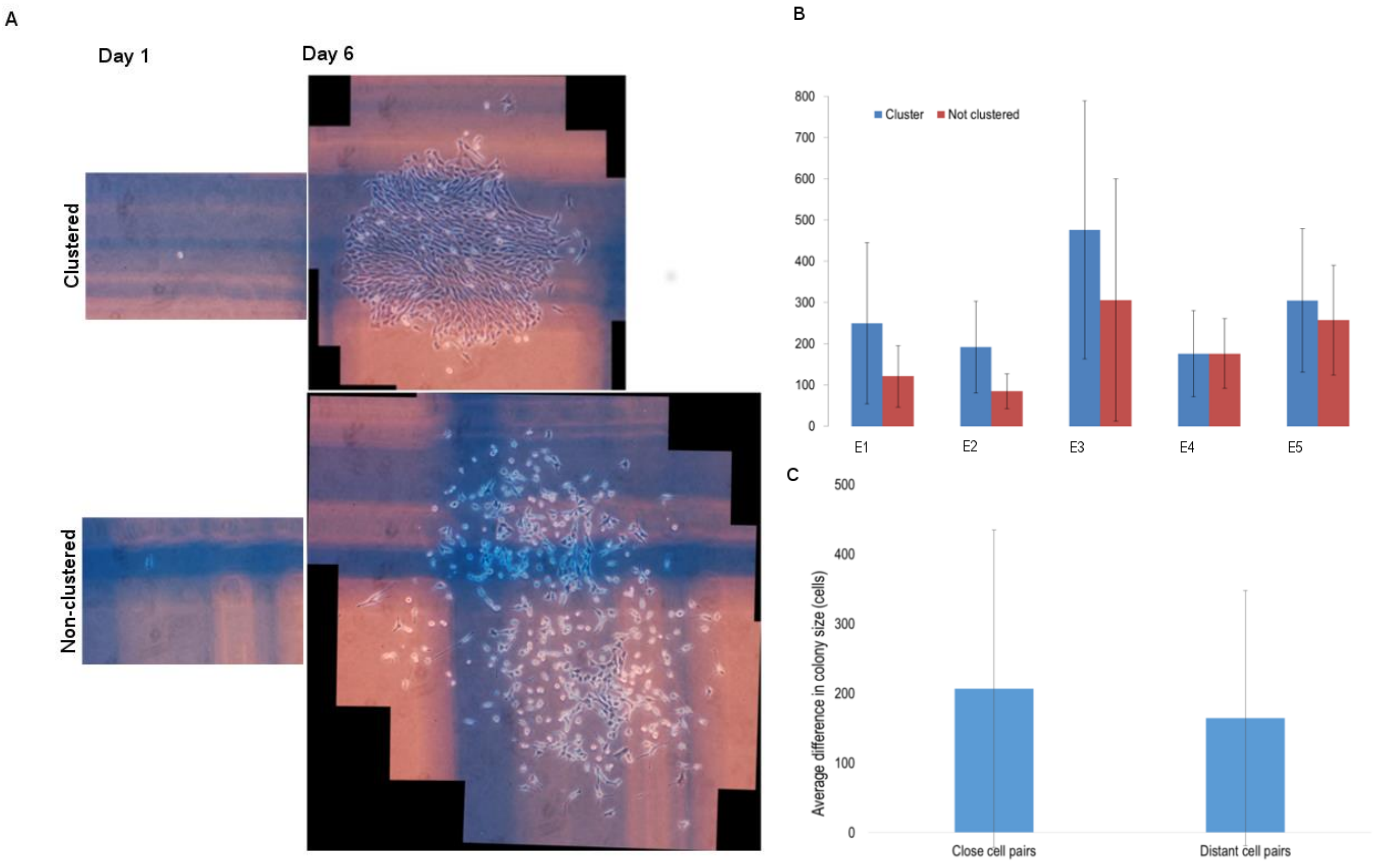
Sample size determination for *in situ* single cell growth experiments. Most power analyses and hypothesis testing use the mean as a point estimate, however single cell growth experiments measure variation in growth. Therefore, an alternate method of determining a sample size is proposed that plots increasing sample size against the standard deviation of single cell growth. As the sample size increases, the variation in the standard deviation decreases and levels off at 30 cells.

Figure 7b: Replicate experiments demonstrating growth heterogeneity of unstable cells



Experimental replicates of unstable Sub1 cells. Each experiment demonstrated a large degree of growth heterogeneity, as cells manifested into bi-modal or multi-modal growth distributions.

Figure 8: Growth heterogeneity unaffected by contact inhibition or spatial proximity to other colonies



Contact inhibition and spatial proximity to other colonies was examined to determine its influence on growth heterogeneity. (A) In 5 randomly chosen experiments, colony growth was categorized as "clustered" if cells grew in a tightly clustered colony, or "non-clustered" if the individual cells of a colony were spread out. (B) Average colony size was compared between clustered and nonclustered in five independent experiments, and there was no significant difference between the two groups. (C) Spatial proximity between individual colonies was also tested to determine if those colonies growing closer together exhibited similar growth patterns compared to those colonies that were more distant. Results show that spatial proximity to other colonies was not a factor in colony growth.

CHAPTER 3: SINGLE CELL HETEROGENEITY IN STABLE AND UNSTABLE CELL POPULATIONS

Introduction

Single cell heterogeneity represents a paradoxical dilemma for current and future biological research (Huang et al. 2009, Heng et al. 2011a, Heng et al. 2011b). Many of the fundamental principles upon which biology is based has been completed by experimental methods and analyses based around the statistical average. While data that has been generated by average-based methods and analyses has contributed much to our current understanding of biology, increased single-cell studies have begun to demonstrate that the "average cell" does not match up to the single cell profiles of a cell population (Wang and Bodovitz 2010, Pelkmans 2012). For example, different cells of the same cell population may use different molecular mechanisms to carry out the same cellular process (Heng et al. 2011a, Stevens et al. 2011). Cellular heterogeneity is a fundamental component for robust biological systems and evolution, and is central to understanding the evolution of somatic cell diseases (Heng 2008, Heng 2013a).

Single cell heterogeneity has been previously studied to understand microbial resistance and evolution. It has been shown in many studies that diverse phenotypic states exist within a cell population, where single cells are grouped into multi-modal distributions. These studies illustrate many interesting findings: 1) distinct cell states are due to the stochastic fluctuations of single cells (Elowitz et al. 2002, Balaban et al. 2004); 2) cell populations are diversified by individual cell states that are important for cellular evolution (Kussell and Leibler 2005, Acar et al. 2008, Cohen et al. 2008); and 3) causal relationships cannot usually be determined in the natural settings where stochasticity, or noise, is a dominant feature, while these relationships are well detected under defined experimental settings (Heng et al. 2006c, Heng et al. 2011a).

Among genomically stable cells, variation among single cells may be attributed to novel regulatory mechanisms (Pelkmans 2012). That cell populations are mainly isogenic is well accepted in biology, and that genetic and nongenetic variation among individual cells contribute to varying population dynamics is a view shared by many (Huang et al. 2009, Gupta et al. 2011). However, many studies have challenged the well-accepted concept of karyotype homogeneity. First, most cancer cell populations exhibit a large degree of heterogeneity, especially karyotype heterogeneity (Heng et al. 2006b, Heng 2009, Heng et al. 2013a). Second, a re-examining asexual reproduction finds that, contrary to conventionally held beliefs, cells that undergo asexual reproduction actually do not genetically produce identical daughter cells (Heng 2007b, Gorelick and Heng 2010, Horne et al. 2013). Many studies have demonstrated that genome heterogeneity exists even in normal tissue, a phenomenon termed somatic chimerism, and is linked with various physiological conditions and somatic cell disease conditions (Heng et al. 2004, Iourov et al. 2008, Heng et al. 2010, Sgaramella 2010, Duncan et al. 2012, Heng 2013c, Heng et al. 2013b, Hulten et al. 2013). Somatic chimerism challenges the idea that cells in normal tissue only contain normal karyotypes. Third, according to the genome theory, the evolutionary unit of selection is the genome system (Heng 2009, Heng et al. 2010, Heng et al. 2011a, Stevens et al. 2011). In the unstable cancer macro-cellular phase, genome replacement drives cancer evolution (Heng 2009, Heng et al. 2011a, Stevens et al. 2011, Heng 2013a).

Altogether, the most significant form of genetic heterogeneity in cancer cells occurs at the genome level. Our first step in biological research is to validate karyotypic integrity at single cell resolution in a cell population. Experimental model systems are often clonal or linear, which are fundamentally different than natural settings where genome-level heterogeneity is more common. The characterization of genome level heterogeneity in these two settings may explain how the two

settings are fundamentally different. The relationship between different types of heterogeneity in cancer can be explained by various proposed multiple level landscape models (Stevens et al. 2011, Heng et al. 2013a, Huang 2013).

In this chapter, the degree of genome heterogeneity and its effects on population dynamics were investigated using primary ovarian surface epithelial cells that originated from wild type and conditionally inactivated Brca1/p53 C57/BL6 mice, in order to examine the importance of single cell analysis in biological research, its effects on population dynamics (specifically growth heterogeneity), and its implications on average-based technical and analytical research methods commonly used in biological and specifically cancer research. Using single cell culture and single-cell spectral karyotyping, a panel of single cells originating from the ovarian surface epithelium in spontaneously transformed wild type and conditionally inactivated Brca1/p53 mice were isolated, and the degree of single cell genome heterogeneity and single cell growth profiles were determined. It was found that unstable cancer cells exhibit a large degree of genome heterogeneity at the single cell level. Additionally, cell populations that exhibit increased genome heterogeneity also exhibit elevated growth heterogeneity. Comparison of single cell analyses to population-level analyses found that the statistical average was inconsistent with single cell karyotype and growth profiles of unstable cell populations. This indicates that the average is a poor measure for cell populations that exhibit a high level of genome heterogeneity. Finally, it is demonstrated that genome heterogeneity is mediated by instability, and is a key feature of cancer cell populations that exhibit a high level of chromosome instability (CIN), and outliers drive population growth.

Materials and Methods

Cell lines: Spontaneously transformed ovarian surface epithelial cells were isolated as previously described (Roby et al. 2000). Briefly, wild type C57/BL6 mice were sacrificed at age 6 weeks, upon which the ovaries were resected and scraped onto 6 well plates. Cells were maintained under continuous culture conditions. Conditionally inactivated Brca1/p53 knockout ovarian surface epithelial primary cells were obtained from the University of Ottawa (Clark-Knowles et al. 2009). Single cells were isolated from various *ex-vivo* cells as described earlier. Briefly, a panel of single cells were isolated from both cell lines through serial dilutions in 96 well plates after 1 year (wild type spontaneously transformed) or 60 days (Brca1/p53 conditional knockouts) in continuous culture. Immediately after single cell dilutions, single cells were identified using a Nikon TMS inverted microscope and confirmed by a second investigator. HCT116 cells were a gift from the lab of Dr. Bert Vogelstein (Lengauer et al. 1997).

Cell culture: Standard cell culture techniques were used for all cell lines. Spontaneously transformed wild type cells and conditionally inactivated Brca1/p53 knockouts were kept in high glucose DMEM, supplemented with 4%FBS, antibiotic, and insulin, transferrin and sodium selenite (ISST) growth supplement. HCT116 cells were maintained in RPMI medium, supplemented with 10% FBS and antibiotics.

Cytogenetic metaphase slide preparation and spectral karyotyping (SKY): Cytogenetic slides and spectral karyotyping were prepared and completed, respectively, as previously described (Heng et al. 1992, Heng et al. 2006a, Heng et al. 200c6, Ye et al. 2009). Briefly, cells were plated into T-75 culture dishes. After 3 days of growth or when cells reached approximately 50%-60% confluency, cells were treated with 100µM colcemid for 2 hours. Next, mitotic cells were obtained through a mitotic shake-off, washed in PBS and re-suspended in a hypotonic solution of 0.4% KCl for 30 minutes at room temperature. After hypotonic treatment, cells were prefixed using a 3:1 methanol:

acetic acid fixative by applying approximately 120µL of fixative to centrifuge tubes. Cells were then resuspended and fixative was applied three times at varying times and temperatures: first and second fixatives applied for 30 minutes at room temperature; and third fixative applied overnight at 4°C. Cells were then dropped onto slides. Metaphase slides were then denatured and hybridized with mouse probes. Images of mitotic structures were captured using a charge coupled device camera.

SKY was completed on metaphase slides according as previously described (Ye et al. 2009). Metaphase slides were first washed in Earl's medium. Slides were then treated in trypsin solution (5g/L trypsin and 2g/L EDTA in Earl's medium) for 20 seconds at room temperature. Following trypsin treatment, slides were washed in water and dehydrated in ethanol series (70%, 80%, 100%) for 2 minutes each. Slides were left to air dry.

After trypsin treatment, slides were ready for chromosome denaturation. Slides were placed in 2X SSC (saline sodium citrate) solution for 2 minutes at room temperature, and then dehydrated in ethanol series 70%, 80%, 100% for 2 minutes each. Slides were left to air dry. Slides were then placed in a coplin jar containing 40mL of heated (72ºC) denaturation solution (70% formamide/2X SSC, pH 7.0) for 60-90 seconds, and then immediately placed in cold ethanol series (70%, 80%, 100%) for 2 minutes each. Slides were left to air dry.

Following chromosome denaturation, SKY (spectral karyotyping) probe was denatured by incubation at 80ºC in a water bath for 7 minutes. Probe was then placed in a 37ºC water bath for 10 minutes. Denatured probes were then applied to metaphase slide. Plastic cover slips were placed over slides and the edges were sealed with rubber cement and placed in a 37ºC humidified incubator overnight. Slides were then washed in 0.4% SSC at 72ºC for 5 minutes. Following this wash, slides were placed in another washing solution (4X SSC/0.1% Tween 20) for 1 minute.

Slides were then allowed to drain, Cy5 applied to slides, cover slip placed on slides and plac incubated at 37ºC for 40 minutes. Following incubation, slides were washed three times in washing solution at 45ºC for 2 minutes per wash. Cy5.5 staining reagent was then applied directly onto slides, cover slip was placed on top of slide and slides were placed in 37ºC incubator for 40 minutes. Following incubation, slides were washed three times in washing solution at 45ºC for 2 minutes per wash. Slides were tilted and drain. Finally, anti-fade/DAPI was applied to slides, glass cover slide was placed, and slides were ready imaging.

Karyotypic analysis: Karyotypic analysis was completed as previously described (Heng et al. 2006b, Heng et al. 2006c). Briefly, SKY analysis was completed for 30 cells for each sample where the number of NCCAs and CCAs were enumerated. NCCAs are either structural (such as translocations) or numerical (such as aneuploidy). A chromosome aberration is designated an NCCA if it is present in 4% or less of the sampled cells. All chromosome structures were visualized on heatmaps. In each heatmap, row corresponds to one single cell.

Population-level counting: Each cell line and single cell derived subpopulation were grown in T-75 flasks. Cells were passaged and enumerated when they reached approximately 80% confluency. For subsequent passages, cells were re-plated in fresh T-75 culture flasks.

In-situ single cell counting: 400 cells were plated in gridded and labeled T-25 cell culture flasks Single cells were identified on day 1 and cell growth was imaged and counted daily.

Statistical analysis: Both normal and derivative chromosome structures were enumerated to determine the degree of karyotypic heterogeneity. The sample size for SKY analysis was determined through two methods to maintain statistical robustness as well as to account for variation of chromosome structures. A power analysis was first completed where $\alpha = 0.95$, $\beta = 0.9$) yielded a sample size of 15 cells. A second method was also used to find the sample size that

took chromosome variation into account. The variation of each chromosome, as measured by the standard deviation, plotted against increased sample size (Figure 8f). As the sample size increased, chromosome variation began to decrease and eventually tapered off at 15 cells. By using these two methods, at least 15 cells were analyzed in each sample.

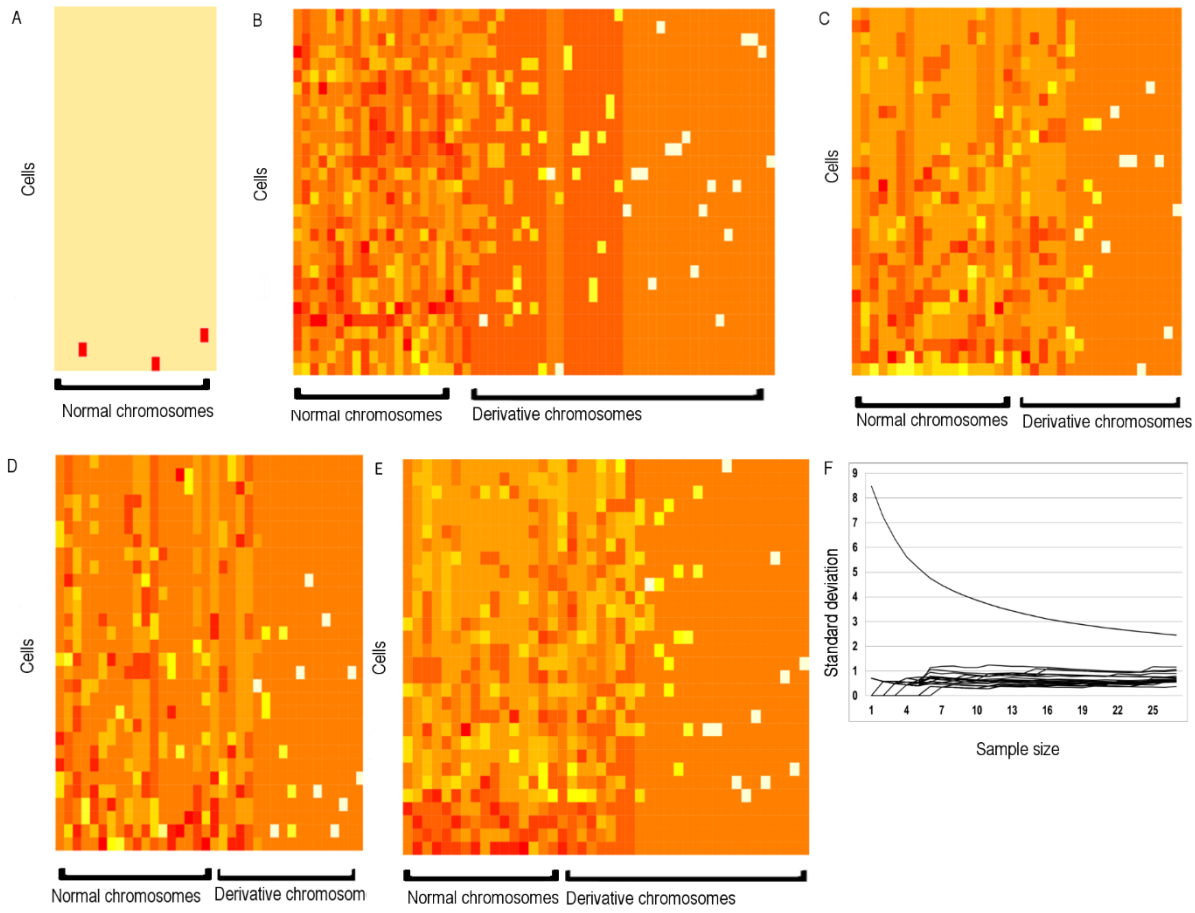
Carboxyflourescin succimydyl ester (CFSE) tri-color stain generation sort Tri-Color Stain: Generation sort was performed using Carboxyflourescin succinmydal ester (CFSE), purchased from Invitrogen and assay performed to manufacturer protocol. First, cells were first synchronized using a double thymidine block. 2 mM thymidine was applied to cells growing in a T-75 culture flask when they reached approximately 40% confluency for 12 hours. This first block keeps cells in S phase. After 12 hours, cells were washed twice with PBS and replated at a 1:3 dilution. Cells were left to grow overnight. After the release, the second thymidine block (2mM) was added for 12 hours. Following the second block, cells were washed twice with PBS and resuspended with CFSE dye at a concentration of 6µL/100,000 cells. Cells were incubated for 10 minutes in the dark at 37ºC. At 1 minute intervals, centrifuge tubes were inverted to ensure CFSE uptake into cells. Following the incubation, cells were quenched with 5X volume of ice cold media and placed on ice for 5 minutes. Cells were then washed and resuspended in fresh media three times. Cells were plated and left to grow for 3 days. On the day of analysis, cells were trypsinized, and incubated with Hoechst 33342 live cell nucleic acid dye for 45 minutes at 37°C. Cells were washed twice, and then propidium iodide was added immediately before analysis. Cells were sorted and analyzed using BD FACSDIVA and BD LSR II and analyzed with Cellquest software.

Results

Unstable cells cannot be cloned

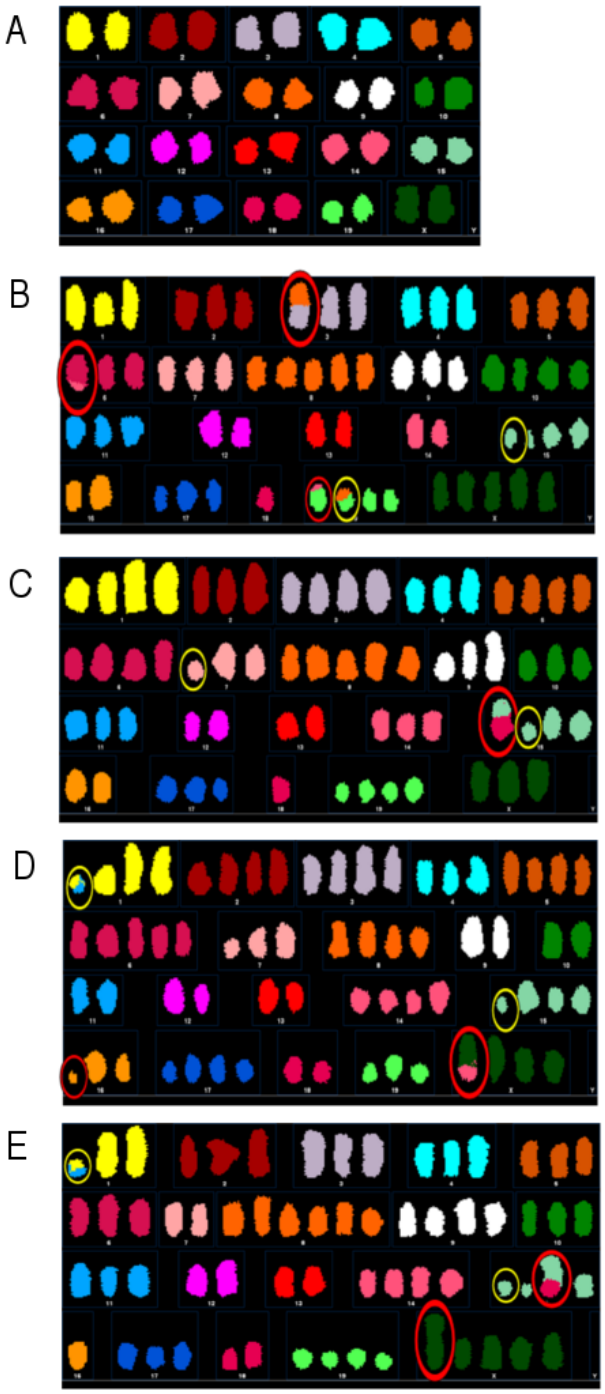
Multi-level heterogeneity is present in most cancer types (Heppner 1984). Specifically, karyotype heterogeneity is a characteristic feature of cancer cell populations (Heng 2009, Heng 2013a). However, despite its ubiquitous presence, its significance has often been overlooked as more effort has been placed on identifying a pattern of clonal evolution. It has been demonstrated that sequential and accumulated mutations in a population of cancer cells can result in tumor growth in a number of cancer models (Nowell 1976, Fearon and Vogelstein 1990, Maley et al. 2006). However recent basic and clinical research studies conducted at population-level and single cell resolution demonstrated that punctuated evolution is much more common in most cancers than clonal evolution. This makes the identification of common biomarkers difficult because each cell exhibits its own unique genomic profile (Heng et al. 2006c, Navin et al. 2011). Most of the heterogeneity found in cancer cell populations is noncloncal, making it difficult to detect using conventional methods based around the average cell profile. Therefore, in order to determine the degree of heterogeneity in various cancer cell populations, and to understand how the degree of genome-level heterogeneity affects population dynamics (specifically cellular growth and evolution), a panel of single cells were isolated using serial dilutions to generate pure, single-cell derived cell populations from wild type mouse ovarian surface epithelial (MOSE) primary cells that spontaneously transformed in cell culture. Cells were kept in continuous cell culture conditions for one year prior to single cell isolation. Spectral karyotyping (SKY) was used to determine the degree of genome-level heterogeneity. The parent population exhibited a large degree of genome-level heterogeneity, as no two cells were karyotypically identical, as compared to early passage wild type primary cells that exhibited very little

Figure 9: Unstable cells are not clonable



Genome-level heterogeneity depicted by heatmaps. A) Karyotype heatmap of early passage wild type mouse ovarian surface epithelial cells after two days in cell culture. As expected, most cells exhibit a normal karyotype. B) Karyotype heatmap of parent cell population of spontaneously transformed wild type ovarian surface epithelial cells after one year in continuous cell culture. C-D) Single cell-derived subpopulations originating from spontaneously transformed wild type parent cell population kept in continuous cell culture conditions for 23 days (C, WT_Sub1) and 40 days (D, WT_Sub2). A high level of genome heterogeneity was observed in both subpopulations. Also, no direct intermediates could be traced to the parent cell population. E) Karyotype heatmap of single cell-derived WT_Sub1 was kept in continuous culture 117 days after single cell isolation. Increased genome heterogeneity was observed. F) Sample size determination for SKY analysis.

Figure 10: Spectral karyotyping of unstable cell populations



Representative spectral karyotypes from each cell population. A) Wild type early passage MOSE cells. B) Spontaneously transformed WT MOSE cells after one year in culture. Single cell-derived subpopulation (C) WT_Sub1 and WT_Sub2 (D) after 23 and 40 days, respectively days in continuous cell culture conditions. E) WT_Sub1, 117 days after single cell isolation. Structural NCCAs are circled in red, structural CCAs are circled in yellow.

Karyotypic heterogeneity was high in both subpopulations. NCCAs (structural and numerical) greatly outnumbered CCAs (Figure 9c, 9d, Figure 10). No two cells were identical, and no common karyotypic intermediates were found in either subpopulation, suggesting that punctuated evolution was dominant. Subsequent analysis at a later time point demonstrated that heterogeneity remained high and increased with time, as indicated by the NCCA index (Figure 9e).

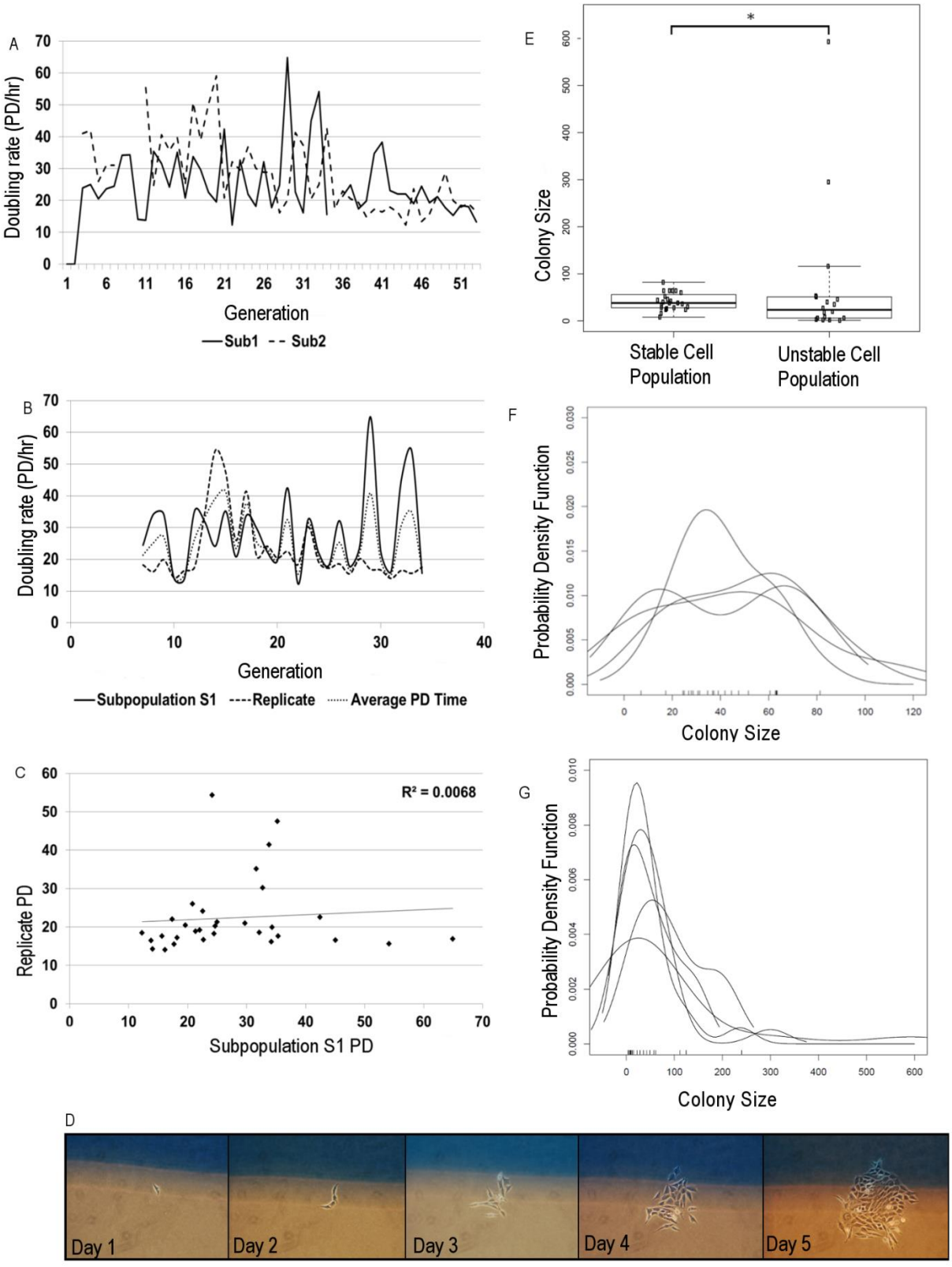
According to models of clonal evolution, it is expected that some clonal karyotypic aberrations should have been detected in subpopulations that were generated from a single cell. In order to maintain the purity of a cell line, it is common practice in cell culture to subclone a cell line after extended periods of cell culture. The data presented suggests that in cell lines where the genome is unstable, karyotypic cloning is not possible even within a short time period. SKY analysis of two single-cell derived subpopulations demonstrated that no two cells had identical karyotypes, even in a short time period, suggesting that the karyotypic cloning of an unstable cell population is not possible. A common critique has been that the results of this data are artifacts of cell culture. To demonstrate that the results presented are not artifacts, the same cell culture and techniques were used on HCT116 (data shown in Chapter 4) and HeLa cell lines with stable genomes. Both cell lines exhibited minimal karyotypic change over time (Shih et al. 2001, Knutsen et al. 2010). Therefore, in unstable cell lines that cannot be karyotypically cloned, heterogeneity is dominant. The data presented is consistent with previous studies that associate a high degree of genome instability with punctuated cancer evolution where stochastic genome alterations are dominant (Heng et al. 2006a, Heng et al. 2006b, Heng et al. 2006c).

Karyotype heterogeneity leads to growth heterogeneity

karyotypic heterogeneity (Figure 9a, 9b, Figure 10). SKY of two single-cell derived pure subpopulations was completed after 3 weeks in continuous culture.

Karyotype heterogeneity can have significant impacts on the dynamics of a cell population. Karyotype heterogeneity has been linked as a causal factor in transcriptional heterogeneity, growth heterogeneity, and survival heterogeneity in a number of experimental systems (Harewood et al. 2010, Pavelka et al. 2010, Levy et al. 2012, Heng et al. 2013a, Stevens et al. 2013). In order to determine the effects of karyotype heterogeneity on cancer cell growth, single cell and population-level growth was compared using karyotypically stable (HCT116 cells) and karyoytpically unstable cell lines (single-cell derived subpopulations of conditionally inactivated Brca1/p53 MOSE primary cells). Single cell growth profiles were compared to the population-level growth profiles to determine how single cells contributed to population level growth. Population doubling (PD) was monitored. PD is a commonly used technique to measure population growth (Mehrara et al. 2007). Monitoring PD of karyotypically unstable single-cell derived subpopulations demonstrated that each subpopulation had a different doubling rate and overall growth (Figure 11a). Next, the growth of two biological replicates were compared. The PD times of replicates found that there was considerable difference between the doubling times, suggesting independent evolution of each replicate (Figure 11b). Regression analysis of the doubling times was completed between the two replicates and no correlation was found (Figure 11c; r^2=0.0068). While some PD values could be considered as outliers or growth anomalies and might be excluded in order to give a weakly positive correlation, the exclusion of any PD time, regardless of how fast or how slow, would not provide an accurate assessment of population-level growth, as it may be that each replicate is independently evolving, therefore exhibiting its

Figure 11: Genome-mediated growth heterogeneity



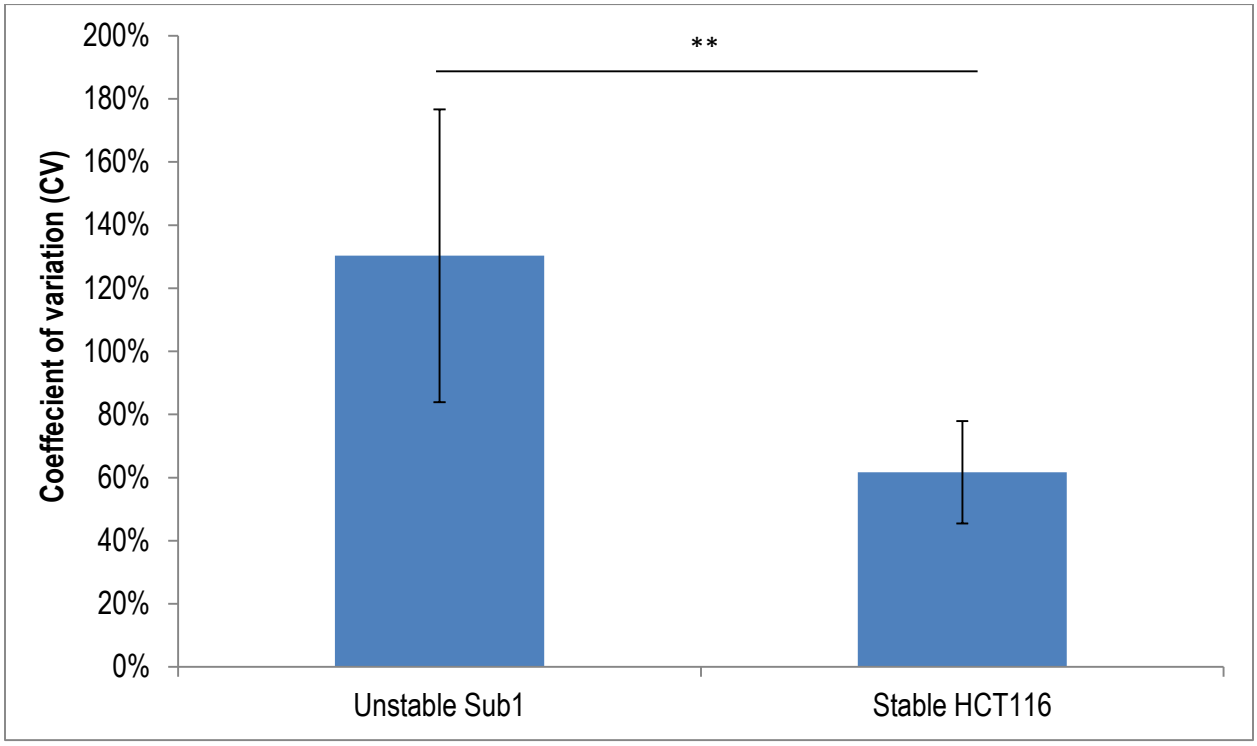
Caption: Growth heterogeneity in unstable cell populations. A) PD rates of single cell populations were monitored, where each subpopulation exhibited unique and fluctuating growth rates. Growth rate variation was moderate, and measured by CV (WT_Sub1, 40%, WT_Sub2, 42%). B) PD rates of two replicates of WT_Sub1 were compared. Both replicates exhibited unique and fluctuating growth rates. (Replicate 1 CV = 44%; replicate 2 CV = 45%, n=2). C) Regression analysis was completed between two replicates displayed no correlation (r^2=0.0068). D) Representative example of *in-situ* single cell growth. Single cells were identified on day 1 and their daily growth was monitored. E) Single cell growth in HCT116 cells and unstable Sub1 cells. Unstable Sub1 cells exhibit significantly higher growth variation (CV=200%) than stable HCT116 cells (CV=44%) (F-test, $p \leq 1.4 \times 10^{-6}$. Kernel density estimates are shown in panels F (stable cells) and G (unstable cells). Stable cell populations are unimodal and exhibit narrow distributions while unstable cells are bimodal or multimodal, exhibiting diverse and broad growth distributions.

own PD rates. Furthermore, the exclusion of any particular PD time-point would remove any contribution that outlier cells could potentially make to the growth of the cell population. Finally, in addition to the overall growth of each subpopulation being different, the PD growth rates fluctuated with every passage. The fluctuating PD rates may be due to the high karyotypic heterogeneity in each subpopulation, suggesting considerable variability in growth within each subpopulation. To determine the degree of growth heterogeneity within a subpopulation, *in-situ* single cell growth was monitored over a period of 6 days (Figure 11d). A karyotypically unstable conditionally inactivated Brca1/p53 knockout primary cell single-cell derived subpopulation (Sub1) (SKY data shown later) was compared against karyotypically stable HCT116 cells (data shown later). A total of 400 cells were plated in gridded and labeled T-25 culture flasks for each cell type. Single cells were identified on day 1 and growth was monitored daily for a total of 6 days or until single-cell derived colonies began to merge. Interestingly, single-cell growth heterogeneity was significantly more variable in unstable Sub 1 cells than in karyotypically stable HCT116 cells (Figure 11e-g, Figure 12). Single-cell derived colonies in the HCT116 cells exhibited the same overall growth (range 8-82 cells), exhibiting low growth heterogeneity, as

measured by the coefficient of variation (CV), 44% (Figure 13). In contrast, single-cell derived unstable Sub1 colonies exhibited a high degree of heterogeneity in

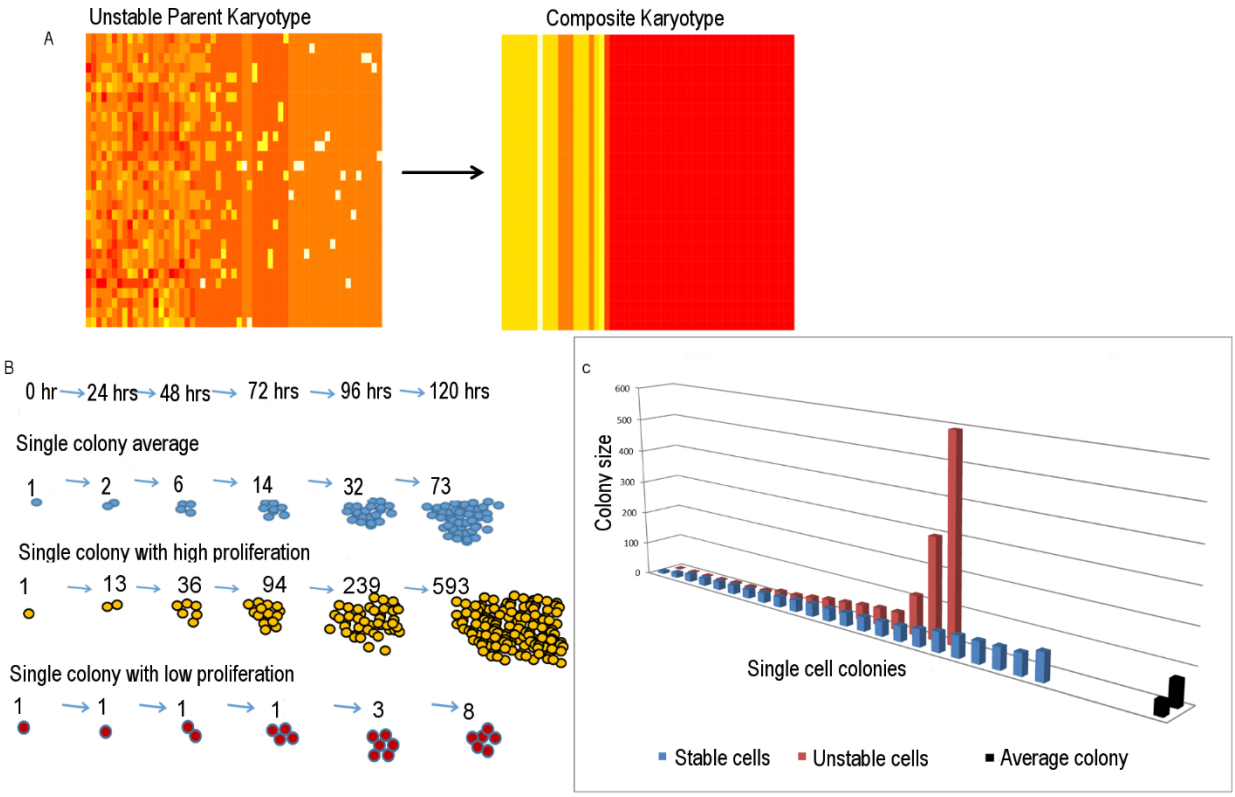
their overall proliferation (CV= 200%; Figure 13). Most cells had slow-to-moderate proliferation and only a few exhibited very high proliferation; additionally, some cells did not grow at all. For example, one highly proliferative cell of the unstable Sub1 colony grew to 593 cells in just 6 days. A significant disparity in growth among unstable Sub1 cells suggests that traditional

Figure 12: Growth heterogeneity comparison between unstable cells and stable cells



Growth heterogeneity was compared between unstable Sub1 cells and stable HCT116 cells. CV was calculated and averaged among 5 experiments. Unstable cells exhibit significantly increased growth heterogeneity than stable cells (students t-test, $p \leq 0.001$)

Figure 13: Average is a poor measure for genomically heterogeneous cell populations



Statistical averages are not representative measures for genomically heterogeneous cell populations. A) A composite karyotype was generated that averaged all normal and derivative chromosome frequencies, thus homogenizing genomic heterogeneity. The composite karyotype exhibits none of the variation that is characteristic of heterogeneous cell populations. B) Diagram of in-situ single colony growth of unstable cells. Cell proliferation was enumerated and averaged over 6 days, higher proliferating cells were compared against low proliferating cells. C) Single colony growth of unstable and stable cells. Most unstable cells grew at a slow to moderate rate and only few cells were highly proliferative. In contrast, stable cells exhibited the same, uniform growth. Average colony growth measured at 73 cells and 41 cells for unstable and stable cells respectively. While average colony growth accurately represent stable cells, unstable cells are not accurately characterized through average-based measures.

methods involving the statistical average may be inappropriate assessments for population-level dynamics.

The arithmetic mean is a poor measure for genomically unstable cell populations

Genome heterogeneity that is mediated by genomic instability has significant biological consequences. Elevated genome heterogeneity leads to increased evolutionary potential, as evidenced by the instability-mediated heterogeneous single cell growth rates and genome instability-mediated transcriptome heterogeneity (Stevens et al. 2014). This high level of genome heterogeneity warrants closer attention, as it presents a problem for most current technical and analytical methods used to characterize cell populations. For example, using average-based research methods may lead to results that do not accurately characterize a cell population where genome instability is high.

To demonstrate the challenges posed by using average-based methods for unstable cell populations, the daily proliferation of all single-cell derived colonies in HCT116 cells and unstable Sub1 cells were averaged and compared against the actual single cell growth profiles (Figure 13). Single cell growth of unstable Sub1 cells manifest in a non-normal growth distribution (Shapiro-Wilkes normality test, $p \leq 1.0 \times 10^{-5}$), while stable HCT116 cells exhibited a normal growth distribution (Shapiro-Wilkes normality test, $p \leq 0.5$). Growth in unstable Sub1 cells was significantly more diverse, as total colony growth had a much larger range than stable cells. In stable HCT116 cells, each colony roughly contributed the same proportion of cells to the overall population. However the dynamics changed for unstable Sub1 cells, as only a few cells were very highly proliferative and were responsible for generating most of the population growth. For example, one single-cell derived colony in Sub1 cells was responsible for generating 70% of sample growth, as compared to stable HCT116 cells where each colony contributed no more than

10% of overall growth. This indicates that average profiles are not suitable for cell populations with a high degree of genome level heterogeneity (Figure 13). In unstable Sub1 cells, the arithmetic mean (AM) is measured at 73 cells per colony and day 6 proliferation ranged between 1-593 cell per colony. A majority of cells sampled fell well short of the 73 cell/colony average because the highly proliferative cells, or in other words, the AM is much greater than the median. In contrast, stable HCT116 cells had a 41 cell per colony average, and total proliferation ranged between 8-82 cells.

Discussion

Most would agree: heterogeneity is a well-accepted feature of cancer cells. Despite its recognition, it is rarely systematically documented and discussed much less for its impact on data presentation. The data presented here demonstrates the existence of heterogeneity at extremely high levels, where single cells of unstable cell populations are very difficult to clone. In addition, a high degree of karyotype heterogeneity has significant effects on cell population dynamics. Karyotype heterogeneity leads to drastic growth heterogeneity, which is in stark contrast to the generally accepted concept that all cells in a cell population grow and divide at relatively the same rate. The data presented earlier clearly demonstrates that, at the population level, unstable cells exhibit different PD rates and single cell proliferation. Furthermore, most unstable cells contributed little to overall population growth, while a few highly proliferative cells generated most of the population growth. These findings clearly demonstrate that karyotype heterogeneity has significant research and clinical significance, and further indicate that a reevaluation of fundamental biological principles is warranted.

While genome-level heterogeneity is common among cancer cells, it is usually ignored or not measured. Current popular research strategies report the AM, and sometimes the geometric

mean. However means and other statistics related to the first moment $\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)$ do not quantify heterogeneity. Rather, they smooth heterogeneity, thus making it invisible. Heterogeneity is represented by the second moment $\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right)$ and its modifications, like variance, standard deviation, coefficient of variation (CV), and average variation. Some second moment indices use first and second moments, while means only use the first moment. Thus, the use of statistical means to characterize and assess genomically heterogeneous cell populations may not produce reliable measures, especially as outliers are typically unaccounted for in many average-based measures or techniques that are commonly used in biomarker discovery and drug design.

In addition, the standard deviation or standard error of the mean are commonly included in the graphical representation of data, however it is not scale-free and therefore cannot be used to compare against the variation of a different sample. In order to demonstrate how the AM is a poor measure of genomically heterogeneous cell populations, a composite karyotype of an unstable cell population was generated by averaging the frequencies of all chromosome structures (Figure 13a). The single-cell genome-level heterogeneity that was so drastically apparent is virtually nonexistent in the composite karyotype. Issues in reporting the AM also extends to growth heterogeneity in unstable cell populations (Figure 13b, 13c). For unstable cell datasets, the geometric may be a better fit as it excludes outliers and only describes the slower growing cells. However even as they may be a better fit for the data set, the geometric mean also represents a poor measure of overall population growth. For example, the geometric and arithmetic means were calculated and compared in stable and unstable cell populations (Table 2). Calculation of the geometric mean results in a 19 cell per colony average, and while this number is more representative of most of the colonies sampled, it does not account for total population growth. Statistical outliers are not

frequent events, and their occurrence may depend on probability. However, there still is considerable variation in unstable cells (Figure 11g). Therefore both statistical outliers and variation are important to population growth, and statistical

Table 2: Inadequacies of statistical mean to describe growth of heterogeneous cell populations

	Mean colony size	*Estimated total population*	*Difference from actual population*
Arithmetic mean, Heterogeneous subpopulation	73	1314 cells	8 cells
Arithmetic, HCT 116 cells	41	943 cells	7 cells
Geometric mean, Heterogeneous subpopulation	19	342 cells	980 cells
Geometric mean, HCT 116 cells	37	806 cells	99 cells

The arithmetic and geometric means were calculated for both stable and unstable cells. The arithmetic mean can be reliably used to assess stable cell populations that are clonal, but not unstable cells that exhibit a large degree of genome heterogeneity. The geometric mean was also calculated, which tends to exclude outliers in its calculation. While the geometric may describe average growth of slower growing unstable colonies, it does not reliability characterize the entire cell population.

means do not accurately characterize population dynamics in unstable cell populations.

Genome heterogeneity and genome heterogeneity-mediated growth heterogeneity have obvious significance related to basic and clinical research. The statistical mean is a measure that can be accurately used to assess cell populations that are genomically homogeneous, like normal homeostatic or developmental conditions. However, statistical means are not appropriate for

profiling cell populations with elevated genome-level heterogeneity, such as cancer cell populations or other conditions of the pathological context. During cancer macro-cellular evolution, there is no "average" cell because system heterogeneity is the key dominant feature. Using the average would generate an inaccurate assessment of the cell population, as averages eliminate heterogeneity, the key defining feature of cancer, which is reflected by the dynamic NCCA/CCA cycle and NCCA frequency. In contrast, statistical averages can be accurately used during micro-cellular evolution (Darwinian evolution) because system heterogeneity is low and most change that occurs is at the gene-level. Various evolutionary cancer models that are linear and eliminate heterogeneity can be accurately assessed by statistical averages (Horne et al. 2013). Therefore, the inappropriate use of the statistical average poses a challenge in cancer modeling and chemotherapy drug design and targeting.

Targets chosen through statistical averages may lead to increased resistance and off-target effects (Heng et al. 2011a, Horne et al. 2013, Abdallah et al. 2014). As an example, while PD rates has its inaccuracies, it is used in clinical settings to measure the growth rates of tumors (Mehrara et al. 2007). In order to bypass the PD and its inaccuracies, a number of other indices have been proposed to measure tumor growth. While some may be superior to PD rates, they still employ the average and stationary growth models that make them fall short of accurately assessing growth heterogeneity of tumors (Mehrara et al. 2007). Rather, the coefficient of variation (CV) and other indices such as standard deviation or variance, should be employed under conditions when genome-level heterogeneity is high. In addition, techniques that measure the average cell across an entire cell population, which include genome sequencing, RNA-sequencing, expression profiling, western blotting, etc., are commonly used in biomarker identification and drug targeting. An increased number of studies are moving to single cell platforms, however, efforts like the

TCGA, the Cancer Genome Sequencing Consortium, and other sequencing projects continue their use of average-based profiling. Applications of these strategies may exhibit some initial success, but long-term monitoring of tumor growth demonstrates that the tumors regrow, suggesting that only more dominant subpopulations are profiled through average-based methods, thus excluding from the initial assessment the outlier cells and more rare subpopulations that may play a more significant role in the long-term evolution of cancer (Abdallah et al. 2014).

It is difficult to investigate single cell dynamics using *in vivo* model systems. *In vitro* experimental systems represent effective strategies to systematically investigate single cell dynamics under various conditions for variable time periods. It has been often noted that the results presented may in fact be due to artifacts of cell culture, and the use of *in-vivo* model systems or human *ex vivo* samples may reveal different findings. This is not the case, as the results presented in this chapter that reflect the dominance of genome-level heterogeneity in cancer have been confirmed by multiple genome sequencing cancer projects and other studies (Navin et al. 2011, Baca et al. 2013). For example, single cell sequencing completed on 100 breast cancer cells revealed that cancer evolution is punctuated, not stepwise (Navin et al. 2011). Individual cells exhibited copy number changes and a large number of gene mutations that were not conserved. Finally a study by Baca et al sequenced prostate cancer tumors to demonstrate that prostate cancer evolves through a stochastic process called chromoplexy (Baca et al. 2013). Chromoplexy is described as the genome-wide reorganization of large segments of the genome. These studies have confirmed the two phases of cancer evolution that was first described based on karyotype analysis (Heng et al. 2006c, Heng et al. 2009, Ye et al. 2009, Heng et al. 2010).

Genome-level heterogeneity is also commonly observed in normal healthy tissue (Iourov et al. 2008). As an example, normal mammalian livers are found to exhibit a large degree of

polyploidy than *in vitro* systems that range from 2N-16N (Duncan et al. 2010, Duncan et al. 2012). Furthermore, the chimeric genome has been associated with a number of human diseases (Iourov et al. 2008, Heng 2013c, Heng et al. 2013a, Campbell et al. 2015). NCCAs are also present in normal human lymphocyte cultures (Heng et al, in preparation). While normal tissues do only exhibit a low degree of genome-level aberrations, NCCAs are elevated under disease conditions. NCCA frequency in human lymphocytes of a number of diseases including cancer, ranged from 20%-40% (Heng et al, in preparation). Furthermore, the monitoring of NCCA frequency in a number of mouse cancer models revealed a high level of genome heterogeneity in each model (Ye et al. 2009). Under normal physiological or homeostatic conditions where the NCCA frequency is low, the statistical average remains an appropriate measure for the characterization and assessment of cell populations. However, the average is ill-suited for conditions where genome-level heterogeneity is high.

The data presented earlier in this chapter challenges the notion that cell populations are mostly isogenic. For the most part, cancer cell populations are not isogenic, as multi-level heterogeneity and especially genome-level heterogeneity are key features of the disease that have significant effects on cancer evolution. A direct link between genome-level heterogeneity and systems-heterogeneity, such as growth heterogeneity, may exist. Heterogeneity exists at different molecular levels, including gene mutations (Lawrence et al. 2013), transcription (Stevens et al. 2013, Stevens et al. 2014), biochemical signaling pathways (Chakraborty and Roose 2013, Hartzell et al. 2013), the tumor microenvironment and the response to drug treatment (Blagosklonny 2006). It has been difficult to assess the relationship between the different types of heterogeneities as most studies do not address genome-level heterogeneity. The genome represents the highest level of genetic organization, and the genome package is responsible for punctuated macroevolution of

cancer, therefore the effort to unify genome-level heterogeneity to systems-heterogeneity is highly significant. This type of analysis may also be used to profile single dominant cells for drug treatment (Blagosklonny 2006).

Finally, a high degree of genome-level heterogeneity may also provide some explanation to the challenges faced with reproducibility (Heng 2013a). While the same cell lines may be used, the genomes may be altered, thus leading to different results. This has been demonstrated by our study that links genome heterogeneity transcriptome heterogeneity (Stevens et al. 2013, Stevens et al. 2014). Many different mouse models have been engineered to investigate the functions of specific genes or dysregulated pathways (Hartzell et al. 2013). However, while the dysregulation of a particular pathway can be highly penetrant (due to any gene member being mutated), pathway switching remains a dominant feature of unstable cancer cells that can be exploited to acquire resistance (Heng et al. 2013a, Stevens et al. 2013, Stevens et al. 2014). Elevated genome-level heterogeneity can also provide an explanation for the large disconnect between experimental systems and clinical samples. Clinical samples are often more heterogeneous and complicated due to increased genome heterogeneity, especially compared to homogenous mouse strains that are inbred for several generations. Taken together, genome-level heterogeneity is a major player of cancer, as genome replacement represents the key feature of cancer evolution and is responsible for the initial growth of tumors, progression, metastasis and drug resistance. A large portion of conventional thought regarding cancer were originally formulated using average-based methods are in need of re-evaluation, as the average cancer cell does not exist.

CHAPTER 4: DETECTION OF FUZZY INHERITANCE

Introduction

The role of heterogeneity is especially significant in several biological processes, which includes the cancer evolutionary process (Heppner 1984, Rubin 1984, Brock et al. 2009, Heng 2009, Marusyk et al. 2012, Heng 2013b, Heng et al. 2013a, Huang 2013). In cancer, genetic/nongenetic heterogeneity is overwhelming, as it exists at multiple levels including the gene, epigenetic and genome levels (Brock et al. 2009, Heng et al. 2011a, Huang 2013). Furthermore, due to the more dominant effect that genome-level changes have on an individual cell, the genetic blueprint of a system is represented by karyotype-defined system inheritance, as opposed to gene-level parts inheritance, thereby calling for the redefinition of inheritance (Heng 2009, Heng et al. 2011a, Heng et al. 2013a).

The stochastic generation of NCCAs, especially in cancer cell populations where they are the dominant form of heterogeneity, creates a dilemma for the understanding how heterogeneity is regulated, and conventional mechanisms of precise inheritance in somatic cell evolution. According to conventional views, genetic variation is largely generated by errors at low frequencies, followed by clonal expansion of cells with the accumulated genetic change. Here, inheritance mechanisms are precise, as a mother cell passes the same karyotype to its daughter cell, and accumulated genetic change is traceable. However, a high degree of genome-level heterogeneity has been documented and is inconsistent with mechanisms of precise inheritance. In fact, the degree of heterogeneity reaches such high levels that it cannot be accounted for by Darwinian evolution (Ling et al. 2015). In most cases, genetic change is not conserved and the degree of heterogeneity is the only consistent feature. Specifically, most cancer cell populations exist in the punctuated macro-cellular evolutionary phase where no specific genome dominates and NCCAs are significantly high. The constant generation of NCCAs indicates that system

inheritance remains discontinued, as an identical genome or karyotype is not passed from mother cell to daughter cell (Heng et al. 2006c, Heng HQ 2016).

Genetic inheritance can be defined as the passing of bio-processes from a mother cell to a daughter cell. Here, a daughter cell inherits material that carries genetic information from the mother cell. However, conventional genetic inheritance mechanisms are challenged by data generated from natural rather than lab settings: first, while a vast majority of cells can inherit an identical karyotype from their mother cells, a given number of cells exhibit de-novo NCCAs that are generated in each cell division; second, during punctuated macro-cellular evolution, a majority of daughter cells do not inherit the same genome as the mother cell, rather daughter cells all exhibit drastically new genomes. If the same karyotype is not inherited, but a given degree of heterogeneity is inherited, there must be a mechanism that permits cells to inherit altered system inheritance. This new mechanism of inheritance is unlike precise inheritance because there is no precision to maintain a specific genome, particularly in the punctuated macro-cellular evolutionary phase.

What evolutionary advantage is associated with passing heterogeneity, or NCCAs, as opposed to stable, clonal genetic change? Cell populations that exhibit elevated NCCA frequency have heightened evolutionary potential, as changes to the karyotype alter the genome context (gene content plus genomic topology) (Heng 2009). During selection events, NCCA frequency significantly increases and remains elevated until a stable cell(s) is selected (Heng et al. 2006a, Heng et al. 2006c, Heng et al. 2011a). After the cell population stabilizes, NCCAs are still present but at a lower frequency that remains constant over multiple passages. The relationship of NCCAs to evolution and their continued presence in a cell population at a stable frequency led to the question, how do cell populations maintain their degree of heterogeneity?

A number of cancer evolution studies conducted using *in vitro* evolutionary models have demonstrated that, while the final karyotypic products or specific chromosomal aberrations, are different, the relationship between NCCAs and CCAs remains the same (Heng et al. 2006c, Lawrenson 2010). Specifically, in different evolutionary runs, the frequency of NCCAs was stable in parallel stages of evolution. Therefore, it is hypothesized that a new form of somatic cell inheritance exists, termed fuzzy inheritance, which regulates and maintains the degree of cell population heterogeneity. Specifically, a given degree of heterogeneity is passed over generations, as opposed to specific changes, and can be seen at the cell population level. Furthermore, the passing of heterogeneity functions to maintain the evolvability of the cell population.

In order to test the mechanism of fuzzy inheritance, a panel of single cells were isolated and their inherited traits were monitored, including karyotype and growth rate, over multiple generations at the single cell level. It was discovered that unstable cell populations and stable cell populations drastically differ regarding how their inherited traits are passed, and ultimately suggested that genome status, or the stability of the genome, is what directs heritable population behavior. First, it is demonstrated that, when the genome is unstable, specific karyotypes cannot be inherited, but a similar degree of karyotype heterogeneity can be passed from generation to generation. In contrast, stable genomes can be passed over multiple generations with high fidelity, and are accompanied with a lower frequency of NCCAs. Single cell growth analysis found that in unstable cell populations, growth rates were significantly more heterogeneous than stable cell populations. Most cells grew at slow or a modest rate while only a few cells were highly proliferative and drove overall population-level growth. Furthermore, a similar degree of karyotype and growth heterogeneity were restored from a single cell, regardless of the initial karyotype or growth rate. Finally, it is demonstrated that single cells can continue to restore

karyotype and growth heterogeneity even after multiple generations. Altogether, these data support that a new type of inheritance exists where, in unstable cell populations, rather than specific changes being passed, a similar degree of heterogeneity is passed, and is seen at the cell population level. This is termed fuzzy inheritance, which can offer an understanding towards cancer evolution and drug resistance by explaining how cell populations retain their evolvability.

Materials and Methods

Cell lines: Conditionally inactivated Brca1/p53 isolated from the surface epithelium of C57/BL6 mouse ovaries were a generous gift from Dr. Barbara Vanderhyden at the University of Ottawa (Clark-Knowles et al. 2007). Cells were maintained in low glucoses DMEM supplemented with 10% FBS and antibiotics. Standard cell culture procedures were used to maintain cell lines. Serial dilutions were performed to isolate a panel of single cells 60 days after ovary resection from mice. HCT116 cells and HCT116-E6 cells were a generous gift from Dr. Bert Vogelstein (Lengauer et al. 1997). HCT116 cells were maintained in RPMI supplemented with 10% FBS and antibiotics. HCT116-E6 cells were maintained in low glucose DMEM supplemented with 10% FBS, antibiotics and G418. Standard cell culture was used to maintain cell lines.

Mefs (mouse embryonic fibroblasts) were a gift from Dr. David Chen from the University of Texas. Cells were transformed using SV40. Cells were maintained in low glucose DMEM supplemented with 10% FBS and antibiotics. Standard cell culture was used to maintain cell lines.

Cytogenic metaphase slide preparation and spectral karyotyping (SKY): Cytogenetic slides were prepared as previously described (Heng et al. 1992).

SKY was completed on metaphase slides according as previously described (Ye et al. 2009).

Karyotypic analysis: Karyotypic analysis was completed as previously described. All normal and abnormal chromosome structures were represented on karyographs (Nicholson and Duesberg

2009). Karyographs are a visual presentation of single cell genome-level heterogeneity. The x-axis represents chromosome structures (normal and abnormal), while the y-axis represents the frequency of each structure. Each line represents a single cell.

Statistical analysis: All chromosome structures (normal and abnormal) are measured in order to determine karyotype heterogeneity. Power calculations were completed as previously described in the methods sections of chapters 2 and 3 (Abdallah et al. 2013).

In situ single cell counting: Single cell growth variation experiments were conducted in gridded T-25 cell culture flasks, as previously described.

Isolation of fast- and slow- growing subcolonies for generation experiments: Heritability of karyotype heterogeneity and single cell growth heterogeneity was completed by selecting fast- and slow- growing colonies over five generations using *in vitro* cell culture techniques (Figure 21). *In situ* single cell growth was first performed in unstable Sub1 cells, HCT116 cells, and stable conditionally inactivated Brca1 mouse ovarian epithelial cells as previously described. Cell growth was monitored for 6 days and colony population sizes were enumerated. Fast- and slow- growing subcolonies were identified for isolation, and then isolated on day 6 using 0.25% trypsin (Mediatech) and sterile cloning rings. After 1-2 weeks of continuous cell culture, cells were harvested for chromosome analysis and *in situ* single cell counting was performed. Each isolation of a fast- and slow- growing colony constituted a generation. This procedure was completed for two generations in HCT 116 cells and 5 generations in unstable Sub1 cells and stable Brca1 knockout cells.

Results

Conceptual considerations

It is first necessary to understand why karyotype change is compared over multiple cell passages to illustrate the mechanism of fuzzy inheritance. According to the Genome Theory, each specific karyotype represents a unique system, as changes to the genome alters the heritable genome context. Because the karyotype is the blueprint that defines a biological system, a mother cell that passes the same karyotype to its daughter cell is passing the same system inheritance. While an identical karyotype is passed, there is still some variation at the parts level (example, gene-level) that can slightly alter or fine tune the system. However, system inheritance is altered if the mother cell passes a different karyotype. Because changes to genome topology alters genome context, a mother cell that passes an altered system inheritance to its daughter cell are both considered different systems. Therefore, the passing of altered system inheritance can be illustrated by monitoring the pattern of karyotype change. In a similar fashion, the pattern of cell growth, cell death, and other inherited traits can also be compared.

It is also necessary to measure the degree of heterogeneity, which can be achieved by measuring stochastic variation. At the genome level, stochastic variation exists in the form of NCCAs (Heng et al. 2006a, Heng et al. 2006c, Heng et al. 2011a), therefore NCCAs are used to measure genome heterogeneity. At the gene level, gene mutation heterogeneity can be measured by the frequency of random gene mutations. In a similar fashion, growth heterogeneity can be measured by the differences in single cell growth rates.

Karyotype heterogeneity is heritable

Heterogeneity is a key feature of cancer and is increasingly studied to understand clonal evolution in cancer (Roschke et al. 2002, Heng et al. 2009, Nicholson and Duesberg 2009, Hanahan and Weinberg 2011, Gerlinger et al. 2012, Greaves and Maley 2012, Huang 2012). Most cancers exist in the unstable macro-cellular punctuated phase of evolution where most genome-level

changes are not clonal. Understanding how genetic and nongenetic changes in a single cell affects an entire cell population is a challenging task. In the natural situation, the fact that most cancer cell populations are not clonal is inconsistent with conventional theories of clonal evolution and precise inheritance mechanisms. It is thus necessary to reevaluate the process of cancer evolution, and specifically the role and significance of stochastic genome-level changes.

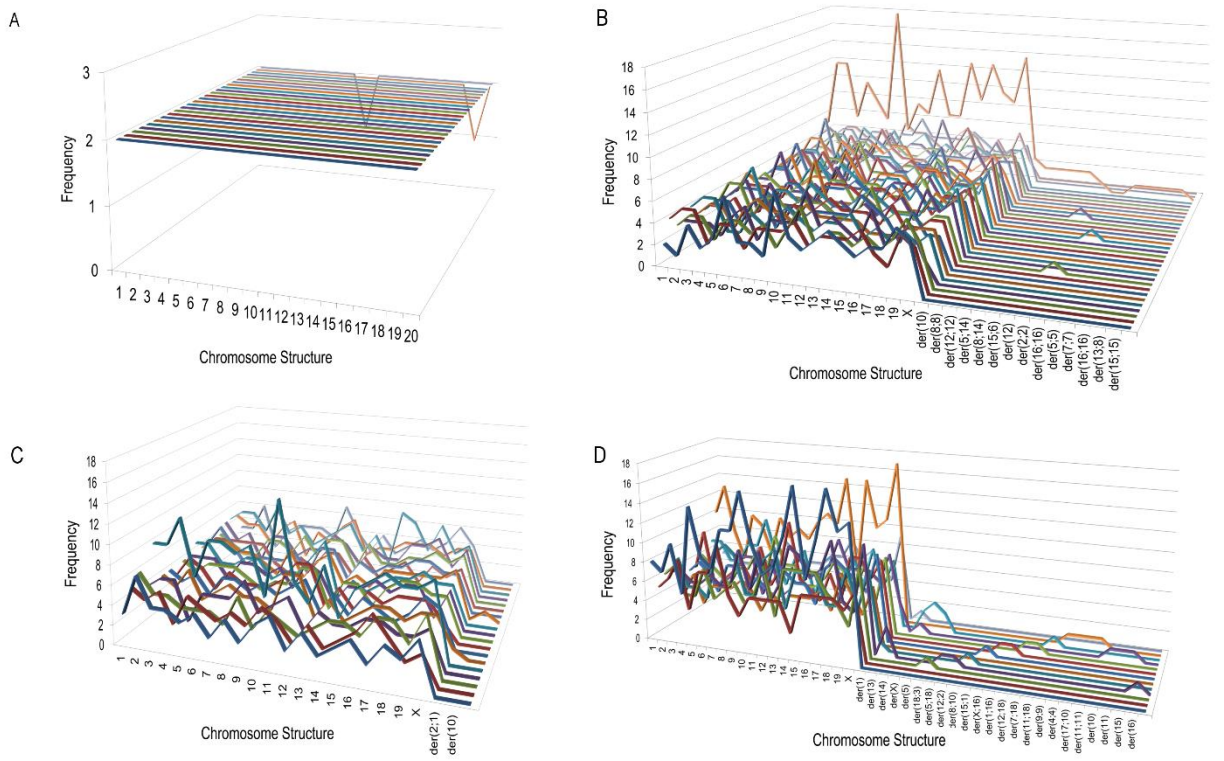
In order to monitor the pattern of genome change during cancer evolution, pure populations of cells were made by cloning single cells from spontaneously transformed, conditionally inactivated Brca1/p53 mouse ovarian surface epithelial (MOSE) primary cells (Clark-Knowles et al. 2009). Isolation of single cells is a preferred method for generating isogenic cell populations, as genetic changes should be minimal in a short period of time. By using SKY, karyotype changes were monitored in order to determine the degree of genome-level change. Within 4-8 weeks of continuous, standard cell culture conditions, all single-cell derived pure populations were found to exhibit a high degree of genome-level heterogeneity (Figure 14). The data also revealed that no two cells had identical genomes. Furthermore, there were no shared karyotypic abnormalities between the parent cell population and two single-cell derived subpopulations, as all cell populations exhibited the same, high degree of novel karyotype change.

The experiment was repeated in unstable human colorectal cancer cells, HCT116-E6 line. Two single-cell derived subpopulations were developed, and each subpopulation exhibited approximately the same degree of shared and novel change as the parent cell population, indicating that the same degree of karyotype change is inherited (Figure 15). These findings demonstrate that the degree of genome-level heterogeneity found in the single-cell derived subpopulations was similar to the parent cell population before cloning, and further suggests that the same degree of heterogeneity can be maintained without passing a specific abnormal karyotype.

Because all cells analyzed exhibited unique genomes, it was concluded that the genomes were altered during cell division. Additionally, the subpopulations all originated from single cells and were under continuous culture conditions for only a short time period, suggesting a single cell that originates from an unstable cell population can generate population-level heterogeneity by passing altered karyotypes to its daughter cells. This conclusion conflicts with conventional views on how heterogeneity is established and maintained, where both normal and abnormal cells will pass that same normal and abnormal karyotypes, respectively, to its daughter cell. In both scenarios, heterogeneity of the cell population is maintained. Interestingly, subclones isolated from karyotypically stable HCT116 cells (which contains an abnormal karyotype) have the same karyotype as the parent cell population. Even after long-term continuous cell culture, HCT116 cells maintained the same karyotype (Shih et al. 2001, Knutsen et al. 2010).

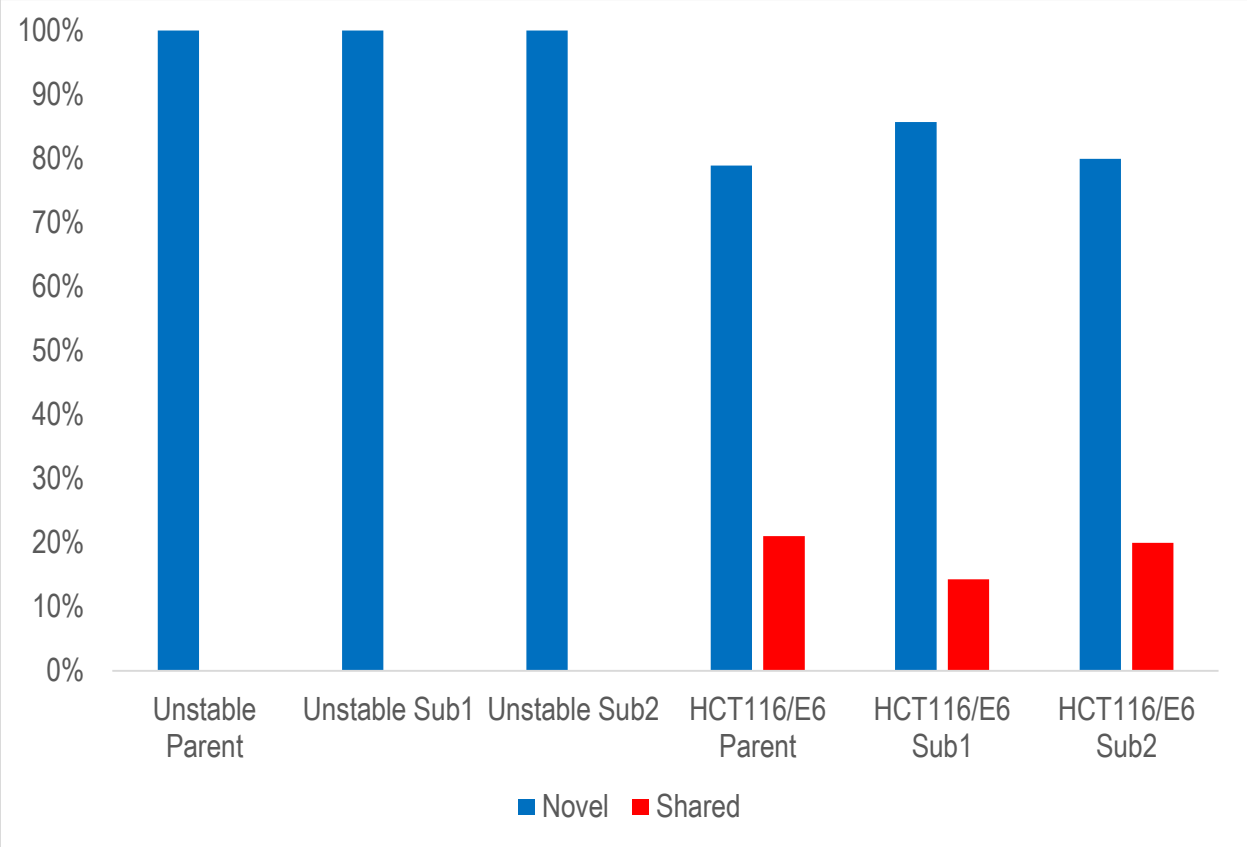
These findings are in agreement with previous studies that demonstrate that, during punctuated macro-cellular evolution, karyotypic change is stochastic, as the same karyotype cannot be inherited (Heng et al. 2006c). Stochastic, punctuated evolution has been confirmed by cancer genome sequencing studies (Navin et al. 2011, Baca et al. 2013, Wang et al. 2014). Ultimately, these results suggest that heterogeneity is a heritable feature, as a specific karyotype was not passed and heterogeneity was restored from a single unstable cell. This also occurs during punctuated evolution, where cell populations are unstable new genomes are generated at

Figure 14: Single cells pass karyotype heterogeneity to daughter cell populations



Karyotype heterogeneity is inherited from mother cell to daughter cell population. A) SKY karyograph of early passage wild type mouse ovarian surface epithelial cells. As expected, most cells exhibit normal karyotypes. B) Karyograph of the parent cell population, conditionally inactivated Brca1/p53 mouse ovarian surface epithelial cells which exhibits high genome heterogeneity. Each cell exhibits unique NCCA frequency. C-D) Karyograph of single cell-derived subpopulations, Sub 1 (C) and Sub2 (D). Each cell in both subpoplations exhibited unique karyotypes. No shared, clonal karyotypic abnormalities were detected in any cell population. Rather than inheriting specific karyotypic structures, karyotype heterogeneity is inherited from a single cell.

Figure 15: Inheritance of karyotype heterogeneity in HCT116/E6 cell line



Spectral karyotyping was completed on HCT116/E6 parent cells and two single cell-derived subpopulations. The number of cells that exhibit shared and novel karyotype structures are approximately the same across all three cell populations, indicating that a single cell restored a similar degree of karyotype heterogeneity, or in other words, the degree of karyotype heterogeneity is inherited.

each cell passage and identical karyotypes cannot be passed. Altogether, this suggests that conventional mechanisms of precise inheritance are unable to describe somatic cell evolution for unstable cell populations. Somatic cell inheritance is likely to be regulated through an alternate mechanism that regulates heterogeneity by passing the degree of change, or a stochastic state.

Growth heterogeneity is heritable

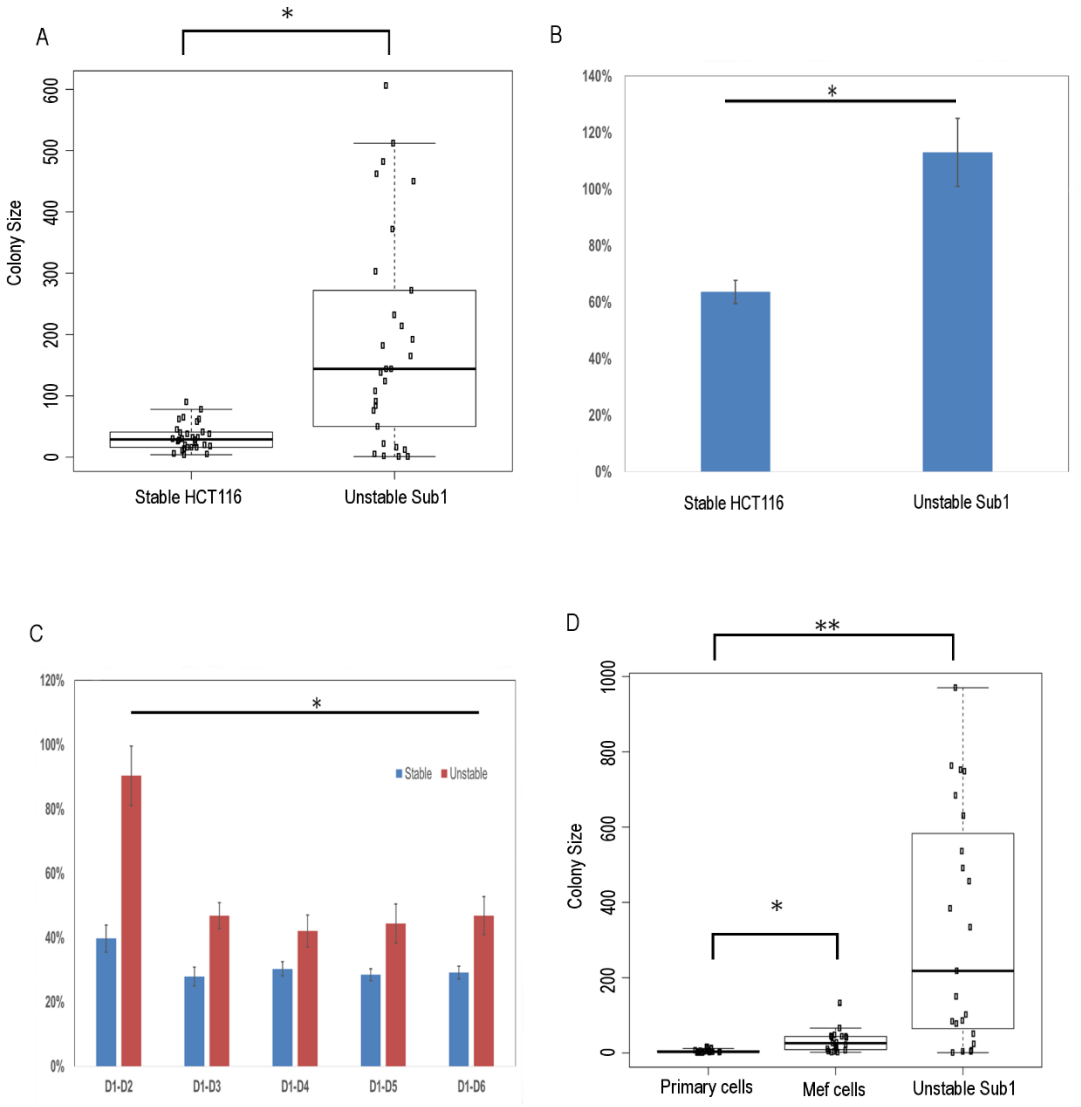
Genome heterogeneity has previously been linked with other cellular features such as transcriptome heterogeneity, tumorigenicity, cancer cell evolution, and drug resistance (Heng et al. 2006a, Heng et al. 2006c, Stevens et al. 2007, Ye et al. 2009, Heng et al. 2013a). If the degree of genome heterogeneity is heritable, then other features of cell population dynamics, such as cell growth, should also be heritable. Previous studies have demonstrated that various genetic manipulations resulted in growth heterogeneity (Kuczek and Axelrod 1987, Abdallah et al. 2013), therefore it is hypothesized that genome heterogeneity and growth heterogeneity are linked and that the degree of cell growth heterogeneity is heritable. First, *in situ* single cell growth was monitored for 6 days. Sub1 cells, which are genomically unstable, exhibited a large degree of variation in their single cell growth rates. Most cells grew at a slow or modest rate, as measured by total cell growth per single cell derived colony after 6 days, and only very few cells were highly proliferative. In contrast, karyotypically stable HCT116 cells exhibited very little variation in their single cell growth rates (Figure 16a, F-test $p \leq 1.0 \times 10^{-18}$, Figure 17b-c). If heterogeneity is a heritable feature, then it is necessary to demonstrate that heterogeneity is present from the earliest cell divisions, as opposed to heterogeneity accumulating over time. In both stable and unstable cells, daily growth rates were calculated to define that heterogeneity was present in the earliest cell divisions and remained constant for the duration of the assay (Figure 16c).

The above data suggest that a link between genome heterogeneity and growth heterogeneity exists. In order to elucidate the relationship genome heterogeneity has with growth heterogeneity, the single cell *in situ* growth heterogeneity of different cell lines with differing degrees of genome instability was compared. Highly unstable Sub1 cells were compared against moderately stable Mef cells (Figure 18) and highly stable early passage MOSE cells. The results demonstrate that cell populations with increased genome heterogeneity exhibited a higher degree of growth heterogeneity (Figure 16d, Figure 17b-d). Additionally, highly proliferative unstable Sub1 cells were significantly more proliferative and more variable compared to moderately unstable Mef cells and early passage MOSE cells. This suggests that a link between genome heterogeneity and growth heterogeneity, where increased genome instability leads to increased heterogeneity of growth.

A single cell restores heterogeneity of parent cell population

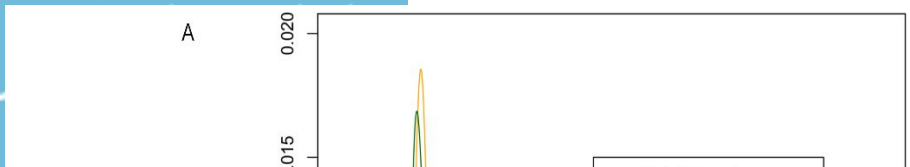
The above data demonstrated that Sub1 cells have highly diverse growth rates, and only select cells are highly proliferative. However it remains undetermined if all cells are able to recapitulate the same degree of heterogeneity in a cell population, or if this feature is reserved for select cells. In order to determine whether any cell in a population can restore the same degree of growth heterogeneity as the parent cell population, a panel of single cells were isolated from stable HCT116 cells and unstable Sub1 cells. *In situ* single cell growth experiments were carried out for each isolated cell. Growth in the unstable Sub1 cell population was expectedly heterogeneous. Single cell growth in unstable Sub1 isolates all exhibited a high degree of growth heterogeneity similar to the parent cell population. In contrast, the parent HCT116 cell population exhibited a small degree of growth heterogeneity. HCT116 isolates exhibited the

Figure 16: Growth heterogeneity is heritable



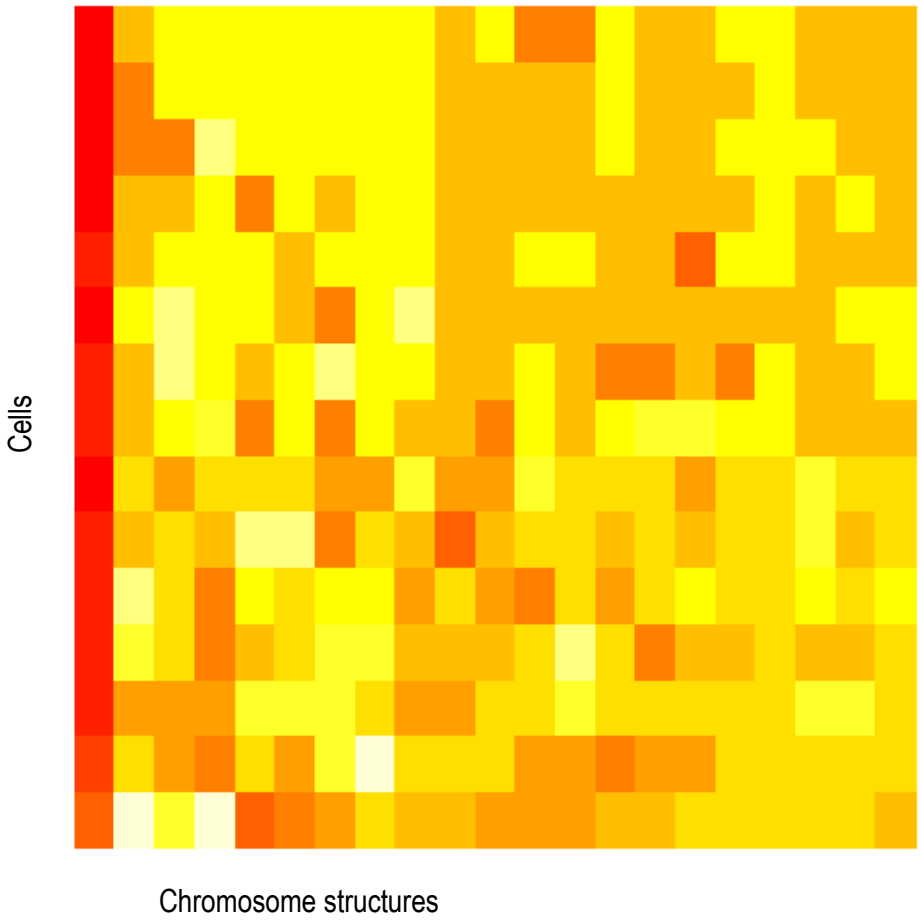
Growth heterogeneity is inherited. A) *In-situ* single cell growth in stable HCT116 cells and unstable Sub1 cells. Stable cells exhibited a uniform growth distribution, as all cells grew at approximately the same rate. Unstable cells exhibited heterogeneous growth and only few cells were highly proliferative. B) Growth heterogeneity, as measured by the coefficient of variation (CV) in stable HCT116 cells (n=3) and unstable Sub1 cells (n=3). Unstable Sub1 cells exhibit greater growth heterogeneity (t-test, $p \leq 0.05$). C) Daily growth rates were calculated in stable and unstable cells. Growth heterogeneity was present at the earliest cell divisions and remained constant during the entire experiment. D) Genome heterogeneity is linked to growth heterogeneity. Genomically homogeneous primary early passage MOSE cells exhibit minimal single cell growth heterogeneity, moderately unstable Mef cells exhibit a small degree of growth heterogeneity, and unstable Sub1 cells exhibit significantly greater growth heterogeneity (t-test, $p \leq 0.05$).

Figure 17: Replicates demonstrating growth heterogeneity of various cell populations



Kernel density estimates of cell growth for various cell populations. A) Kernel density estimates of parent cells and single cell derived subpopulations in both stable HCT116 and unstable conditionally inactivated Brca1/p53 mouse ovarian surface epithelial cells. B) Kernel density estimates of stable HCT116 cells, displaying uniform cell growth among replicates. C) Kernel density estimates of unstable Sub1, displaying bi-modal heterogeneous growth distributions among replicates. Most cells grew slowly, while only few cells grew at a fast rate. D) Kernel density estimates of moderately stable Mef cells. E) Kernel density estimates of early passage wild type MOSE cells, displaying uniform cell growth among replicates.

Figure 18: SKY of moderately unstable Mef cells



Heatmap karyograph of Mef cells. Mef cells exhibit mostly numerical NCCAs (aneuploidy), conferring moderately unstable genomes.

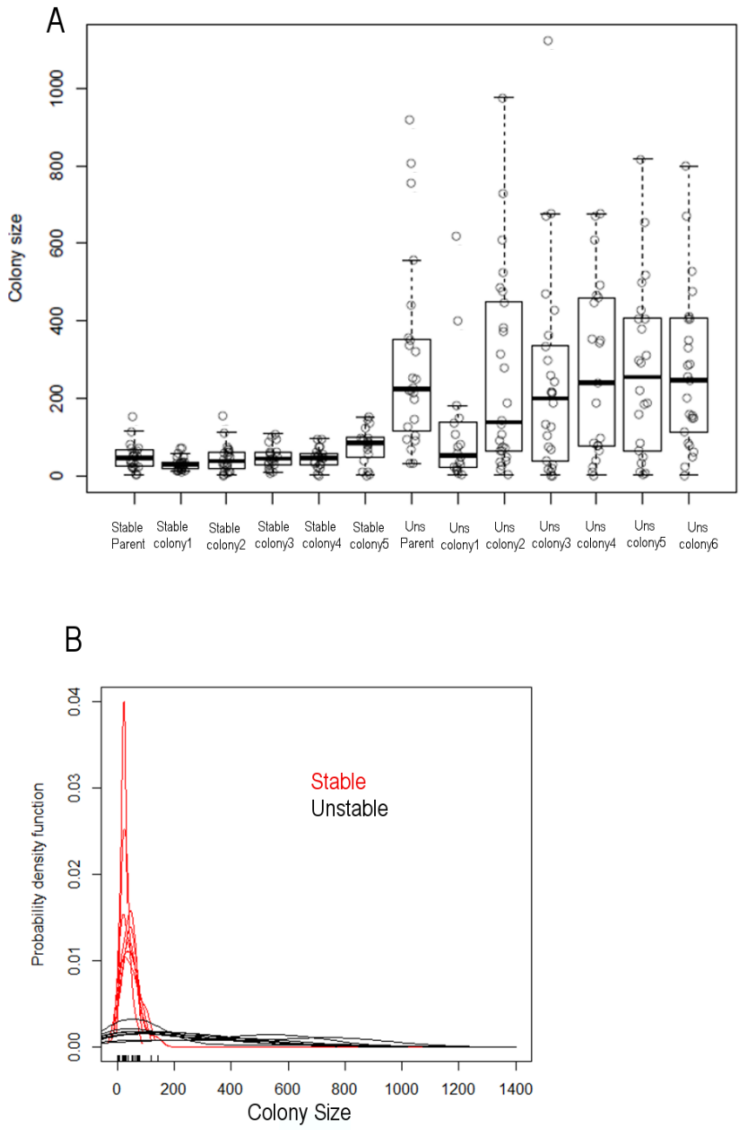
same degree of minimal growth heterogeneity found in the parent cell population. Given the considerable difference in the degree of growth heterogeneity between each cell line, the data also suggests that each cell population has different degrees of evolutionary potential.

Heterogeneity can be inherited over multiple generations

The previous data suggested that any cell can restore the same degree of growth heterogeneity as the parent cell population. In order to determine whether a specific growth rate or growth heterogeneity can be passed down to future cell populations, fast- and slow- growing subcolonies were isolated from Sub1 cells and HC116 cells, and *in situ* single cell growth experiments were performed (Figure 20a). Interestingly in unstable Sub1 cells, fast- and slow-growing subcolonies did not pass on their respective growth rates. Fast- and slow- growing subcolonies both restored the degree of growth heterogeneity found in the parent cell population, generating cells with a multitude of growth rates. Fast- and slow- growing isolates of HCT116 cells both displayed single cell growth profiles with significantly less variation. In addition, the growth profiles in both fast and slow subcolonies was similar to the parent cell population. (Figure 20b, Figure 18a).

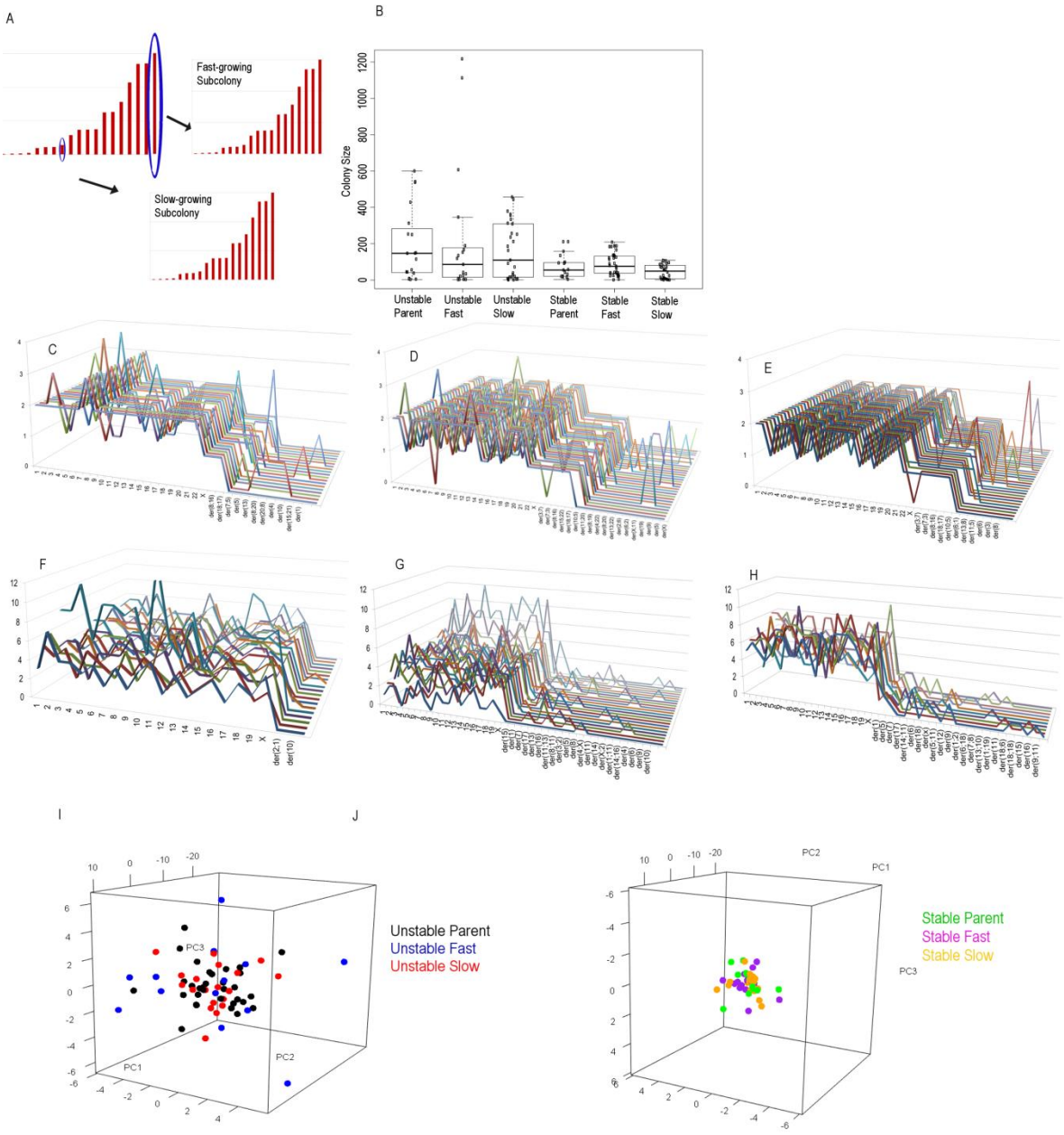
Spectral karyotyping was performed on fast- and slow- growing subcolonies of Sub1 and HCY116 cells in order to understand the genomic mechanism of the restoration of growth heterogeneity. Fast- and slow- growing subcolonies of unstable Sub1 cells exhibited a high degree of growth heterogeneity that is similar to the parent cell population. No two cells were identical and no direct karyotypic intermediates could be traced. On the other hand, fast- and slow- growing subcolonies of stable HCT116 cells exhibited a minimal degree of karyotypic change. Cells exhibited very similar genome profiles and many shared karyotypic changes were

Figure 19: Each colony restores cell growth heterogeneity



Each single cell is capable of restoring cell population growth heterogeneity. *In-situ* single cell growth experiments were conducted in stable HCT116 cells and unstable Sub1 cells. As expected, stable parent cells exhibited a uniform growth distribution and unstable cells exhibited heterogeneous growth. In order to demonstrate that each cell is capable of restoring growth heterogeneity, each single-cell derived colony was extracted and single cell growth was tested. In stable HCT116 cells, each single-cell colony exhibited uniform growth, similar to parent cells. In unstable Sub1 cells, each single-cell derived colony exhibited the same heterogeneous growth as the parent cells. Altogether this indicates that each cell is capable of restoring growth heterogeneity.

Figure 20: Genome and growth heterogeneity restored

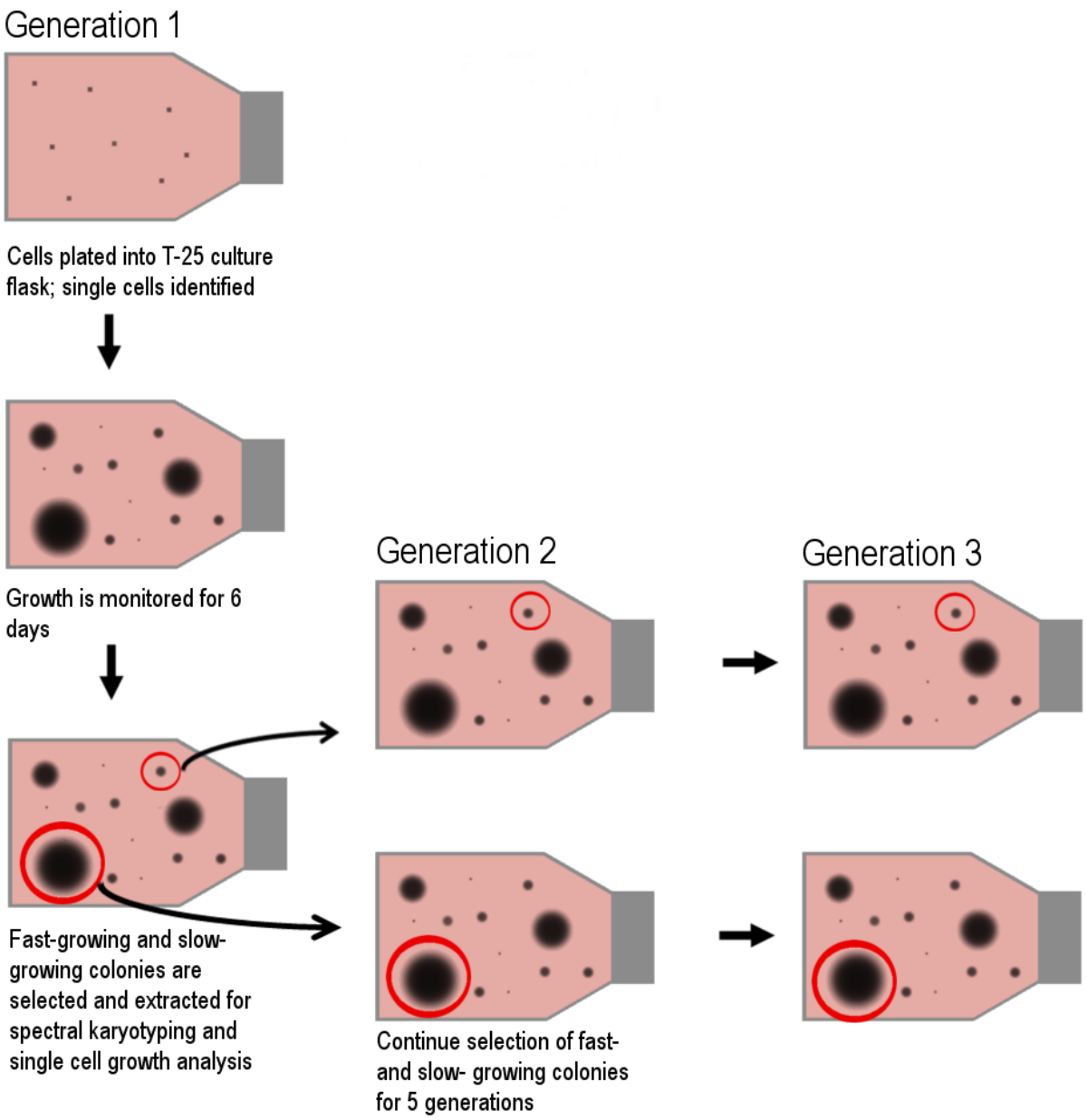


Restoration of genome and growth heterogeneity. A) Diagram of experiment design. *In-situ* single cell growth experiments are conducted, and fast- and slow- growing subcolonies are isolated to determine whether they can restore growth heterogeneity or specific growth rate. B) Results of restoration of growth heterogeneity. In stable cells, both fast- and slow- growing subcolonies restored a similar degree of growth heterogeneity (homogeneity). In unstable cells, both fast- and slow- growing cells restored the same heterogeneous growth profile as the parent cell population. SKY analysis was also completed on all subpopulations. The parent stable cell population (C) was mostly clonal with little genome heterogeneity. A similar degree of genome heterogeneity was restored by fast-growing (D) and slow-growing (E) subcolonies. Unstable parent cells (F) exhibited a high level of genome heterogeneity that was similarly restored in fast-growing (G) and slow-growing (H) subcolonies. PCA analysis was completed on unstable (I) and stable (J) cell populations. In unstable cells, single cells were distantly spaced in each of the cell populations, indicating a large degree of variation and heterogeneity. In contrast, stable cells (J) were tightly clustered, indicating a small degree of variation between individual cells.

present in all cell populations (Figure 20c-h). To determine whether the amount of variation between the parent cell population was similar the fast- and slow- growing subcolonies, PCA analysis was completed (Figure 20i-j). Single cells of the unstable Sub1 parent population and both fast- and slow- growing subcolonies were similarly spaced far apart, indicating a wide degree of variation and heterogeneity. In contrast, stable HCT116 parent and fast- and slow- growing subcolonies were more tightly clustered, indicating a much lesser degree of heterogeneity and more similar relatedness between cells.

The heritability of genome heterogeneity and growth heterogeneity was further confirmed by isolating fast- and slow- growing single cell-derived colonies over multiple generations in unstable Sub1 cells and in a stable, MOSE single cell derived subline conditionally inactivated for Brca1. *In situ* single cell growth heterogeneity was profiled at every generation (Figure 21). Briefly, single cell *in situ* growth experiments were perfumed; this first experiment is referred to as generation 1 (G1) or parent generation. After 6 days, fast- and slow- growing colonies were identified, isolated, and cultivated. These isolated colonies were referred to as generation 2 fast

Figure 21: Diagram of fast-and slow- selection experiment for multiple generations



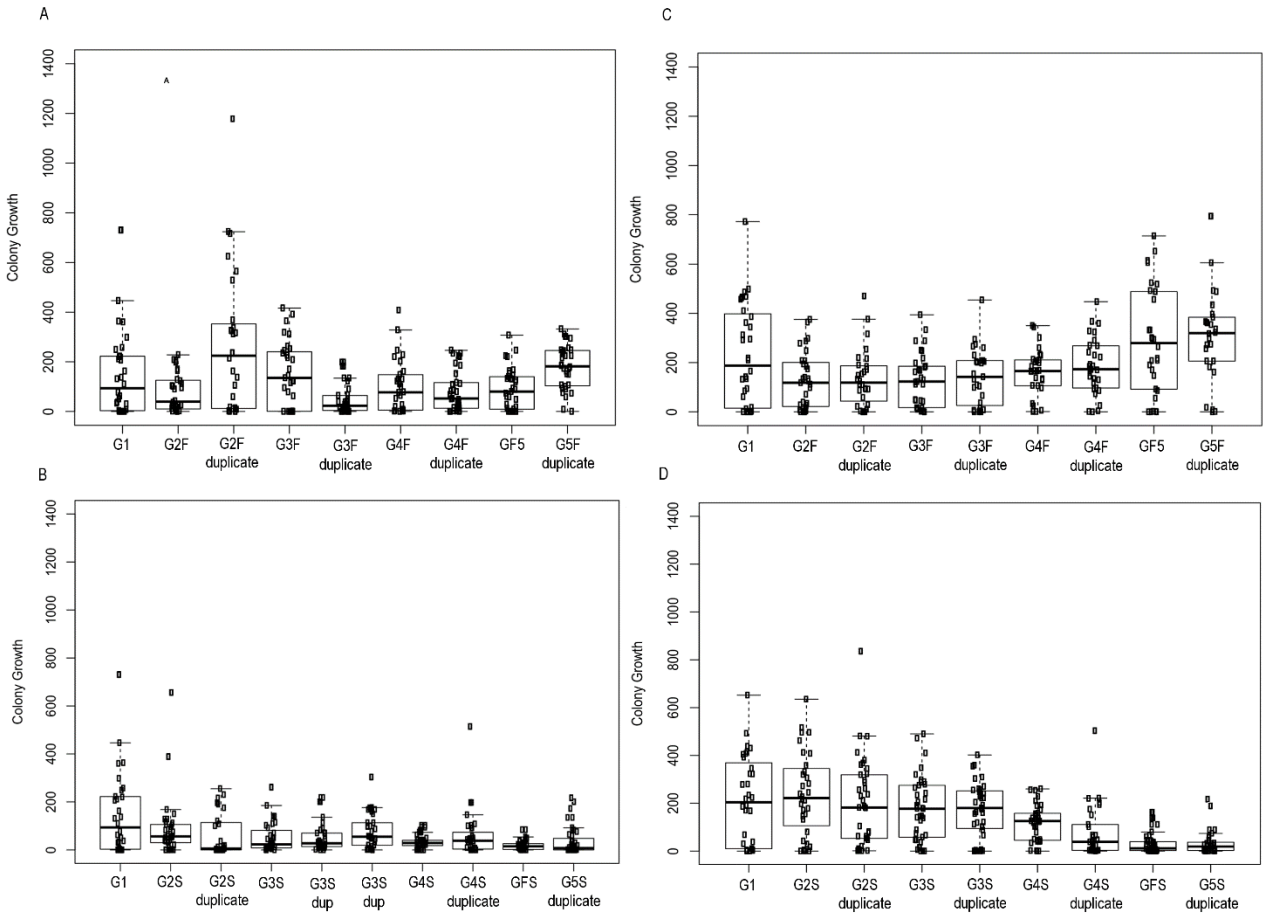
Fast- and slow- selection experiment. *In situ* single cell growth was performed on unstable and stable cells; this first growth experiment is referred to as generation 1 (G1). After 6 days of growth, fast- and slow- growing colonies were identified and isolated, to form their own subpopulations. These subpopulations are referred to as generation 2 fast (G2F) or generation 2 slow (G2S). Selection continued for five generations.

(G2F) or generation 2 slow (G2S). Selection of fast- and slow- growing colonies continued for five generations. In unstable cells, the continuous selection of fast- and slow- growing single cell-derived colony resulted in the restoration of growth heterogeneity instead of the selection of a specific growth rate, as each respective population consistently generated fast- and slow- growing cells regardless of the initial growth rate (Figure 22a-b, 23a-b). In stable cells, the continuous selection of fast- and slow- growing cells led to the selection of an increased or decreased growth rate after multiple successive selections. This suggests that a growth rate can be selected in stable cells (Figure 22c-d, 23a-b). Spectral karyotyping was next performed after 5 selections to determine if a specific genome or genome heterogeneity is passed in stable and unstable cells. In unstable cells, fast- and slow- growing cells after five successive selections restored the degree of genome heterogeneity found in the parent population. The frequency of novel structural chromosomal change specific to only a particular generation was the same across all cell populations and was significantly greater than shared structural chromosomal change (Figure 24). The dynamics change in stable cell populations. In stable cells, shared structural chromosome change significantly outnumbered novel chromosome change and remained consistent across the multiple generations (Figure 24).

It is interesting to note that when the previous experiment was repeated in stable cells, the growth rate saw an initial increase over the first three generations similar to the first trial. However, the growth pattern changed at the fourth and fifth selections of fast-growing cells. Rather than the growth rate steadily increasing, single cell growth became heterogeneous and resembled the growth profile of an unstable cell population (Figure 25a-b). The increased growth heterogeneity was accompanied by an increase in the growth CV (Figure 26b). Because genome heterogeneity and growth heterogeneity are linked, this suggests that the genome instability of the cell population

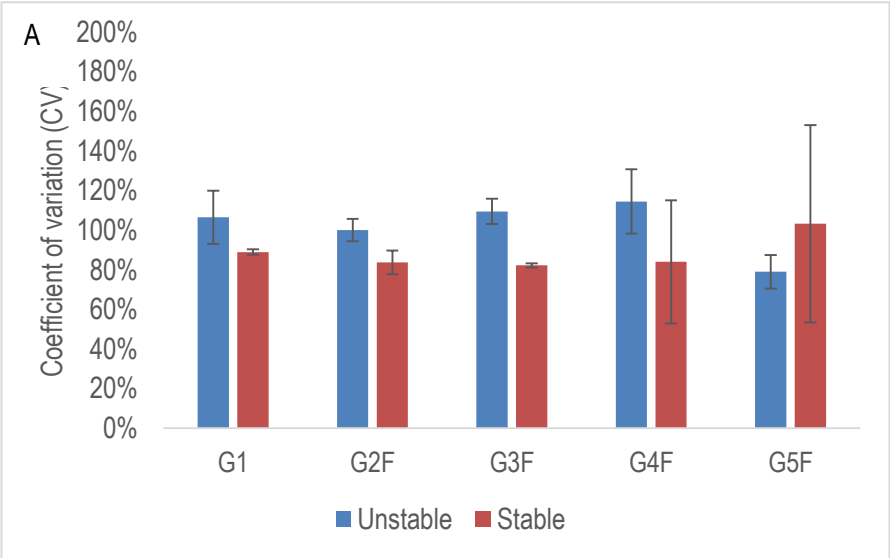
had slightly increased. To confirm whether genome instability had increased, spectral karyotyping was completed on generations 1 and 5 in both replicates (Figure 26a). After 5 selections, SKY analysis showed a higher percentage of novel structural chromosome change in run 2, suggesting the genome status had changed to become slightly more unstable and thus altering the degree of heterogeneity that can be inherited.

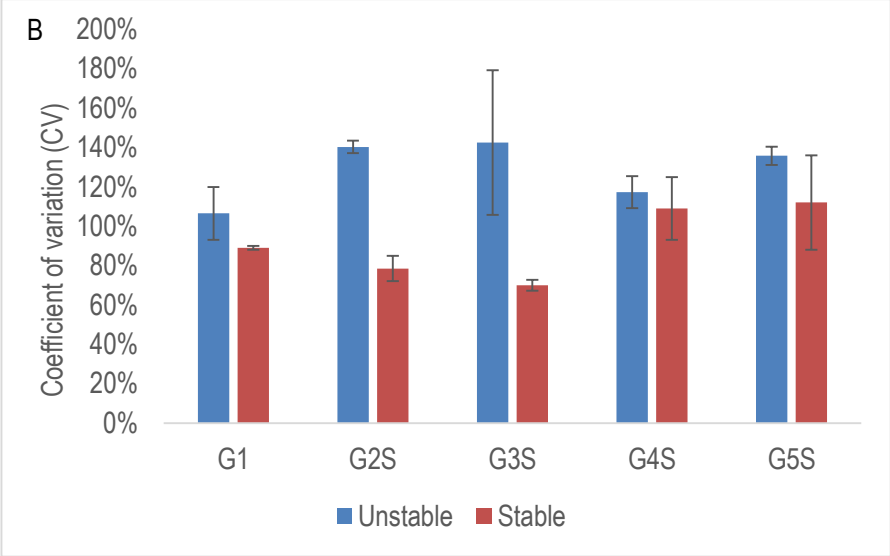
Figure 22: Growth heterogeneity passed over several generations



Growth heterogeneity is passed even after multiple selections. Fast- and slow- growing single cell-derived colonies were continuously selected over several generations. In unstable Sub1 cells, selection of a fast- or slow- growing colony resulted in the redistribution of growth heterogeneity, as daughter cell populations exhibited both fast- and slow- growing cells. In stable cells, selection of fast- and slow- growing single cell-derived colonies over several generations led to the selection of a fast- and slow- growing phenotype.

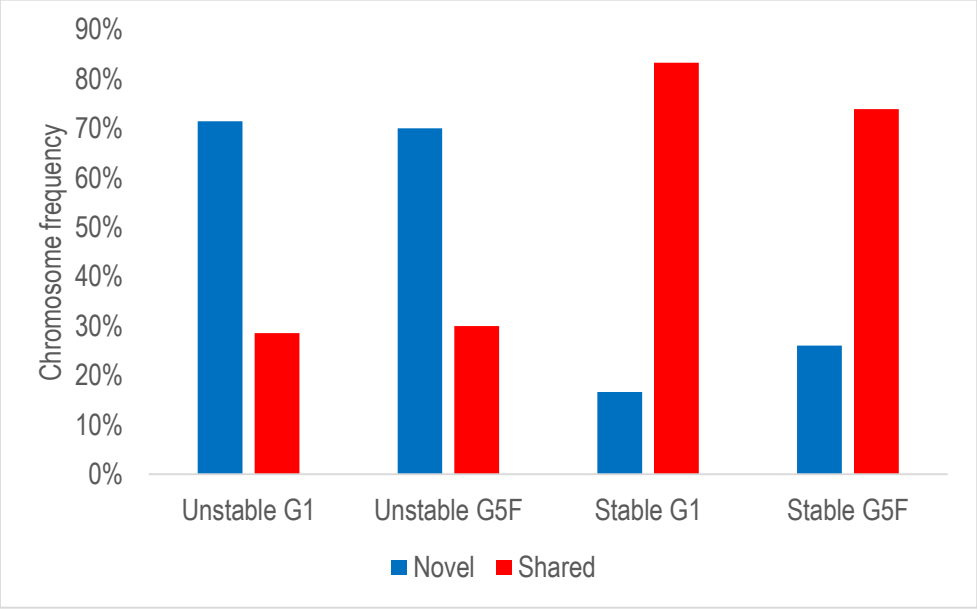
Figure 23: Growth heterogeneity of fast- and slow- growing colonies over multiple selections





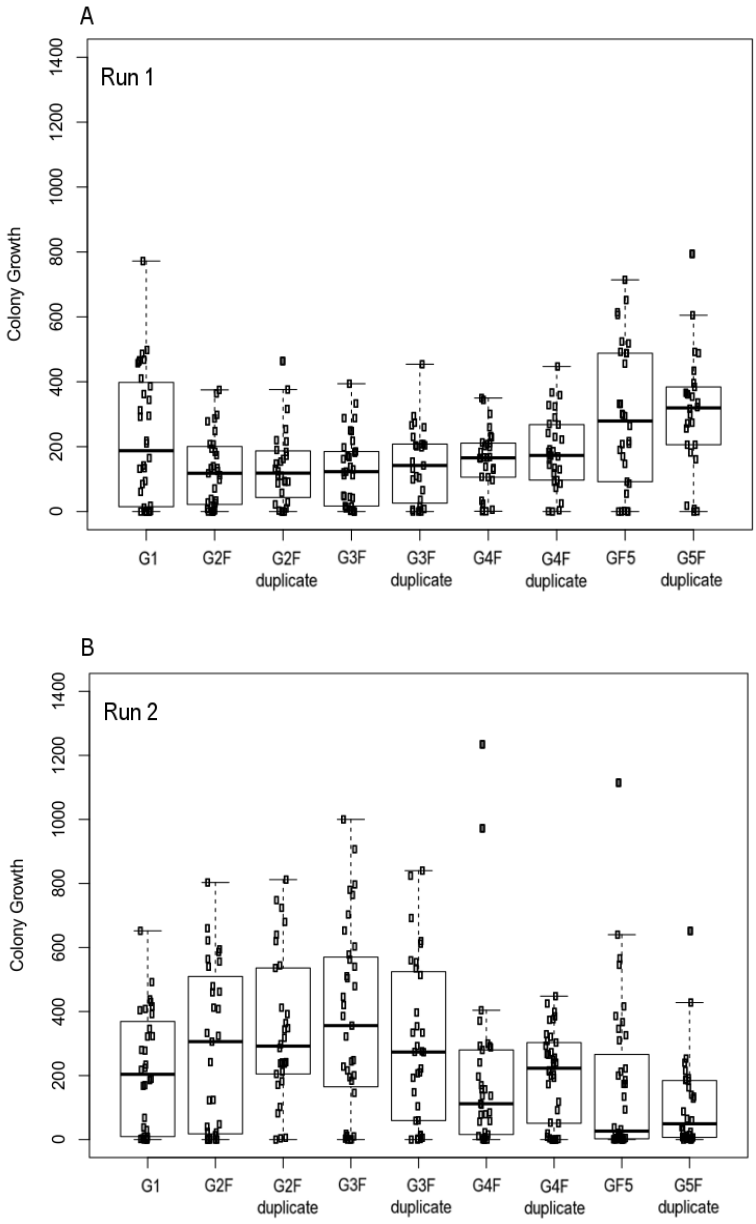
Growth heterogeneity, as measured by the coefficient of variation, for five selections of fast- (A) and slow-growing (B) colonies in unstable and stable cells. The CV largely remained the same across all selections for stable and unstable cells, suggesting that the same degree of growth was passed over several generations.

Figure 24: SKY analysis of generations 1 and 5 of unstable and stable cells



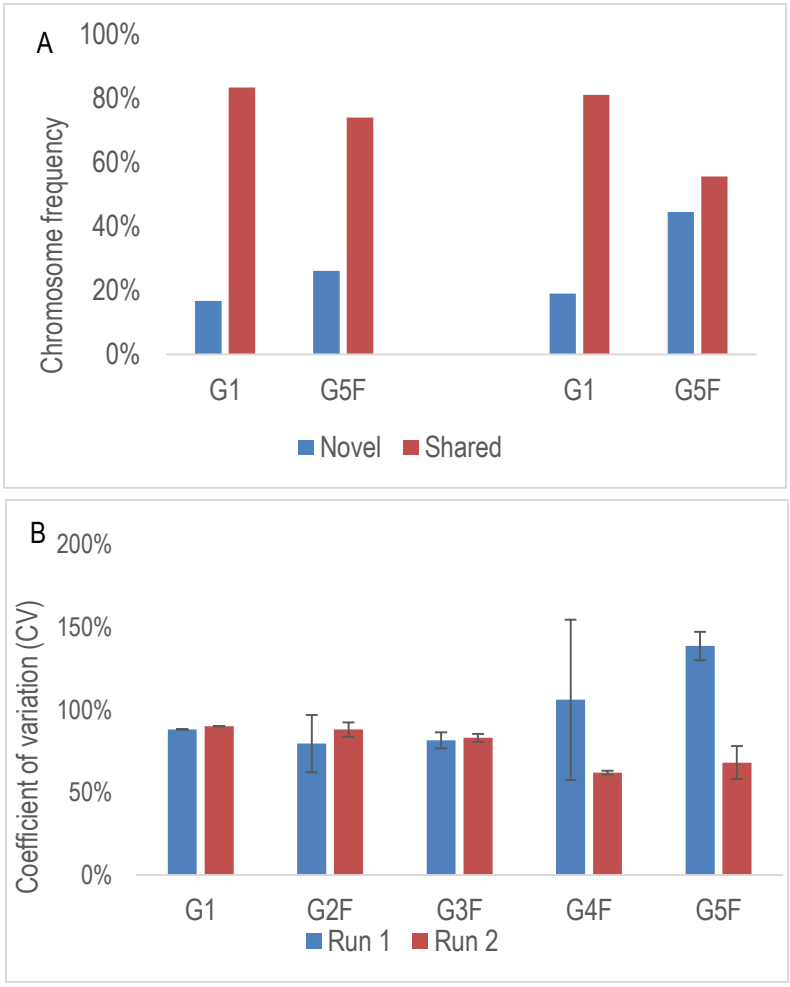
SKY analysis was completed on the parent generations (G1: generation 1) and the fifth selection of fast-growing colonies (G5F: generation 5, fast). In unstable cells, both G1 and G5F exhibited the same degree of novel and shared chromosome structural change. Novel changes outnumbered shared changes, and there was no significant difference in the degree of novel change even after multiple selections, indicating that the potential to generate genome heterogeneity is passed. The dynamics change in stable cells, as shared changes were dominant and novel changes were few. The same degree of shared and novel changes remained constant even after multiple selections, indicating that a minimal degree of genome heterogeneity was passed, even after multiple selections.

Figure 25: Biological duplicates of growth selection experiment in stable cells



Biological replicates in stable cells representing single cell growth *in situ* experiments after 5 selections of fast-growing colonies. A) In the first run, the growth rate increased after multiple selections of fast-growing colonies, indicating that a growth rate can be selected when the genome is unstable. B) However, the dynamics change in a duplicate experiment. While the growth rate increased after 3 selections, single cell growth became heterogeneous, resembling the single cell growth dynamics of unstable cells, after the fourth and fifth selections. Altogether, this suggests that a stable growth rate can be selected so long as the genome remains the stable.

Figure 26: SKY analysis and growth heterogeneity of biological replicates of stable cells



SKY and growth CV of stable cell replicates. A) SKY results describing the frequency of novel and shared chromosome change in biological replicates of stable cells. In the first run, there was no significant difference between novel and shared chromosome change after five selections. In the second run, the novel and shared chromosome change ratios changed, as the frequency of novel change increased after five selections. This indicates that genome instability had increased slightly, altering the genome status and the degree of genome heterogeneity that can be inherited. B) CV measuring growth heterogeneity for the selection of fast-growing colonies in each biological replicate in stable cells. In the first run, the CV largely remained the same for the first three generations, then decreased in the last two generations. The decreased CV was accompanied with an increase in growth rate, as more cells exhibited the same faster growing phenotype after selection. In the second run, the first three generations exhibited the similar CV values, however the CV increased at generations 4 and 5. The increased CV is accompanied with increased growth heterogeneity and increased novel chromosome change.

Taken together, this data suggests that a single cell can restore the degree of both genome and growth heterogeneity, and that these heterogeneities can be passed over multiple generations at a given rate, so long as the genome remains stable. In other words, a specific growth rate or karyotype, may not be passed, but the overall degree of heterogeneity can be passed when the genome is unstable. This data also demonstrates that the stability of a cell population is not a static feature, and as the degree of genome instability changes, the degree of heritable heterogeneity will also change, which can be reflected in the frequency of novel, stochastic chromosome change.

Discussion

The mechanism of fuzzy inheritance has significant implications in the biological sciences. Notably, it explains how cell populations maintain and regulate their heterogeneity, which is necessary for understanding and treating most cancers. At the surface these results may seem obvious, as it is commonly known that unstable cells will generate unstable cells. However, the data presented above demonstrated the mechanism of why an unstable cell population can pass heterogeneity without directly passing the same genome (system inheritance).

Can heterogeneity be inherited?

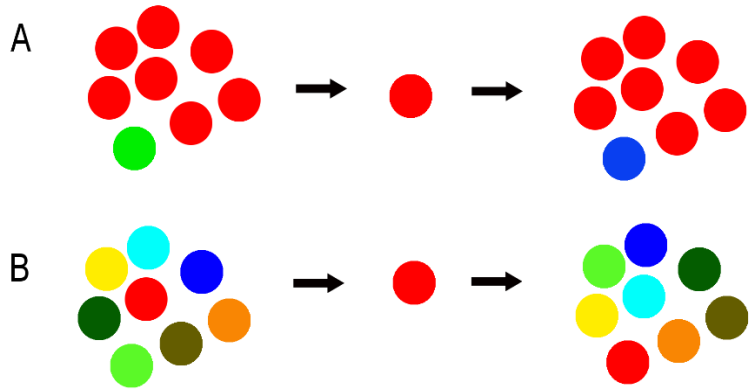
Can heterogeneity be considered a heritable feature? Inheritance can be defined as the actual passing of traits from the parent(s) to offspring. Here, a mother cell represents the parent and the daughter cell population is the offspring. The level of heterogeneity in a cell population represents an important biological trait. Therefore, the level of heterogeneity is a form of bio-inheritance. There are different types of inheritance that can be passed. The first is “parts inheritance” which is mediated by the DNA-level (Heng et al. 2011a, Heng et al. 2011b, Heng et al. 2013a). Under normal conditions, DNA is passed from mother cell to daughter cell, where the daughter cell receives a near identical copy of the mother cell’s DNA sequence. When the cells

are unstable, as in most cancers, a near identical sequence is not passed from mother cell to daughter cell due to genome and gene-level aberrations that continuously change. A second type of inheritance is genome-defined system inheritance (Heng 2009, Heng et al. 2011a). Under normal conditions, a mother cell will pass its genome-defined system inheritance (represented by a specific karyotype) to its daughter cell. In cancer, most of which are genomically unstable, a different genome is passed, and therefore system inheritance is lost. Taken together, the above types of inheritance embody a form of genetic transmission for genomically unstable cells. The genetic information from both the DNA parts level and the genome level, but with much less accuracy than normal cells with genomically stable cells. Along with the "parts" and altered system inheritance, what else is passed?

The pattern of inherited traits (karyotype, growth rate) was compared at the single cell level for cells that were "clonal" (by descent). However, for each of the traits analyzed, a specific growth rate or karyotype were not passed. Rather, the level of heterogeneity was passed. This is demonstrated in a model shown in Figure 27, which explains how fuzzy inheritance is linked to different types of heterogeneities in both normal stable and unstable cell populations. Each circle with a different color represents a cell with a unique karyotype. In the normal situation when cells are genomically stable, most cells have the same karyotype. There are, however, some cells with aberrant genomes that are outliers. Here, a single isolated cell can pass the same genome-defined system inheritance (karyotype) to its daughter cells. As in the parent cell population, cells with aberrant genomes will be present in the same frequency, but the specific aberrant genome will be different than the one found in the parent cell population. Thus, the heterogeneity of the parent cell population is restored. In unstable cell populations, an isolated single cell is unable to pass a specific karyotype. Rather, the unstable cell will pass down the ability to generate

heterogeneity, as all cells in the daughter cell population each have unique karyotypes that were not present in the parent cell population. The specific karyotypes will be different, but a similar degree of heterogeneity is inherited. Therefore, in both stable and unstable cell populations, the degree of heterogeneity is restored. The data demonstrates the restoration of heterogeneity, as cells did not pass down a specific growth rate or genome, but a range of altered genomes and growth rates. Therefore, the ability to generate heterogeneity was passed.

Figure 27: Model of fuzzy inheritance



Model of fuzzy inheritance. In the diagram, each different colored circle represents a cell with a unique genome. Each cell population exhibits a given degree of heterogeneity, and that same degree of heterogeneity can be passed to future generations. A) A smaller degree of heterogeneity is present in stable cell populations. A single cell isolated from a stable cell population will pass that same genome to its daughter cell population, along with a small degree of heterogeneity in the form of outlier cells. The outlier cells that exhibit stochastic genome level change are present in the quantity, but the specific stochastic genome level changes will not be passed. B) Unstable cell populations exhibit a larger degree of heterogeneity. A single cell isolated from an unstable cell population will not pass the same genome to its daughter cell population, but rather an array of heterogeneous genomes. The specific NCCAs will not be passed between generations, but the frequency of cells with stochastic genome level change will remain the same.

Another important feature of fuzzy inheritance is that the pattern of inheritance can be predicted from generation to generation, even though the specific karyotypes are not predictable. A smaller degree of heterogeneity is in stable cell populations and can be expected in its future daughter cell populations. This was observed in karyotypically stable HCT116 cells. Isolation of fast- and slow- growing subclones displayed the same small degree of genome and growth heterogeneity, or homogeneity, as the parent cell population. In large part, the same karyotype is expected to be passed down, along with a given number of cells with aberrant genomes. In a similar fashion, fast- and slow- growing subclones isolated from unstable Sub1 cells displayed the same high degree of genome and growth heterogeneity as the parent cell population. In unstable Sub1 cells, which specific karyotype that will become dominant cannot be predicted, however the degree of novel change can be predicted. The predictability of NCCA frequency has been observed in a mouse model of Atm -/- mice (Heng et al. 2006c). Spectral karyotyping was completed on cells kept in continuous cell culture isolated originating Atm -/- mice over several time-points. Genomic analysis was completed, and it was demonstrated that it was difficult to predict which specific genome would become dominant. The NCCA frequency, however, was found to be predictable.

Another important feature of fuzzy inheritance is that different cell populations have different degrees of heterogeneity, which is defined by the ratio of NCCAs to CCAs, or the karyotype profile. The frequency of NCCAs displayed by any given cell population or cell line is related to chromosome instability. Interestingly, a certain degree of genome heterogeneity has been found even among normal tissue. For example, mammalian liver cells exhibit a high degree of polyploidy that range from 2N-16N (Biesterfeld et al. 1994, Gupta 2000, Duncan et al. 2010, Duncan et al. 2012). Additionally, a small number of NCCAs have been found in human white

blood cells (Heng, unpublished data). Fuzzy inheritance is also found in normal cells that have identical karyotypes.

The data presented in this chapter demonstrates the mechanism of fuzzy inheritance at the karyotype level. Many published studies also support the existence of fuzzy inheritance at other genetic levels, such as the gene and epigenetic levels. For example, large scale single cell of breast cancer cells have discovered a large degree of de novo gene mutations (Wang et al. 2014). At the epigenetic level, heterogeneity of methylation patterns have been shown in multiple normal and disease tissue types (Landan et al. 2012, Schultz et al. 2015). Finally, a recent study conducted yeast documented that a low degree of karyotype heterogeneity is maintained, and that the degree of heterogeneity is higher than the baseline mitotic error rate (Zhu et al. 2015).

The significance of fuzzy inheritance

The biggest obstacle facing currently facing cancer research is heterogeneity. Fuzzy inheritance explains the cellular basis of heterogeneity; therefore, these findings are crucial to understanding most cancer types. Clonal expansion is the mechanism by which normal cells and cancers with stable cells can inherit the same system inheritance, or identical karyotype. However, in unstable cancers, clonal expansion does not occur as a single cell that is unstable cannot pass down an identical karyotype to its daughter cell. These variant cells pass down the ability to generate heterogeneity, leading to daughter cell populations with altered genomes but the same degree of genomic diversity at the population level. The new daughter cell population has a similar degree of genome level heterogeneity, which is inherited, but each cell has completely different genomes, which are not inherited. Therefore, in unstable cells, evolvability, or the potential generate change, is passed down while the genome-based system inheritance is not passed.

Fuzzy inheritance may offer a reason why theoretical models of clonal evolution do not accurately describe cancer evolution in the natural situation. Clonal evolution has been observed in more stable cancer types like chronic myelogenous leukemia (CML). However, the CML example can be considered as an exception (Horne et al. 2013). Most cancer types are genomically unstable and evolve in the macro-cellular punctuated phase where genome-level heterogeneity is high and the direct passage of genetic change is not traceable. Models of clonal evolution are based on the assumption that cell populations are genomically clonal, however this is not the case in the natural situation as clonal genomic aberrations cannot be detected due to the instability of most cancer types.

Fuzzy inheritance also addresses cancer therapy and why unstable cancers are very difficult to treat. Heterogeneity is the biggest obstacle for the effective treatment of cancer. The mechanism of fuzzy inheritance illustrated that the degree of heterogeneity is passed, giving future cell populations evolutionary potential by restoring the degree of genome heterogeneity. When a tumor is treated with any cancer treatment, most cells are killed and only a few survivors remain. These survivors have the ability to recapitulate the same degree of genomic heterogeneity that was present before initial treatment. Even though the specific karyotypes are different, the degree of heterogeneity and evolutionary potential is still the same and can be applied against new treatment (Heng 2016). Therefore, using the same initial treatment will be as useful in treating future cell populations. Fuzzy inheritance also explains why clonal cell populations are much easier to treat, as most cells display specific targets that can be eliminated and the amount of instability-mediated fuzzy inheritance is significantly lower. Cell populations that have a high degree of genome heterogeneity are often unstable and can survive through the mechanism of fuzzy inheritance. Furthermore, drug treatment can also increase the degree of genome-level heterogeneity (Liu et al.

2014). Fuzzy inheritance can explain why clonal expansion can be detected in cancers with more stable karyotypes. Additionally, fuzzy inheritance can explain why clonal expansion can be detected under experimental laboratory conditions, as highly homogeneous animal models are used and a single factor tested favors the emergence of clonal karyotypes. However in natural settings, multiple factors are involved in the formation of a solid tumor which are highly dynamic, and instability-mediated fuzzy inheritance becomes the dominant feature (Heng 2013b).

Fuzzy inheritance is most visible in cancer, however it is a feature that is universally shared in all cell populations. In contrast to conventional inheritance mechanisms that are precise, fuzzy inheritance is less precise that allows for a given degree of heterogeneity in a cell population. The heterogeneity that is passed has significant biological importance, as it gives a cell the potential to adapt to fluctuating environments. Cells populations that exhibit more precise inheritance, having more clonal karyotypes, are less likely to survive stress conditions. Many examples of fuzzy inheritance are found in the literature and are listed in Table 3. For example, under growth restricted conditions, bacteria cell populations will employ an error-prone DNA polymerase that will induce mutations (Ponder et al. 2005). The mutagenic mechanism that employs the error-prone polymerase lends the cell population adaptability necessary for survival by providing greater variability in the passing of its system inheritance. When the cell population returns to homeostatic conditions, the cell will switch back to a normal polymerase that allows a bacterial cell to divide with high fidelity. Interestingly, pathological conditions found in nature are often under stress conditions, while most laboratory settings are stable. This suggests that inheritance in natural conditions are fuzzier, and fuzzy inheritance represents the principle form of inheritance under very stressful conditions (for example most cancers) for somatic cell populations.

It is necessary to highlight an important distinction between the inheritance processes of germline cells and somatic cells. Inheritance in germline cells requires precision. This is because, for the survival and reproductive success of an individual, an identical karyotype must be passed from parents to offspring. However, for somatic cells, inheritance need not be precise. In fact, it may actually be advantageous for a cell population to exhibit genome-level heterogeneity, even in normal somatic tissue (Duncan et al. 2012, Horne et al. 2014). If every cell in a cell population had identical genomes, the cell population as a whole would not survive a stress event. But those cell populations that exhibit some degree of genome-level heterogeneity can actually survive, as they exhibit diversity and thus adaptability that are necessary for survival (Horne et al. 2014). Fuzzy inheritance facilitates this process, as it allows for a cell to pass a given degree of heterogeneity or instability to future cell populations. Altogether, this indicates that there is a type of inheritance for germline cells that requires precision to maintain or pass the same karyotype over multiple generations, and a different type of inheritance, fuzzy inheritance, that is needed and dominant during an individual's lifespan.

Fuzzy inheritance is also demonstrated at the DNA sequence level (Lodato et al. 2015). Single cell sequencing was performed on 36 neurons isolated from the cerebral cortex of three healthy individuals. Single cell single nucleotide variants (SNVs) were identified and occurred at a very high rate. On average, each neuron exhibited 1,685-1793 SNVs, where the number of SNVs were tightly clustered among single neurons of the same individual. This supports fuzzy inheritance, as the degree of SNVs is the same across each cell for each individual.

Table 3: Examples of Fuzzy inheritance in the literature

Information level	Citation
Genome level	(Heng et al. 2006c)
Genome level	(Heng 2007c)
Genome level	(Lawrenson 2010)
Genome level	(Liu et al. 2014)
Genome level	(Zhu et al. 2015)
Genome level	(Duesberg and McCormack 2013)
Genome level	(Bakker et al. 2016)
Transcriptome level	(Stevens et al. 2014)
Transcriptome level	(Lawrenson 2010)
Transcriptome level	(Creekmore et al. 2011)
Transcriptome level	(Lee et al. 2014)
Transcriptome level	(Durfee et al. 2010)
Transcriptome level	(Chang et al. 2008)
Gene expression	(Gupta et al. 2011)
DNA sequence level	(Lodato et al. 2015)
DNA sequence level	(Cannella et al. 2009)
DNA sequence level	(Ponder et al. 2005)
DNA sequence level	(Long et al. 2016)
Epigenetic level	(Cerulus et al. 2016)
Epigenetic level	(Farlik et al. 2015)
Cell growth	(Keren et al. 2015)
Cell growth	(Sandler et al. 2015)

\

Published studies that support or are evidence of fuzzy inheritance.

Another example of fuzzy inheritance illustrates the inheritance of growth heterogeneity (Sandler et al. 2015). L1210 cells, a genomically unstable murine leukemia cell line that forms tumors in mice (Teicher 2006) were transfected using Fucci vectors, and the growth rates of single cells were monitored over several generations. Fucci vectors allow for the monitoring of cell cycle duration of living cells in real time, as G1 is distinguished from S/G2 by color. Red fluorescent protein (RFP) is tagged onto cdt1, which is active in G1, therefore cells in G1 are red. Geminin is tagged with green fluorescent protein (GFP), and as Cdt1 is degraded and geminin is expressed in S and G2, the cell turns green. Single cell times were analyzed over several generations, and it was found that cell cycle times between mother cells and daughter cells were not correlated, as daughter cells exhibited a variety of growth rates that were not similar to the mother cell growth rate. This supports fuzzy inheritance, as a mother cell could not pass down a specific growth rate, but did pass a range of growth rates.

While we only demonstrate fuzzy inheritance at the genome level, fuzzy inheritance can be observed and should be quantitatively demonstrated at the gene and epigenetic levels. Further research is also needed to understand the relationship between different types of inheritances (at different genetic levels) and whether system inheritance can be stably passed under stressful conditions. Fuzzy inheritance can lead to new and exciting developments in biological research, and specifically cancer research.

CHAPTER 5: CONCLUSION AND DISCUSSION

Summary

In this dissertation, the novel mechanism of fuzzy inheritance was presented. According to fuzzy inheritance, a cell population exhibits a given degree of heterogeneity, and that same degree of heterogeneity is passed to future cell generations. The following observations support this hypothesis. First, it was demonstrated that single cells isolated from genomically unstable cell populations could not pass a specific karyotype (system inheritance). Instead, single unstable cells passed instability-mediated heterogeneity, and generated daughter cell populations with an array of heterogeneous karyotypes. This was demonstrated in chapters 3 and 4 in single-cell derived subpopulations that were isolated from unstable *ex-vivo* wild type and conditionally inactivated Brca1/p53 MOSE cell populations. SKY analysis of each single-cell derived subpopulation that originated from a genomically unstable parent cell population exhibited distinctly different karyotypes consistent with punctuated evolution, where the degree of novel change was the same for each subpopulation. In contrast, cells isolated from stable cell populations could be cloned, as the same karyotype was passed to future cell populations. This was demonstrated in chapter 4, where single cells isolated from karyotypically stable HCT116 cells all exhibited similar karyotypes and minimal genome-level change.

Second, a single unstable cell could not pass a specific growth rate. In chapters 3 and 4, single cell *in-situ* growth experiments were completed in single-cell derived cell populations with varying degrees of genome instability. Among unstable cells, single cells exhibited a large degree of growth heterogeneity, where most cells grew at a slow or moderate rate and only few cells exhibited very high proliferation. To detect if heterogeneity was present at the earliest cell divisions, daily growth rates were calculated and demonstrated that heterogeneity was present at

the earliest cell divisions and remained constant for the duration of the assay. Altogether this indicates that single cells did not pass down a single growth rate, but rather passed growth heterogeneity. In contrast, stable cells were able to pass a specific growth rate, as cells all exhibited relatively the same growth rate.

Third, it was demonstrated that unstable cells did not inherit a specific growth rate or genome, even after repeated selection. Single cell *in-situ* growth experiments (Chapter 4), where fast- and slow- growing colonies were isolated over several generations to determine if a specific genome or growth rate can be inherited. In unstable cells, repeated selection of fast- or slow-growing colonies always produced cell populations that exhibited a heterogeneous array of growth rates, not a specific growth rate. Spectral karyotyping of select generations found that no specific genome was passed after repeated selection, as cells with heterogeneous genomes were present. In contrast, repeated selection in stable cells found that a growth rate and specific genome could be selected. This indicates that, so long as the genome is unstable, specific features like growth rate or karyotype cannot be inherited, and heterogeneity is passed.

Fourth, it was demonstrated that genome heterogeneity is linked to other heritable features of the system, like growth heterogeneity. In chapters 3 and 4, single cell *in-situ* growth experiments were completed in different cell lines with different degrees of genome instability. Results demonstrated that growth heterogeneity increased in cell populations with higher levels of genome instability, providing a direct link between genome heterogeneity and growth heterogeneity.

Fifth, outliers dominate cell population dynamics in unstable cell populations. Single cell *in-situ* growth assays in unstable cells completed in chapters 3 and 4 demonstrated that most cells grew at a slow to moderate rate while only few outlier cells were highly proliferative. These highly proliferative outlier cells were largely responsible for repopulating the majority of the next

generation of cells. In contrast, single cell growth assays of genomically stable cells demonstrated that each single cell made approximately the same contribution in repopulating the next generation of cells. This indicates that outlier cells dominate cell population dynamics in cell populations that are genomically unstable, as they are the cells that generate future cell populations.

Finally, it was demonstrated that the statistical average is not a suitable measure for cell populations that are unstable, however it can reliably be used for genomically stable cell populations. Single cell SKY and growth assays in chapter 3 demonstrated that an average, composite karyotype or average cell growth is not representative of the entire cell population. However, in stable cells, the average could be used to accurately assess the cell population. This indicates that in cell populations that are unstable, single cell analysis is needed to accurately characterize a cell population.

Altogether, the above supports the novel cellular mechanism of fuzzy inheritance. Figure 23 (in Chapter 4) shows a mechanism for how both stable and unstable cell populations can pass a given degree of heterogeneity to daughter cell populations. A small degree of heterogeneity exists in stable cell populations. A single cell isolated from the stable cell population will pass that same karyotype, producing a daughter cell population that is mostly genomically clonal, but also exhibits the same degree of heterogeneity. The specific features of the outlier cell may be different, but the frequency will remain the same. In other words, the specific NCCAs present will not be passed, they will be replaced by new NCCAs at every generation. In contrast, a single cell isolated from an unstable cell population will produce a daughter cell population of unstable cells with the same degree of genome-level heterogeneity. Again, the specific NCCAs will not be passed and will change in the next generation, but their frequency will remain the same so long as the genome status is unchanged.

Future directions

The work completed in this dissertation characterizes the mechanism of fuzzy inheritance at the genome level. However, future work is needed in order to directly detect fuzzy inheritance at other genetic and non-genetic levels, such as the DNA sequence level and epigenetic levels, to determine how different types of heterogeneities are linked. For example, how does changing the rate of heterogeneity at one genetic level affect the rate of heterogeneity at other levels? Genome heterogeneity has been previously linked with transcriptome heterogeneity and growth heterogeneity. Therefore, it is likely that a similar relationship between genome heterogeneity and DNA sequence heterogeneity, or genome heterogeneity and epigenetic heterogeneity also exist.

In order to detect fuzzy inheritance at the DNA sequence and epigenetic levels and understand how it relates to genome-level fuzzy inheritance, a single-cell based approach that analyzes the frequency of *de novo* change at each molecular level (genome, DNA sequence, epigenetic) is proposed. An *in-vitro* cell culture system is the most ideal experimental system, as it allows for the monitoring of evolution in action in real time. Cells can be isolated at any time-point for analysis including different phases of cellular evolution. For example, a paired cell culture system with known long-term stability and instability is crucial for detecting fuzzy inheritance at different cellular passages. Paired cell culture systems that fit these criteria are: MDAH-041 cells at passage 25 (which is an unstable passage), and passage 54 (which has known long-term stability). Wild type HCT116 cells are also known for the long-term karyotypic stability, and can be paired with HCT116 with a p53 knockout. The karyotypic stability of the cell line must be validated over multiple time-points to ensure long term stability prior to any experiments. Finally, the detection of fuzzy inheritance must be completed at the single cell level, as average-

based approaches would smooth out the variation that is needed to identify marks of fuzzy inheritance.

In order to detect fuzzy inheritance in single cells over multiple generations, single cells will be isolated from stable and unstable cell populations and cultivated, similar to the experiments completed in this dissertation. Single cell growth experiments will be performed and specific colonies of varying sizes will be repeatedly selected over multiple generations. At each generation, the growth (single cell *in situ* growth will be monitored), genome (spectral karyotyping), DNA sequencing (whole genome sequencing) and epigenetic profiling (whole genome bisulfite sequencing) will be completed. The following parameters will be measured to detect fuzzy inheritance: growth heterogeneity (single cell growth); NCCA frequency (genome); *de novo* mutations (whole genome sequencing); and differentially methylated regions (epigenetic profiling).

It is anticipated that the integration of all data types will demonstrate that different types of heterogeneities are linked in a dynamic relationship. By comparing the rate of stochastic genome-level change to the rate of stochastic gene mutation or epigenetic change, it will be revealed that fuzzy inheritance operates at other molecular levels by passing a given degree of change. Specifically, in unstable cell populations where the rate of genome level stochastic change is high or increases, the rate of *de novo* gene mutations or the number of differentially methylated regions will also increase. In a similar fashion, as the frequency of genome level stochastic change decreases, the rate of *de novo* gene mutations and differentially methylated regions will also decrease.

At the same time, the integration of this data will also put into context the importance and impacts of different types of heterogeneities. Specifically, genome level change impacts the global

cellular and molecular dynamics of a cell, whereas a gene mutation or epigenetic change impacts molecular dynamics on a more local scale. In other words, a single balanced or unbalanced structural change can alter the activities of hundreds to thousands of genes. The emergent properties and cellular networks have changed. In comparison, a single gene or epigenetic change may only alter the effect of a handful genes, keeping the global cellular networks intact. In a population full of single, unstable cells. where each cell exhibits balanced and/or unbalanced change, the complexity of the cellular dynamics is magnified. Altogether, this data would confirm that genome level change is the significant type of change, and the monitoring of stochastic genome level change is the first step towards understanding the maintenance of cellular heterogeneity.

A number of studies support the above anticipated findings. First, genome heterogeneity has been previously linked to transcriptome heterogeneity (Stevens et al. 2014). Second, a number of genome sequencing studies have completed concurrent copy number profiling and gene expression found that some different heterogeneities may be linked (Bashashati et al. 2013, Gerlinger et al. 2014, Bakker et al. 2016). It was found that patients with increased heterogeneity of copy number alterations between regions also exhibited increased heterogeneity of transcriptome profiles or mutation heterogeneity, suggesting that the different types of heterogeneities are linked. Third, our unpublished data has suggested that a high degree of stress can elevate both genome alterations and DNA dynamics, based on the status of SNPs.

Surely, this new information will not only shape the knowledge basis of future genetics and genomics research, but also identify the most suitable level for studying the genetic basis for specific types of human diseases.

Precision medicine

Precision medicine is the latest initiative intended to find new ways to effectively fight and treat cancer and other common/complex diseases (Collins and Varmus 2015). Precision medicine refers to the integration and application of genomic technologies into improved and more precise clinical diagnostics and treatment recommendations, as well as the incorporation of more genomic techniques in drug development. It is similar to the previous initiative, personalized medicine, in that it seeks to translate precise molecular information for clinical use, but also different as it emphasizes a systems approach. As 'omic technologies become less expensive, the integration of patient genetic and nongenetic profiles for improved clinical care has become the new reality. Precision medicine has seen some success stories (Katsnelson 2013) and several clinical trials are currently underway.

Despite the renewed hopes that precision medicine has generated, its successes may be limited (Heng 2016). First, this approach does not consider heterogeneity. Cancer is an evolutionary process where multi-level heterogeneity is a characteristic feature. As thoroughly discussed throughout this dissertation, the average profile does not accurately characterize a tumor cell population. The genetic/nongenetic profiles generated by the various 'omic technologies represent an average population of a snapshot of the evolutionary process. While this does provide useful information about some cell population dynamics, it is still average-based, and reflects only more dominant subpopulations, thereby excluding some outlier cells that may emerge as dominant players in cancer evolution.

Furthermore, following treatment, a phenomenon known as genome chaos occurs (Liu et al. 2014). Genome chaos is an adaptive macro-cellular evolutionary strategy that describes the massive genome rearrangement that occurs after a stress event, including chemotherapeutic treatment. After a major stressor is applied, chromosomes are fragmented, followed by a rejoining

process where chromosome fragments are randomly pieced together. The process of fragmentation and rejoining continues until a cell with a more stable genome is formed. Fuzzy inheritance mediates the process of genome chaos. During the rejoining process, chaotic genomes are repeatedly passed for a period of several weeks before a stable clone is selected, thereby providing the necessary variation for survival and evolution. Genome chaos is not about specific genetic features, but rather providing a variety of new genome systems for selection and evolution. This is consistent with fuzzy inheritance, as unstable cells pass an array of genomically heterogeneous cells, thereby increasing the evolutionary potential of a cell population.

The clinical implications of genome chaos indicate that initiatives like precision medicine may see limited results. It is a paradox that most cancer researchers do not recognize: nearly all cancer researchers agree that cancer is an evolutionary process, but it is a stochastic process as well. A highly specific and precise genetic approach will not drastically improve unstable cancer diagnosis and treatment because as cancer evolves, the chromosome topology will change, and these changes cannot be predicted. Genetic/nongenetic variants that emerge as statistically significant or clinically actionable remain significant under that specific genome context (specific genome), and with new genome changes, new significant variants will emerge. In other words, the idea that cancer diagnosis and treatment can be predicted by several genetic/nongenetic markers may be less significant when these genes are put into the context of massively rearranged genomes. What can we precisely target if the targets themselves are always changing? A re-evaluation of cancer therapy that recognizes the ultimate importance of the genome and genome-level rearrangements is necessary for new discoveries in cancer treatment and diagnosis (Horne S.D. 2015b).

Closing remarks

Detecting fuzzy inheritance was a difficult and arduous task. Believing in fuzzy inheritance was another feat altogether. Regardless of what I believed to be true, I also had to trust and believe in my data. We were taught at a young age that inheritance is precise, that knowledge of our genes will unlock the secrets of our evolutionary history as well as our potential for the future, that our genomes are these rigid structures made to withstand massive amounts of stress. However, after examining the process of inheritance, especially in cancer cells, that inheritance is perplexing is an understatement. At the surface, it would seem that our experimental platforms and research results are working and are building upon our knowledge. However, synthesizing conceptual inheritance with the massive data that has been generated indicates that perhaps, inheritance is understood less than we believe. Genetic inheritance is not the transmission of the beautiful and picturesque chromosomes so perfectly positioned in a nucleus, particularly in cancer. In unstable cells, it is messy and chaotic. Our DNA sequence may not carry the untold secrets of life, as currently, heritability is still missing. And our genome is not an unbreakable structure that has withstood thousands of years of evolution. It changes more frequently than expected, maybe more so. While some may find this disappointing, it is actually quite exciting, as it paves the way for new and exciting research. It renews my passion for scientific discovery, and I am excited for what the future will bring.

APPENDIX A

PUBLICATIONS INCLUDED IN THIS DISSERTATION

Abdallah BY, Horne SD, Liu G, Stevens JB, Ye CJ, Bremer SW, Gorelick,R, Krawetz SA, Heng HH. "Fuzzy inheritance: a novel form of inheritance that regulates tumor heterogeneity." (submitted)

Abdallah BY, Horne SD, Kurkinen M, Stevens JB, Liu G, Ye KJ, Barbat J, Bremer SW Heng HH. (2014) "Ovarian cancer evolution through stochastic genome alterations: defining the genomic role in ovarian cancer." Systems Biology in Reproductive Medicine. 60(1): 2-13.

Abdallah BY, Horne SD, Stevens JB, Liu G, Ying AY, Vanderyhyden B, Krawetz SA, Heng HH. (2013) "Single cell heterogeneity: why unstable genomes are incompatible with average profiles." Cell Cycle. 12(23): 3640-3649.

REFERENCES

Abdallah, B. Y., S. D. Horne, M. Kurkinen, J. B. Stevens, G. Liu, C. J. Ye, J. Barbat, S. W. Bremer and H. H. Heng (2014). "Ovarian cancer evolution through stochastic genome alterations: defining the genomic role in ovarian cancer." Syst Biol Reprod Med **60**(1): 2-13.

Abdallah, B. Y., S. D. Horne, J. B. Stevens, G. Liu, A. Y. Ying, B. Vanderhyden, S. A. Krawetz, R. Gorelick and H. H. Heng (2013). "Single cell heterogeneity: why unstable genomes are incompatible with average profiles." Cell Cycle **12**(23): 3640-3649.

Acar, M., J. T. Mettetal and A. van Oudenaarden (2008). "Stochastic switching as a survival strategy in fluctuating environments." Nat Genet **40**(4): 471-475.

Almendro, V., A. Marusyk and K. Polyak (2013). "Cellular heterogeneity and molecular evolution in cancer." Annu Rev Pathol **8**: 277-302.

AmericanCancerSociety (2016). Cancer Facts & Figures 2016. A. C. Society. Atlanta.

Baca, S. C., D. Prandi, M. S. Lawrence, J. M. Mosquera, A. Romanel, Y. Drier, K. Park, N. Kitabayashi, T. Y. Macdonald, M. Ghandi, E. Van Allen, G. V. Kryukov, A. Sboner, J. P. Theurillat, T. D. Soong, E. Nickerson, D. Auclair, A. Tewari, H. Beltran, R. C. Onofrio, G. Boysen, C. Guiducci, C. E. Barbieri, K. Cibulskis, A. Sivachenko, S. L. Carter, G. Saksena, D. Voet, A. H. Ramos, W. Winckler, M. Cipicchio, K. Ardlie, P. W. Kantoff, M. F. Berger, S. B. Gabriel, T. R. Golub, M. Meyerson, E. S. Lander, O. Elemento, G. Getz, F. Demichelis, M. A. Rubin and L. A. Garraway (2013). "Punctuated evolution of prostate cancer genomes." Cell **153**(3): 666-677.

Bai, H., D. Cao, J. Yang, M. Li, Z. Zhang and K. Shen (2016). "Genetic and epigenetic heterogeneity of epithelial ovarian cancer and the clinical implications for molecular targeted therapy." (1582-4934 (Electronic)).

Bakker, B., A. Taudt, M. E. Belderbos, D. Porubsky, D. C. Spierings, T. V. de Jong, N. Halsema, H. G. Kazemier, K. Hoekstra-Wakker, A. Bradley, E. S. de Bont, A. van den Berg, V. Guryev, P. M. Lansdorp, M. Colome-Tatche and F. Foijer (2016). "Single-cell sequencing reveals karyotype heterogeneity in murine and human malignancies." Genome Biol **17**(1): 115.

Balaban, N. Q., J. Merrin, R. Chait, L. Kowalik and S. Leibler (2004). "Bacterial persistence as a phenotypic switch." Science **305**(5690): 1622-1625.

Baranzini, S. E., J. C. Mudge J Fau - van Velkinburgh, P. van Velkinburgh Jc Fau - Khankhanian, I. Khankhanian P Fau - Khrebtukova, N. A. Khrebtukova I Fau - Miller, L. Miller Na Fau - Zhang, A. D. Zhang L Fau - Farmer, C. J. Farmer Ad Fau - Bell, R. W. Bell Cj Fau - Kim, G. D. Kim Rw Fau - May, J. E. May Gd Fau - Woodward, S. J. Woodward Je Fau - Caillier, J. P. Caillier Sj Fau - McElroy, R. McElroy Jp Fau - Gomez, M. J. Gomez R Fau - Pando, L. E. Pando Mj Fau - Clendenen, E. E. Clendenen Le Fau - Ganusova, F. D. Ganusova Ee Fau - Schilkey, T. Schilkey Fd Fau - Ramaraj, O. A. Ramaraj T Fau - Khan, J. J. Khan Oa Fau - Huntley, S. Huntley Jj Fau - Luo, P.-Y. Luo S Fau - Kwok, T. D. Kwok Py Fau - Wu, G. P. Wu Td Fau - Schroth, J. R. Schroth Gp Fau - Oksenberg, S. L. Oksenberg Jr Fau - Hauser, S. F. Hauser Sl Fau - Kingsmore and S. F. Kingsmore (2010). "Genome, epigenome and RNA sequences of monozygotic twins discordant for multiple sclerosis." (1476-4687 (Electronic)).

Bashashati, A., G. Ha, A. Tone, J. Ding, L. M. Prentice, A. Roth, J. Rosner, K. Shumansky, S. Kalloger, J. Senz, W. Yang, M. McConechy, N. Melnyk, M. Anglesio, M. T. Luk, K. Tse, T. Zeng, R. Moore, Y. Zhao, M. A. Marra, B. Gilks, S. Yip, D. G. Huntsman, J. N. McAlpine and S. P. Shah (2013). "Distinct evolutionary trajectories of primary high-grade serous ovarian cancers revealed through spatial mutational profiling." J Pathol **231**(1): 21-34.

Bast, R. C., Jr., B. Hennessy and G. B. Mills (2009). "The biology of ovarian cancer: new opportunities for translation." Nat Rev Cancer **9**(6): 415-428.

Bast, R. C., Jr., F. J. Xu, Y. H. Yu, S. Barnhill, Z. Zhang and G. B. Mills (1998). "CA 125: the past and the future." Int J Biol Markers **13**(4): 179-187.

Bayani, J., J. D. Brenton, P. F. Macgregor, B. Beheshti, M. Albert, D. Nallainathan, J. Karaskova, B. Rosen, J. Murphy, S. Laframboise, B. Zanke and J. A. Squire (2002). "Parallel analysis of sporadic primary ovarian carcinomas by spectral karyotyping, comparative genomic hybridization, and expression microarrays." Cancer Res **62**(12): 3466-3476.

Bayani, J., J. Paderova, J. Murphy, B. Rosen, M. Zielenska and J. A. Squire (2008). "Distinct patterns of structural and numerical chromosomal instability characterize sporadic ovarian cancer." Neoplasia **10**(10): 1057-1065.

Bell Jt Fau - Spector, T. D. and T. D. Spector (2014). "DNA methylation studies using twins: what are they telling us?" (1474-760X (Electronic)).

Bell, J. T. and T. D. Spector (2011). "A twin approach to unraveling epigenetics." (0168-9525 (Print)).

Biesterfeld, S., K. Gerres, G. Fischer-Wein and A. Bocking (1994). "Polyploidy in non-neoplastic tissues." J Clin Pathol **47**(1): 38-42.

Birrer, M. J. (2010). "The origin of ovarian cancer-is it getting clearer?" N Engl J Med **363**(16): 1574-1575.

Blagosklonny, M. V. (2006). "Target for cancer therapy: proliferating cells or stem cells." Leukemia **20**(3): 385-391.

Bonome, T., D. A. Levine, J. Shih, M. Randonovich, C. A. Pise-Masison, F. Bogomolniy, L. Ozbun, J. Brady, J. C. Barrett, J. Boyd and M. J. Birrer (2008). "A gene signature predicting for survival in suboptimally debulked patients with ovarian cancer." Cancer Res **68**(13): 5478-5486.

Brock, A., H. Chang and S. Huang (2009). "Non-genetic heterogeneity--a mutation-independent driving force for the somatic evolution of tumours." Nat Rev Genet **10**(5): 336-342.

Byrne, E. M., A. K. Carrillo-Roa T Fau - Henders, L. Henders Ak Fau - Bowdler, A. F. Bowdler L Fau - McRae, A. C. McRae Af Fau - Heath, N. G. Heath Ac Fau - Martin, G. W. Martin Ng Fau - Montgomery, L. Montgomery Gw Fau - Krause, N. R. Krause L Fau - Wray and N. R. Wray (2013). "Monozygotic twins affected with major depressive disorder have greater variance in methylation than their unaffected co-twin." (2158-3188 (Electronic)).

Campbell, I. M., C. A. Shaw, P. Stankiewicz and J. R. Lupski (2015). "Somatic mosaicism: implications for disease and transmission genetics." (0168-9525 (Print)).

Cannella, M., T. Maglione V Fau - Martino, G. Martino T Fau - Ragona, L. Ragona G Fau - Frati, G.-M. Frati L Fau - Li, F. Li Gm Fau - Squitieri and F. Squitieri (2009). "DNA instability in replicating Huntington's disease lymphoblasts." (1471-2350 (Electronic)).

Capp, J. P. (2005). "Stochastic gene expression, disruption of tissue averaging effects and cancer as a disease of development." Bioessays **27**(12): 1277-1285.

Cerulus, B., A. M. New, K. Pougach and K. J. Verstrepen (2016). "Noise and Epigenetic Inheritance of Single-Cell Division Times Influence Population Fitness." (1879-0445 (Electronic)).

Chakraborty, A. K. and J. P. Roose (2013). "Biochemical heterogeneity and developmental varieties in T-cell leukemia." Cell Cycle **12**(10): 1480-1481.

Chan, J., K. Fuh, J. Shin, M. Cheung, C. Powell, L. M. Chen, D. Kapp and K. Osann (2008). "The treatment and outcomes of early-stage epithelial ovarian cancer: have we made any progress?" Br J Cancer **98**(7): 1191-1196.

Chang, H. H., M. Hemberg, M. Barahona, D. E. Ingber and S. Huang (2008). "Transcriptome-wide noise controls lineage choice in mammalian progenitor cells." Nature **453**(7194): 544-547.

Cheng, K. W., W.-L. Lahad Jp Fau - Kuo, A. Kuo Wl Fau - Lapuk, K. Lapuk A Fau - Yamada, N. Yamada K Fau - Auersperg, J. Auersperg N Fau - Liu, K. Liu J Fau - Smith-McCune, K. H. Smith-McCune K Fau - Lu, D. Lu Kh Fau - Fishman, J. W. Fishman D Fau - Gray, G. B. Gray Jw Fau - Mills and G. B. Mills (2004). "The RAB25 small GTPase determines aggressiveness of ovarian and breast cancers." (1078-8956 (Print)).

Cho, H. Y., S. H. Park, Y. H. Park, H. B. Kim, J. B. Kang, S. H. Hong and M. S. Kyung (2015). "Comparison of HE4, CA125, and Risk of Ovarian Malignancy Algorithm in the Prediction of Ovarian Cancer in Korean Women." J Korean Med Sci **30**(12): 1777-1783.

Clark-Knowles, K. V., K. Garson, J. Jonkers and B. C. Vanderhyden (2007). "Conditional inactivation of Brca1 in the mouse ovarian surface epithelium results in an increase in preneoplastic changes." Exp Cell Res **313**(1): 133-145.

Clark-Knowles, K. V., M. K. Senterman, O. Collins and B. C. Vanderhyden (2009). "Conditional inactivation of Brca1, p53 and Rb in mouse ovaries results in the development of leiomyosarcomas." PLoS One **4**(12): e8534.

Cohen, A. A., N. Geva-Zatorsky, E. Eden, M. Frenkel-Morgenstern, I. Issaeva, A. Sigal, R. Milo, C. Cohen-Saidon, Y. Liron, Z. Kam, L. Cohen, T. Danon, N. Perzov and U. Alon (2008). "Dynamic proteomics of individual cancer cells in response to a drug." Science **322**(5907): 1511-1516.

Collins, F. S. and A. D. Barker (2007). "Mapping the cancer genome. Pinpointing the genes involved in cancer will help chart a new course across the complex landscape of human malignancies." Sci Am **296**(3): 50-57.

Collins, F. S. and H. Varmus (2015). "A new initiative on precision medicine." (1533-4406 (Electronic)).

Cooke, S. L., C. K. Ng, N. Melnyk, M. J. Garcia, T. Hardcastle, J. Temple, S. Langdon, D. Huntsman and J. D. Brenton (2010). "Genomic analysis of genetic heterogeneity and evolution in high-grade serous ovarian carcinoma." Oncogene **29**(35): 4905-4913.

Creekmore, A. L., W. T. Silkworth, D. Cimini, R. V. Jensen, P. C. Roberts and E. M. Schmelz (2011). "Changes in gene expression and cellular architecture in an ovarian cancer progression model." PLoS One **6**(3): e17676.

Cuellar-Partida, G., Y. Lu, S. C. Dixon, P. A. Fasching, A. Hein, S. Burghaus, M. W. Beckmann, D. Lambrechts, E. Van Nieuwenhuysen, I. Vergote, A. Vanderstichele, J. A. Doherty, M. A. Rossing, J. Chang-Claude, A. Rudolph, S. Wang-Gohrke, M. T. Goodman, N. Bogdanova, T. Dork, M. Durst, P. Hillemanns, I. B. Runnebaum, N. Antonenkova, R. Butzow, A. Leminen, H. Nevanlinna, L. M. Pelttari, R. P. Edwards, J. L. Kelley, F. Modugno, K. B. Moysich, R. B. Ness, R. Cannioto, E. Hogdall, C. Hogdall, A. Jensen, G. G. Giles, F. Bruinsma, S. K. Kjaer, M. A. Hildebrandt, D. Liang, K. H. Lu, X. Wu, M. Bisogna, F. Dao, D. A. Levine, D. W. Cramer, K. L. Terry, S. S. Tworoger, M. Stampfer, S. Missmer, L. Bjorge, H. B. Salvesen, R. K. Kopperud, K. Bischof, K. K. Aben, L. A. Kiemeney, L. F. Massuger, A. Brooks-Wilson, S. H. Olson, V. McGuire, J. H. Rothstein, W. Sieh, A. S. Whittemore, L. S. Cook, N. D. Le, C. Blake Gilks, J. Gronwald, A. Jakubowska, J. Lubinski, T. Kluz, H. Song, J. P. Tyrer, N. Wentzensen, L. Brinton, B. Trabert, J. Lissowska, J. R. McLaughlin, S. A. Narod, C. Phelan, H. Anton-Culver, A. Ziogas, D. Eccles, I. Campbell, S. A. Gayther, A. Gentry-Maharaj, U. Menon, S. J. Ramus, A. H. Wu, A. Dansonka-Mieszkowska, J. Kupryjanczyk, A. Timorek, L. Szafron, J. M. Cunningham, B. L. Fridley, S. J. Winham, E. V. Bandera, E. M. Poole, T. K. Morgan, E. L. Goode, J. M. Schildkraut, C. L. Pearce, A. Berchuck, P. D. Pharoah, P. M. Webb, G. Chenevix-Trench, H. A. Risch and S. MacGregor (2016). "Assessing the genetic

architecture of epithelial ovarian cancer histological subtypes." (1432-1203 (Electronic)).

Druker, B. J., F. Guilhot, S. G. O'Brien, I. Gathmann, H. Kantarjian, N. Gattermann, M. W. Deininger, R. T. Silver, J. M. Goldman, R. M. Stone, F. Cervantes, A. Hochhaus, B. L. Powell, J. L. Gabrilove, P. Rousselot, J. Reiffers, J. J. Cornelissen, T. Hughes, H. Agis, T. Fischer, G. Verhoef, J. Shepherd, G. Saglio, A. Gratwohl, J. L. Nielsen, J. P. Radich, B. Simonsson, K. Taylor, M. Baccarani, C. So, L. Letvak and R. A. Larson (2006). "Five-year follow-up of patients receiving imatinib for chronic myeloid leukemia." N Engl J Med **355**(23): 2408-2417.

Dubeau, L. (2008). "The cell of origin of ovarian epithelial tumours." Lancet Oncol **9**(12): 1191-1197.

Duesberg, P. and A. McCormack (2013). "Immortality of cancers: a consequence of inherent karyotypic variations and selections for autonomy." Cell Cycle **12**(5): 783-802.

Duncan, A. W., W. Hanlon Newell Ae Fau - Bi, M. J. Bi W Fau - Finegold, S. B. Finegold Mj Fau - Olson, A. L. Olson Sb Fau - Beaudet, M. Beaudet Al Fau - Grompe and M. Grompe (2012). "Aneuploidy as a mechanism for stress-induced liver adaptation." (1558-8238 (Electronic)).

Duncan, A. W., A. E. Hanlon Newell, L. Smith, E. M. Wilson, S. B. Olson, M. J. Thayer, S. C. Strom and M. Grompe (2012). "Frequent aneuploidy among normal human hepatocytes." Gastroenterology **142**(1): 25-28.

Duncan, A. W., M. H. Taylor, R. D. Hickey, A. E. Hanlon Newell, M. L. Lenzi, S. B. Olson, M. J. Finegold and M. Grompe (2010). "The ploidy conveyor of mature hepatocytes as a source of genetic variation." Nature **467**(7316): 707-710.

Elowitz, M. B., A. J. Levine, E. D. Siggia and P. S. Swain (2002). "Stochastic gene expression in a single cell." Science **297**(5584): 1183-1186.

Ercan, S., O. Kaymaz, N. Yucel and A. Orcun (2012). "Serum concentrations of CA 125, CA 15-3, CA 19-9 and CEA in normal pregnancy: a longitudinal study." Arch Gynecol Obstet **285**(3): 579-584.

Farlik, M., N. C. Sheffield, A. Nuzzo, P. Datlinger, A. Schonegger, J. Klughammer and C. Bock (2015). "Single-cell DNA methylome sequencing and bioinformatic inference of epigenomic cell-state dynamics." Cell Rep **10**(8): 1386-1397.

Fearon, E. R. and B. Vogelstein (1990). "A genetic model for colorectal tumorigenesis." Cell **61**(5): 759-767.

Fong, M. Y. and S. S. Kakar (2009). "Ovarian cancer mouse models: a summary of current models and their limitations." J Ovarian Res **2**(1): 12.

Gerlinger, M., S. Horswell, J. Larkin, A. J. Rowan, M. P. Salm, I. Varela, R. Fisher, N. McGranahan, N. Matthews, C. R. Santos, P. Martinez, B. Phillimore, S. Begum, A. Rabinowitz, B. Spencer-Dene, S. Gulati, P. A. Bates, G. Stamp, L. Pickering, M. Gore, D. L. Nicol, S. Hazell, P. A. Futreal, A. Stewart and C. Swanton (2014). "Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing." Nat Genet **46**(3): 225-233.

Gerlinger, M., A. J. Rowan, S. Horswell, J. Larkin, D. Endesfelder, E. Gronroos, P. Martinez, N. Matthews, A. Stewart, P. Tarpey, I. Varela, B. Phillimore, S.

Begum, N. Q. McDonald, A. Butler, D. Jones, K. Raine, C. Latimer, C. R. Santos, M. Nohadani, A. C. Eklund, B. Spencer-Dene, G. Clark, L. Pickering, G. Stamp, M. Gore, Z. Szallasi, J. Downward, P. A. Futreal and C. Swanton (2012). "Intratumor heterogeneity and branched evolution revealed by multiregion sequencing." N Engl J Med **366**(10): 883-892.

Gevaert, O., F. De Smet, T. Van Gorp, N. Pochet, K. Engelen, F. Amant, B. De Moor, D. Timmerman and I. Vergote (2008). "Expression profiling to predict the clinical behaviour of ovarian cancer fails independent evaluation." BMC Cancer **8**: 18.

Gillies, R. J., D. Verduzco and R. A. Gatenby (2012). "Evolutionary dynamics of carcinogenesis and why targeted therapy does not work." Nat Rev Cancer **12**(7): 487-493.

Gordon, L., J. E. Joo Je Fau - Powell, M. Powell Je Fau - Ollikainen, B. Ollikainen M Fau - Novakovic, X. Novakovic B Fau - Li, R. Li X Fau - Andronikos, M. N. Andronikos R Fau - Cruickshank, K. N. Cruickshank Mn Fau - Conneely, A. K. Conneely Kn Fau - Smith, R. S. Smith Ak Fau - Alisch, R. Alisch Rs Fau - Morley, P. M. Morley R Fau - Visscher, J. M. Visscher Pm Fau - Craig, R. Craig Jm Fau - Saffery and R. Saffery (2012). "Neonatal DNA methylation profile in human twins is specified by a complex interplay between intrauterine environmental and genetic factors, subject to tissue-specific influence." (1549-5469 (Electronic)).

Gorelick, R. and H. H. Heng (2010). "Sex reduces genetic variation: a multidisciplinary review." Evolution **65**(4): 1088-1098.

Greaves, M. and C. C. Maley (2012). "Clonal evolution in cancer." Nature **481**(7381): 306-313.

Gupta, P. B., C. M. Fillmore, G. Jiang, S. D. Shapira, K. Tao, C. Kuperwasser and E. S. Lander (2011). "Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells." Cell **146**(4): 633-644.

Gupta, S. (2000). "Hepatic polyploidy and liver growth control." Semin Cancer Biol **10**(3): 161-171.

Hanahan, D. and R. A. Weinberg (2011). "Hallmarks of cancer: the next generation." Cell **144**(5): 646-674.

Harewood, L., F. Schutz, S. Boyle, P. Perry, M. Delorenzi, W. A. Bickmore and A. Reymond (2010). "The effect of translocation-induced nuclear reorganization on gene expression." Genome Res **20**(5): 554-564.

Hartzell, C., O. Ksionda, E. Lemmens, K. Coakley, M. Yang, M. Dail, R. C. Harvey, C. Govern, J. Bakker, T. L. Lenstra, K. Ammon, A. Boeter, S. S. Winter, M. Loh, K. Shannon, A. K. Chakraborty, M. Wabl and J. P. Roose (2013). "Dysregulated RasGRP1 responds to cytokine receptor input in T cell leukemogenesis." Sci Signal **6**(268): ra21.

Heng, H. H. (2007a). "Cancer genome sequencing: the challenges ahead." Bioessays **29**(8): 783-794.

Heng, H. H. (2007b). "Elimination of altered karyotypes by sexual reproduction preserves species identity." Genome **50**(5): 517-524.

Heng, H. H. (2007c). Karyotypic Chaos, A Form Of Non-Clonal Chromosome Aberrations, Plays A Key Role For Cancer Progression And Drug Resistance.

FASEB: Nuclear Structure and Cancer, Vermont Academy, Saxtons River, Vermont.

Heng, H. H. (2008). "The conflict between complex systems and reductionism." JAMA **300**(13): 1580-1581.

Heng, H. H. (2009). "The genome-centric concept: resynthesis of evolutionary theory." Bioessays **31**(5): 512-525.

Heng, H. H. (2010). "Missing heritability and stochastic genome alterations." Nat Rev Genet **11**(11): 813.

Heng, H. H. (2013a). Biocomplexity: challenging reductionism. Handbook on Systems and Complexity in Health. J. P. S. C. M. Martin. New York, Springer**:** 193-208.

Heng, H. H. (2013b). "Genomics: HeLa genome versus donor's genome." Nature **501**(7466): 167.

Heng, H. H. (2013c). "The contribution of genomic heterogeneity. Preface." Cytogenet Genome Res **139**(3): 141-143.

Heng, H. H. (2016). "Heterogeneity-mediated cellular adaptation and its trade-off: searching for the general principles of diseases." Journal of Evaluation in Clinical Practice **In press**.

Heng, H. H., J. Squire and L. C. Tsui (1992). "High-resolution mapping of mammalian genes by in situ hybridization to free chromatin." Proc Natl Acad Sci U S A **89**(20): 9509-9513.

Heng, H. H., J. B. Stevens, G. Liu, S. W. Bremer and C. J. Ye (2004). "Imaging genome abnormalities in cancer research." Cell Chromosome **3**(1): 1.

Heng, H. H., S. W. Bremer, J. Stevens, K. J. Ye, F. Miller, G. Liu and C. J. Ye (2006a). "Cancer progression by non-clonal chromosome aberrations." J Cell Biochem **98**(6): 1424-1435.

Heng, H. H., G. Liu, S. Bremer, K. J. Ye, J. Stevens and C. J. Ye (2006b). "Clonal and non-clonal chromosome aberrations and genome variation and aberration." Genome **49**(3): 195-204.

Heng, H. H., J. B. Stevens, G. Liu, S. W. Bremer, K. J. Ye, P. V. Reddy, G. S. Wu, Y. A. Wang, M. A. Tainsky and C. J. Ye (2006c). "Stochastic cancer progression driven by non-clonal chromosome aberrations." J Cell Physiol **208**(2): 461-472.

Heng, H. H., S. W. Bremer, J. B. Stevens, K. J. Ye, G. Liu and C. J. Ye (2009). "Genetic and epigenetic heterogeneity in cancer: a genome-centric perspective." J Cell Physiol **220**(3): 538-547.

Heng, H. H., J. B. Stevens, S. W. Bremer, K. J. Ye, G. Liu and C. J. Ye (2010). "The evolutionary mechanism of cancer." J Cell Biochem **109**(6): 1072-1084.

Heng, H. H., J. B. Stevens, S. W. Bremer, G. Liu, B. Y. Abdallah and C. J. Ye (2011a). "Evolutionary mechanisms and diversity in cancer." Adv Cancer Res **112**: 217-253

Heng, H. H., G. Liu, J. B. Stevens, S. W. Bremer, K. J. Ye, B. Y. Abdallah, S. D. Horne and C. J. Ye (2011b). "Decoding the genome beyond sequencing: the new phase of genomic research." Genomics **98**(4): 242-252.

Heng, H. H., S. W. Bremer, J. B. Stevens, S. D. Horne, G. Liu, B. Y. Abdallah, K. J. Ye and C. J. Ye (2013a). "Chromosomal instability (CIN): what it is and why it is crucial to cancer evolution." Cancer Metastasis Rev **32**(3-4): 325-340.

Heng, H. H., G. Liu, J. B. Stevens, B. Y. Abdallah, S. D. Horne, K. J. Ye, S. W. Bremer, S. K. Chowdhury and C. J. Ye (2013b). "Karyotype heterogeneity and unclassified chromosomal abnormalities." Cytogenet Genome Res **139**(3): 144-157.

Heng, H. H., S. M. Regan, G. Liu and C. J. Ye (2016). "Why it is crucial to analyze non clonal chromosome aberrations or NCCAs?" Mol Cytogenet **9**: 15.

Heng, H. H. Q. (2016). Debating Cancer: The Paradox in Cancer Research, World Scientific.

Heppner, G. H. (1984). "Tumor heterogeneity." Cancer Res **44**(6): 2259-2265.

Hoglund, M., D. Gisselsson, G. B. Hansen, T. Sall and F. Mitelman (2003). "Ovarian carcinoma develops through multiple modes of chromosomal evolution." Cancer Res **63**(12): 3378-3385.

Horne, S. D., J. B. Stevens, B. Y. Abdallah, G. Liu, S. W. Bremer, C. J. Ye and H. H. Heng (2013). "Why imatinib remains an exception of cancer research." J Cell Physiol **228**(4): 665-670.

Horne, S. D., S. K. Chowdhury and H. H. Heng (2014). "Stress, genomic adaptation, and the evolutionary trade-off." Front Genet **5**: 92.

Horne, S. D., S. A. Pollick and H. H. Heng (2015). "Evolutionary mechanism unifies the hallmarks of cancer." Int J Cancer **136**(9): 2012-2021.

Horne S.D., Y. C. J., Abdallah B.Y., Liu G., Heng, H.H. (2015). "Cancer genome evolution." Translational Cancer Research **4**: 303-313.

Horne, S. D., Ye, C.J., Heng, H.H. (2015). Chromosomal instability (CIN) in cancer. eL*S***:** 1-9.

Horne S.D., L. G., Abdallah B.Y., Stevens J.B., Ye C.J.,Heng, H.H. (2016). "Effective drug treatment induces drug resistance through rapid genome alteration-mediated cancer evolution." Submitted.

Hough, C. D., C. A. Sherman-Baust, E. S. Pizer, F. J. Montz, D. D. Im, N. B. Rosenshein, K. R. Cho, G. J. Riggins and P. J. Morin (2000). "Large-scale serial analysis of gene expression reveals genes differentially expressed in ovarian cancer." Cancer Res **60**(22): 6281-6287.

Huang, S. (2012). "Tumor progression: Chance and necessity in Darwinian and Lamarckian somatic (mutationless) evolution." Prog Biophys Mol Biol **110**(1): 69-86.

Huang, S. (2013). "Genetic and non-genetic instability in tumor progression: link between the fitness landscape and the epigenetic landscape of cancer cells." Cancer Metastasis Rev **32**(3-4): 423-448.

Huang, S., I. Ernberg and S. Kauffman (2009). "Cancer attractors: a systems view of tumors from a gene network dynamics and developmental perspective." Semin Cell Dev Biol **20**(7): 869-876.

Hulten, M. A., J. Jonasson, E. Iwarsson, P. Uppal, S. G. Vorsanova, Y. B. Yurov and I. Y. Iourov (2013). "Trisomy 21 mosaicism: we may all have a touch of Down syndrome." Cytogenet Genome Res **139**(3): 189-192.

Iourov, I. Y., S. G. Vorsanova and Y. B. Yurov (2008). "Chromosomal mosaicism goes global." Mol Cytogenet **1**: 26.

Jacob, F., M. Meier, R. Caduff, D. Goldstein, T. Pochechueva, N. Hacker, D. Fink and V. Heinzelmann-Schwarz (2011). "No benefit from combining HE4 and CA125 as ovarian tumor markers in a clinical setting." Gynecol Oncol **121**(3): 487-491.

Jazaeri, A. A., C. J. Yee, C. Sotiriou, K. R. Brantley, J. Boyd and E. T. Liu (2002). "Gene expression profiles of BRCA1-linked, BRCA2-linked, and sporadic ovarian cancers." J Natl Cancer Inst **94**(13): 990-1000.

Jin, Y., H. Zhang, S. W. Tsao, C. Jin, M. Lv, B. Strombeck, J. Wiegant, T. S. Wan, P. W. Yuen and Y. L. Kwong (2004). "Cytogenetic and molecular genetic characterization of immortalized human ovarian surface epithelial cell lines: consistent loss of chromosome 13 and amplification of chromosome 20." Gynecol Oncol **92**(1): 183-191.

Kafali, H., H. Artuc and N. Demir (2004). "Use of CA125 fluctuation during the menstrual cycle as a tool in the clinical diagnosis of endometriosis; a preliminary report." Eur J Obstet Gynecol Reprod Biol **116**(1): 85-88.

Kaiser, J. (2012). "Profile: Bert Vogelstein. Cancer genetics with an edge." Science **337**(6092): 282-284.

Kanchi, K. L., K. J. Johnson, C. Lu, M. D. McLellan, M. D. Leiserson, M. C. Wendl, Q. Zhang, D. C. Koboldt, M. Xie, C. Kandoth, J. F. McMichael, M. A. Wyczalkowski, D. E. Larson, H. K. Schmidt, C. A. Miller, R. S. Fulton, P. T. Spellman, E. R. Mardis, T. E. Druley, T. A. Graubert, P. J. Goodfellow, B. J. Raphael, R. K. Wilson and L. Ding (2014). "Integrated analysis of germline and somatic variants in ovarian cancer." Nat Commun **5**: 3156.

Katsnelson, A. (2013). "Momentum grows to make 'personalized' medicine more 'precise'." (1546-170X (Electronic)).

Kerem, B., J. A. Rommens Jm Fau - Buchanan, D. Buchanan Ja Fau - Markiewicz, T. K. Markiewicz D Fau - Cox, A. Cox Tk Fau - Chakravarti, M. Chakravarti A Fau - Buchwald, L. C. Buchwald M Fau - Tsui and L. C. Tsui (1989). "Identification of the cystic fibrosis gene: genetic analysis." (0036-8075 (Print)).

Keren, L., D. van Dijk, S. Weingarten-Gabbay, D. Davidi, G. Jona, A. Weinberger, R. Milo and E. Segal (2015). "Noise in gene expression is coupled to growth rate." Genome Res **25**(12): 1893-1902.

Khalique, L., A. Ayhan, M. E. Weale, I. J. Jacobs, S. J. Ramus and S. A. Gayther (2007). "Genetic intra-tumour heterogeneity in epithelial ovarian cancer and its implications for molecular diagnosis of tumours." J Pathol **211**(3): 286-295.

Kim, D., H. Shin, K. A. Sohn, A. Verma, M. D. Ritchie and J. H. Kim (2014). "Incorporating inter-relationships between different levels of genomic data into cancer clinical outcome prediction." Methods **67**(3): 344-353.

Knutsen, T., H. M. Padilla-Nash, D. Wangsa, L. Barenboim-Stapleton, J. Camps, N. McNeil, M. J. Difilippantonio and T. Ried (2010). "Definitive molecular cytogenetic characterization of 15 colorectal cancer cell lines." Genes Chromosomes Cancer **49**(3): 204-223.

Kobel, M., S. E. Kalloger, N. Boyd, S. McKinney, E. Mehl, C. Palmer, S. Leung, N. J. Bowen, D. N. Ionescu, A. Rajput, L. M. Prentice, D. Miller, J. Santos, K. Swenerton, C. B. Gilks and D. Huntsman (2008). "Ovarian carcinoma subtypes

are different diseases: implications for biomarker studies." PLoS Med **5**(12): e232.

Konstantinopoulos, P. A., D. Spentzos and S. A. Cannistra (2008). "Gene-expression profiling in epithelial ovarian cancer." Nat Clin Pract Oncol **5**(10): 577-587.

Koul, A., S. Malander, N. Loman, T. Pejovic, S. Heim, R. Willen, O. Johannsson, H. Olsson, M. Ridderheim and A. A. Borg (2000). "BRCA1 and BRCA2 mutations in ovarian cancer: Covariation with specific cytogenetic features." Int J Gynecol Cancer **10**(4): 289-295.

Kuczek, T. and D. E. Axelrod (1987). "Tumor cell heterogeneity: divided-colony assay for measuring drug response." Proc Natl Acad Sci U S A **84**(13): 4490-4494.

Kussell, E. and S. Leibler (2005). "Phenotypic diversity, population growth, and information in fluctuating environments." Science **309**(5743): 2075-2078.

Landan, G., N. M. Cohen, Z. Mukamel, A. Bar, A. Molchadsky, R. Brosh, S. Horn-Saban, D. A. Zalcenstein, N. Goldfinger, A. Zundelevich, E. N. Gal-Yam, V. Rotter and A. Tanay (2012). "Epigenetic polymorphism and the stochastic formation of differentially methylated regions in normal and cancerous tissues." Nat Genet **44**(11): 1207-1214.

Lawrence, M. S., P. Stojanov, P. Polak, G. V. Kryukov, K. Cibulskis, A. Sivachenko, S. L. Carter, C. Stewart, C. H. Mermel, S. A. Roberts, A. Kiezun, P. S. Hammerman, A. McKenna, Y. Drier, L. Zou, A. H. Ramos, T. J. Pugh, N. Stransky, E. Helman, J. Kim, C. Sougnez, L. Ambrogio, E. Nickerson, E. Shefler, M. L. Cortes, D. Auclair, G. Saksena, D. Voet, M. Noble, D. Dicara, P. Lin, L. Lichtenstein, D. I. Heiman, T. Fennell, M. Imielinski, B. Hernandez, E. Hodis, S. Baca, A. M.

Dulak, J. Lohr, D. A. Landau, C. J. Wu, J. Melendez-Zajgla, A. Hidalgo-Miranda, A. Koren, S. A. McCarroll, J. Mora, R. S. Lee, B. Crompton, R. Onofrio, M. Parkin, W. Winckler, K. Ardlie, S. B. Gabriel, C. W. Roberts, J. A. Biegel, K. Stegmaier, A. J. Bass, L. A. Garraway, M. Meyerson, T. R. Golub, D. A. Gordenin, S. Sunyaev, E. S. Lander and G. Getz (2013). "Mutational heterogeneity in cancer and the search for new cancer-associated genes." Nature **499**: 214-218.

Lawrenson, L. (2010). Tracking profiles of genomic instability in spontaneous transformation and tumorigenesis. PhD, Wayne State University.

Lee, M. C., F. J. Lopez-Diaz, S. Y. Khan, M. A. Tariq, Y. Dayn, C. J. Vaske, A. J. Radenbaugh, H. J. Kim, B. M. Emerson and N. Pourmand (2014). "Single-cell analyses of transcriptional heterogeneity during drug tolerance transition in cancer cells by RNA sequencing." Proc Natl Acad Sci U S A **111**(44): E4726-4735.

Lengauer, C., K. W. Kinzler and B. Vogelstein (1997). "Genetic instability in colorectal cancers." Nature **386**(6625): 623-627.

Levy, S. F., N. Ziv and M. L. Siegal (2012). "Bet hedging in yeast by heterogeneous, age-correlated expression of a stress protectant." PLoS Biol **10**(5): e1001325.

Ling, S., Z. Hu, Z. Yang, F. Yang, Y. Li, P. Lin, K. Chen, L. Dong, L. Cao, Y. Tao, L. Hao, Q. Chen, Q. Gong, D. Wu, W. Li, W. Zhao, X. Tian, C. Hao, E. A. Hungate, D. V. Catenacci, R. R. Hudson, W. H. Li, X. Lu and C. I. Wu (2015). "Extremely high genetic diversity in a single tumor points to prevalence of non-Darwinian cell evolution." (1091-6490 (Electronic)).

Liu, G., J. B. Stevens, S. D. Horne, B. Y. Abdallah, K. J. Ye, S. W. Bremer, C. J. Ye, D. J. Chen and H. H. Heng (2014). "Genome chaos: survival strategy during crisis." Cell Cycle **13**(4): 528-537.

Lodato, M. A., M. B. Woodworth, S. Lee, G. D. Evrony, B. K. Mehta, A. Karger, S. Lee, T. W. Chittenden, A. M. D'Gama, X. Cai, L. J. Luquette, E. Lee, P. J. Park and C. A. Walsh (2015). "Somatic mutation in single human neurons tracks developmental and transcriptional history." (1095-9203 (Electronic)).

Long, H., S. F. Miller, C. Strauss, C. Zhao, L. Cheng, Z. Ye, K. Griffin, R. Te, H. Lee, C. C. Chen and M. Lynch (2016). "Antibiotic treatment enhances the genome-wide mutation rate of target cells." Proc Natl Acad Sci U S A **113**(18): E2498-2505.

Lynch, M. (2009). "Rate, molecular spectrum, and consequences of human mutation." (1091-6490 (Electronic)).

Maher, B. (2008). "Personal genomes: The case of the missing heritability." (1476-4687 (Electronic)).

Maley, C. C., P. C. Galipeau, J. C. Finley, V. J. Wongsurawat, X. Li, C. A. Sanchez, T. G. Paulson, P. L. Blount, R. A. Risques, P. S. Rabinovitch and B. J. Reid (2006). "Genetic clonal diversity predicts progression to esophageal adenocarcinoma." Nat Genet **38**(4): 468-473.

Mankoo, P. K., R. Shen, N. Schultz, D. A. Levine and C. Sander (2011). "Time to recurrence and survival in serous ovarian tumors predicted from integrated genomic profiles." PLoS One **6**(11): e24709.

Manolio, T. A., F. S. Collins, N. J. Cox, D. B. Goldstein, L. A. Hindorff, D. J. Hunter, M. I. McCarthy, E. M. Ramos, L. R. Cardon, A. Chakravarti, J. H. Cho, A. E.

Guttmacher, A. Kong, L. Kruglyak, E. Mardis, C. N. Rotimi, M. Slatkin, D. Valle, A. S. Whittemore, M. Boehnke, A. G. Clark, E. E. Eichler, G. Gibson, J. L. Haines, T. F. Mackay, S. A. McCarroll and P. M. Visscher (2009). "Finding the missing heritability of complex diseases." Nature **461**(7265): 747-753.

Marusyk, A., V. Almendro and K. Polyak (2012). "Intra-tumour heterogeneity: a looking glass for cancer?" Nat Rev Cancer **12**(5): 323-334.

McClellan, J. and M. C. King (2010). "Genetic heterogeneity in human disease." (1097-4172 (Electronic)).

Meden, H. and A. Fattahi-Meibodi (1998). "CA 125 in benign gynecological conditions." Int J Biol Markers **13**(4): 231-237.

Mehrara, E., E. Forssell-Aronsson, H. Ahlman and P. Bernhardt (2007). "Specific growth rate versus doubling time for quantitative characterization of tumor growth rate." Cancer Res **67**(8): 3970-3975.

Merlo, L. M., J. W. Pepper, B. J. Reid and C. C. Maley (2006). "Cancer as an evolutionary and ecological process." Nat Rev Cancer **6**(12): 924-935.

Moss, E. L., J. Hollingworth and T. M. Reynolds (2005). "The role of CA125 in clinical practice." J Clin Pathol **58**(3): 308-312.

Navin, N., J. Kendall, J. Troge, P. Andrews, L. Rodgers, J. McIndoo, K. Cook, A. Stepansky, D. Levy, D. Esposito, L. Muthuswamy, A. Krasnitz, W. R. McCombie, J. Hicks and M. Wigler (2011). "Tumour evolution inferred by single-cell sequencing." Nature **472**: 90-94.

Negrini, S., T. D. Gorgoulis Vg Fau - Halazonetis and T. D. Halazonetis (2010). "Genomic instability--an evolving hallmark of cancer." (1471-0080 (Electronic)).

Nicholson, J. M. and P. Duesberg (2009). "On the karyotypic origin and evolution of cancer cells." Cancer Genet Cytogenet **194**(2): 96-110.

Nordling, C. O. (1953). "A new theory on cancer-inducing mechanism." Br J Cancer **7**(1): 68-72.

Nossov, V., M. Amneus, F. Su, J. Lang, J. M. Janco, S. T. Reddy and R. Farias-Eisner (2008). "The early detection of ovarian cancer: from traditional methods to proteomics. Can we really do better than serum CA-125?" Am J Obstet Gynecol **199**(3): 215-223.

Nowell, P. C. (1976). "The clonal evolution of tumor cell populations." Science **194**(4260): 23-28.

O'Huallachain, M., K. J. Karczewski, S. M. Weissman, A. E. Urban and M. P. Snyder (2012). "Extensive genetic variation in somatic human tissues." Proc Natl Acad Sci U S A **109**(44): 18018-18023.

Pack, S. D., A. O. Weil Rj Fau - Vortmeyer, W. Vortmeyer Ao Fau - Zeng, J. Zeng W Fau - Li, H. Li J Fau - Okamoto, M. Okamoto H Fau - Furuta, E. Furuta M Fau - Pak, I. A. Pak E Fau - Lubensky, E. H. Lubensky Ia Fau - Oldfield, Z. Oldfield Eh Fau - Zhuang and Z. Zhuang (2005). "Individual adult human neurons display aneuploidy: detection by fluorescence in situ hybridization and single neuron PCR." (1551-4005 (Electronic)).

Pavelka, N., G. Rancati, J. Zhu, W. D. Bradford, A. Saraf, L. Florens, B. W. Sanderson, G. L. Hattem and R. Li (2010). Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. Nature. England. **468:** 321-325.

Pejovic, T., S. Heim, N. Mandahl, B. Baldetorp, B. Elmfors, U. M. Floderus, S. Furgyik, G. Helm, A. Himmelmann, H. Willen and et al. (1992). "Chromosome aberrations in 35 primary ovarian carcinomas." Genes Chromosomes Cancer **4**(1): 58-68.

Pelkmans, L. (2012). "Cell Biology. Using cell-to-cell variability--a new era in molecular biology." Science **336**(6080): 425-426.

Ponder, R. G., N. C. Fonville and S. M. Rosenberg (2005). "A switch from high-fidelity to error-prone DNA double-strand break repair underlies stress-induced mutation." Mol Cell **19**(6): 791-804.

Rao, P. H., C. P. Harris, X. Yan Lu, X. N. Li, S. C. Mok and C. C. Lau (2002). "Multicolor spectral karyotyping of serous ovarian adenocarcinoma." Genes Chromosomes Cancer **33**(2): 123-132.

Roberts, N. J., J. T. Vogelstein, G. Parmigiani, K. W. Kinzler, B. Vogelstein and V. E. Velculescu (2012). "The predictive capacity of personal genome sequencing." Sci Transl Med **4**(133): 133ra158.

Roby, K. F., C. C. Taylor, J. P. Sweetwood, Y. Cheng, J. L. Pace, O. Tawfik, D. L. Persons, P. G. Smith and P. F. Terranova (2000). "Development of a syngeneic mouse model for events related to ovarian cancer." Carcinogenesis **21**(4): 585-591.

Roschke, A. V., K. Stover, G. Tonon, A. A. Schaffer and I. R. Kirsch (2002). "Stable karyotypes in epithelial cancer cell lines despite high rates of ongoing structural and numerical chromosomal instability." Neoplasia **4**(1): 19-31.

Rubin, H. (1984). "Early origin and pervasiveness of cellular heterogeneity in some malignant transformations." Proc Natl Acad Sci U S A **81**(16): 5121-5125.

Sandler, O., S. P. Mizrahi, N. Weiss, O. Agam, I. Simon and N. Q. Balaban (2015). "Lineage correlations of single cell division time as a probe of cell-cycle dynamics." Nature **519**(7544): 468-471.

Schultz, M. D., Y. He, J. W. Whitaker, M. Hariharan, E. A. Mukamel, D. Leung, N. Rajagopal, J. R. Nery, M. A. Urich, H. Chen, S. Lin, Y. Lin, I. Jung, A. D. Schmitt, S. Selvaraj, B. Ren, T. J. Sejnowski, W. Wang and J. R. Ecker (2015). "Human body epigenome maps reveal noncanonical DNA methylation variation." Nature **523**(7559): 212-216.

Sgaramella, V. (2010). "Variability of our somatic (epi)genomes." Science **329**(5987): 32-33.

Shih, I. M., W. Zhou, S. N. Goodman, C. Lengauer, K. W. Kinzler and B. Vogelstein (2001). "Evidence that genetic instability occurs at an early stage of colorectal tumorigenesis." Cancer Res **61**(3): 818-822.

Sieben, N. L., P. Macropoulos, G. M. Roemen, S. M. Kolkman-Uljee, G. Jan Fleuren, R. Houmadi, T. Diss, B. Warren, M. Al Adnani, A. P. De Goeij, T. Krausz and A. M. Flanagan (2004). "In ovarian neoplasms, BRAF, but not KRAS, mutations are restricted to low-grade serous tumours." J Pathol **202**(3): 336-340.

Siegel, R., D. Naishadham and A. Jemal (2012). "Cancer statistics, 2012." CA Cancer J Clin **62**(1): 10-29.

Soslow, R. A. (2008). "Histologic subtypes of ovarian carcinoma: an overview." Int J Gynecol Pathol **27**(2): 161-174.

Stephens, P. J., C. D. Greenman, B. Fu, F. Yang, G. R. Bignell, L. J. Mudie, E. D. Pleasance, K. W. Lau, D. Beare, L. A. Stebbings, S. McLaren, M. L. Lin, D. J.

McBride, I. Varela, S. Nik-Zainal, C. Leroy, M. Jia, A. Menzies, A. P. Butler, J. W. Teague, M. A. Quail, J. Burton, H. Swerdlow, N. P. Carter, L. A. Morsberger, C. Iacobuzio-Donahue, G. A. Follows, A. R. Green, A. M. Flanagan, M. R. Stratton, P. A. Futreal and P. J. Campbell (2011). "Massive genomic rearrangement acquired in a single catastrophic event during cancer development." Cell **144**(1): 27-40.

Stevens, J. B., G. Liu, S. W. Bremer, K. J. Ye, W. Xu, J. Xu, Y. Sun, G. S. Wu, S. Savasan, S. A. Krawetz, C. J. Ye and H. H. Heng (2007). "Mitotic cell death by chromosome fragmentation." Cancer Res **67**(16): 7686-7694.

Stevens, J. B., B. Y. Abdallah, G. Liu, C. J. Ye, S. D. Horne, G. Wang, S. Savasan, M. Shekhar, S. A. Krawetz, M. Huttemann, M. A. Tainsky, G. S. Wu, Y. Xie, K. Zhang and H. H. Heng (2011). "Diverse system stresses: common mechanisms of chromosome fragmentation." Cell Death Dis **2**: e178.

Stevens, J. B., S. D. Horne, B. Y. Abdallah, C. J. Ye and H. H. Heng (2013). "Chromosomal instability and transcriptome dynamics in cancer." Cancer Metastasis Rev **32**(3-4): 391-402.

Stevens, J. B., G. Liu, B. Y. Abdallah, S. D. Horne, K. J. Ye, S. W. Bremer, C. J. Ye, S. A. Krawetz and H. H. Heng (2014). "Unstable genomes elevate transcriptome dynamics." Int J Cancer **134**(9): 2074-2087.

Taetle, R., M. Aickin, L. Panda, J. Emerson, D. Roe, F. Thompson, J. Davis, J. Trent and D. Alberts (1999). "Chromosome abnormalities in ovarian adenocarcinoma: II. Prognostic impact of nonrandom chromosome abnormalities in 244 cases." Genes Chromosomes Cancer **25**(1): 46-52.

Talens, R. P., K. Christensen, H. Putter, G. Willemsen, L. Christiansen, D. Kremer, H. E. Suchiman, P. E. Slagboom, D. I. Boomsma and B. T. Heijmans (2012). "Epigenetic variation during the adult lifespan: cross-sectional and longitudinal data on monozygotic twin pairs." Aging Cell **11**(4): 694-703.

Tang, J., F. Fang, D. F. Miller, J. M. Pilrose, D. Matei, T. H. Huang and K. P. Nephew (2015). "Global DNA methylation profiling technologies and the ovarian cancer methylome." Methods Mol Biol **1238**: 653-675.

TCGA (2011). "Integrated genomic analyses of ovarian carcinoma." Nature **474**(7353): 609-615.

Teicher, B. A. (2006). "Tumor models for efficacy determination." (1535-7163 (Print)).

Tomasetti, C., L. Marchionni, M. A. Nowak, G. Parmigiani and B. Vogelstein (2015). "Only three driver gene mutations are required for the development of lung and colorectal cancers." Proc Natl Acad Sci U S A **112**(1): 118-123.

Tothill, R. W., A. V. Tinker, J. George, R. Brown, S. B. Fox, S. Lade, D. S. Johnson, M. K. Trivett, D. Etemadmoghadam, B. Locandro, N. Traficante, S. Fereday, J. A. Hung, Y. E. Chiew, I. Haviv, D. Gertig, A. DeFazio and D. D. Bowtell (2008). "Novel molecular subtypes of serous and endometrioid ovarian cancer linked to clinical outcome." Clin Cancer Res **14**(16): 5198-5208.

van Dongen, J., M. G. Nivard, G. Willemsen, J. J. Hottenga, Q. Helmer, C. V. Dolan, E. A. Ehli, G. E. Davies, M. van Iterson, C. E. Breeze, S. Beck, H. E. Suchiman, R. Jansen, J. B. van Meurs, B. T. Heijmans, P. E. Slagboom and D. I. Boomsma (2016). "Genetic and environmental influences interact with age and sex in shaping the human methylome." Nat Commun **7**: 11115.

Van Gorp, T., I. Cadron, E. Despierre, A. Daemen, K. Leunen, F. Amant, D. Timmerman, B. De Moor and I. Vergote (2011). "HE4 and CA125 as a diagnostic test in ovarian cancer: prospective validation of the Risk of Ovarian Malignancy Algorithm." Br J Cancer **104**(5): 863-870.

Vanderhyden, B. C., T. J. Shaw and J. F. Ethier (2003). "Animal models of ovarian cancer." Reprod Biol Endocrinol **1**: 67.

Venet, D., J. E. Dumont and V. Detours (2011). "Most random gene expression signatures are significantly associated with breast cancer outcome." PLoS Comput Biol **7**(10): e1002240.

Vereczkey, I., O. Serester, J. Dobos, M. Gallai, O. Szakacs, Z. Szentirmay and E. Toth (2011). "Molecular characterization of 103 ovarian serous and mucinous tumors." Pathol Oncol Res **17**(3): 551-559.

Verhaak, R. G., P. Tamayo, J. Y. Yang, D. Hubbard, H. Zhang, C. J. Creighton, S. Fereday, M. Lawrence, S. L. Carter, C. H. Mermel, A. D. Kostic, D. Etemadmoghadam, G. Saksena, K. Cibulskis, S. Duraisamy, K. Levanon, C. Sougnez, A. Tsherniak, S. Gomez, R. Onofrio, S. Gabriel, L. Chin, N. Zhang, P. T. Spellman, Y. Zhang, R. Akbani, K. A. Hoadley, A. Kahn, M. Kobel, D. Huntsman, R. A. Soslow, A. Defazio, M. J. Birrer, J. W. Gray, J. N. Weinstein, D. D. Bowtell, R. Drapkin, J. P. Mesirov, G. Getz, D. A. Levine and M. Meyerson (2013). "Prognostically relevant gene signatures of high-grade serous ovarian carcinoma." J Clin Invest **123**(1): 517-525.

Visscher, P. M., M. A. R. Medland Se Fau - Ferreira, K. I. Ferreira Ma Fau - Morley, G. Morley Ki Fau - Zhu, B. K. Zhu G Fau - Cornes, G. W. Cornes Bk Fau - Montgomery, N. G.

Montgomery Gw Fau - Martin and N. G. Martin (2006). "Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings." (1553-7404 (Electronic)).

Visscher, P. M. (2008). "Sizing up human height variation." (1546-1718 (Electronic)).

Voet, T., P. Kumar, P. Van Loo, S. L. Cooke, J. Marshall, M. L. Lin, M. Zamani Esteki, N. Van der Aa, L. Mateiu, D. J. McBride, G. R. Bignell, S. McLaren, J. Teague, A. Butler, K. Raine, L. A. Stebbings, M. A. Quail, T. D'Hooghe, Y. Moreau, P. A. Futreal, M. R. Stratton, J. R. Vermeesch and P. J. Campbell (2013). "Single-cell paired-end genome sequencing reveals structural variation per cell cycle." Nucleic Acids Res **41**(12): 6119-6138.

Vogelstein, B., N. Papadopoulos, V. E. Velculescu, S. Zhou, L. A. Diaz, Jr. and K. W. Kinzler (2013). "Cancer genome landscapes." Science **339**(6127): 1546-1558.

Wan, Y. W., C. M. Mach, G. I. Allen, M. L. Anderson and Z. Liu (2014). "On the reproducibility of TCGA ovarian cancer microRNA profiles." (1932-6203 (Electronic)).

Wang, D. and S. Bodovitz (2010). "Single cell analysis: the new frontier in 'omics'." Trends Biotechnol **28**(6): 281-290.

Wang, Y., J. Waters, M. L. Leung, A. Unruh, W. Roh, X. Shi, K. Chen, P. Scheet, S. Vattathil, H. Liang, A. Multani, H. Zhang, R. Zhao, F. Michor, F. Meric-Bernstam and N. E. Navin (2014). "Clonal evolution in breast cancer revealed by single nucleus genome sequencing." Nature **512**(7513): 155-160.

Watson, J. (2013). "Oxidants, antioxidants and the current incurability of metastatic cancers." Open Biol **3**(1): 120144.

Weinstein Jn Fau - Collisson, E. A., G. B. Collisson Ea Fau - Mills, K. R. M. Mills Gb Fau - Shaw, B. A. Shaw Kr Fau - Ozenberger, K. Ozenberger Ba Fau - Ellrott, I. Ellrott K Fau - Shmulevich, C. Shmulevich I Fau - Sander, J. M. Sander C Fau - Stuart and J. M. Stuart (2013). "The Cancer Genome Atlas Pan-Cancer analysis project." (1546-1718 (Electronic)).

Welsh, J. B., P. P. Zarrinkar, L. M. Sapinoso, S. G. Kern, C. A. Behling, B. J. Monk, D. J. Lockhart, R. A. Burger and G. M. Hampton (2001). "Analysis of gene expression profiles in normal and neoplastic ovarian tissue samples identifies candidate molecular markers of epithelial ovarian cancer." Proc Natl Acad Sci U S A **98**(3): 1176-1181.

Widschwendter, M., C. Jiang G Fau - Woods, H. M. Woods C Fau - Muller, H. Muller Hm Fau - Fiegl, G. Fiegl H Fau - Goebel, C. Goebel G Fau - Marth, E. Marth C Fau - Muller-Holzner, A. G. Muller-Holzner E Fau - Zeimet, P. W. Zeimet Ag Fau - Laird, M. Laird Pw Fau - Ehrlich and M. Ehrlich (2004). "DNA hypomethylation and ovarian cancer biology." (0008-5472 (Print)).

Wood, A. R., T. Esko, J. Yang, S. Vedantam, T. H. Pers, S. Gustafsson, A. Y. Chu, K. Estrada, J. Luan, Z. Kutalik, N. Amin, M. L. Buchkovich, D. C. Croteau-Chonka, F. R. Day, Y. Duan, T. Fall, R. Fehrmann, T. Ferreira, A. U. Jackson, J. Karjalainen, K. S. Lo, A. E. Locke, R. Magi, E. Mihailov, E. Porcu, J. C. Randall, A. Scherag, A. A. Vinkhuyzen, H. J. Westra, T. W. Winkler, T. Workalemahu, J. H. Zhao, D. Absher, E. Albrecht, D. Anderson, J. Baron, M. Beekman, A. Demirkan, G. B. Ehret, B. Feenstra, M. F. Feitosa, K. Fischer, R. M. Fraser, A. Goel, J. Gong, A. E. Justice, S. Kanoni, M. E. Kleber, K. Kristiansson, U. Lim, V.

Lotay, J. C. Lui, M. Mangino, I. Mateo Leach, C. Medina-Gomez, M. A. Nalls, D. R. Nyholt, C. D. Palmer, D. Pasko, S. Pechlivanis, I. Prokopenko, J. S. Ried, S. Ripke, D. Shungin, A. Stancakova, R. J. Strawbridge, Y. J. Sung, T. Tanaka, A. Teumer, S. Trompet, S. W. van der Laan, J. van Setten, J. V. Van Vliet-Ostaptchouk, Z. Wang, L. Yengo, W. Zhang, U. Afzal, J. Arnlov, G. M. Arscott, S. Bandinelli, A. Barrett, C. Bellis, A. J. Bennett, C. Berne, M. Bluher, J. L. Bolton, Y. Bottcher, H. A. Boyd, M. Bruinenberg, B. M. Buckley, S. Buyske, I. H. Caspersen, P. S. Chines, R. Clarke, S. Claudi-Boehm, M. Cooper, E. W. Daw, P. A. De Jong, J. Deelen, G. Delgado, J. C. Denny, R. Dhonukshe-Rutten, M. Dimitriou, A. S. Doney, M. Dorr, N. Eklund, E. Eury, L. Folkersen, M. E. Garcia, F. Geller, V. Giedraitis, A. S. Go, H. Grallert, T. B. Grammer, J. Grassler, H. Gronberg, L. C. de Groot, C. J. Groves, J. Haessler, P. Hall, T. Haller, G. Hallmans, A. Hannemann, C. A. Hartman, M. Hassinen, C. Hayward, N. L. Heard-Costa, Q. Helmer, G. Hemani, A. K. Henders, H. L. Hillege, M. A. Hlatky, W. Hoffmann, P. Hoffmann, O. Holmen, J. J. Houwing-Duistermaat, T. Illig, A. Isaacs, A. L. James, J. Jeff, B. Johansen, A. Johansson, J. Jolley, T. Juliusdottir, J. Junttila, A. N. Kho, L. Kinnunen, N. Klopp, T. Kocher, W. Kratzer, P. Lichtner, L. Lind, J. Lindstrom, S. Lobbens, M. Lorentzon, Y. Lu, V. Lyssenko, P. K. Magnusson, A. Mahajan, M. Maillard, W. L. McArdle, C. A. McKenzie, S. McLachlan, P. J. McLaren, C. Menni, S. Merger, L. Milani, A. Moayyeri, K. L. Monda, M. A. Morken, G. Muller, M. Muller-Nurasyid, A. W. Musk, N. Narisu, M. Nauck, I. M. Nolte, M. M. Nothen, L. Oozageer, S. Pilz, N. W. Rayner, F. Renstrom, N. R. Robertson, L. M. Rose, R. Roussel, S. Sanna, H. Scharnagl, S.

Scholtens, F. R. Schumacher, H. Schunkert, R. A. Scott, J. Sehmi, T. Seufferlein, J. Shi, K. Silventoinen, J. H. Smit, A. V. Smith, J. Smolonska, A. V. Stanton, K. Stirrups, D. J. Stott, H. M. Stringham, J. Sundstrom, M. A. Swertz, A. C. Syvanen, B. O. Tayo, G. Thorleifsson, J. P. Tyrer, S. van Dijk, N. M. van Schoor, N. van der Velde, D. van Heemst, F. V. van Oort, S. H. Vermeulen, N. Verweij, J. M. Vonk, L. L. Waite, M. Waldenberger, R. Wennauer, L. R. Wilkens, C. Willenborg, T. Wilsgaard, M. K. Wojczynski, A. Wong, A. F. Wright, Q. Zhang, D. Arveiler, S. J. Bakker, J. Beilby, R. N. Bergman, S. Bergmann, R. Biffar, J. Blangero, D. I. Boomsma, S. R. Bornstein, P. Bovet, P. Brambilla, M. J. Brown, H. Campbell, M. J. Caulfield, A. Chakravarti, R. Collins, F. S. Collins, D. C. Crawford, L. A. Cupples, J. Danesh, U. de Faire, H. M. den Ruijter, R. Erbel, J. Erdmann, J. G. Eriksson, M. Farrall, E. Ferrannini, J. Ferrieres, I. Ford, N. G. Forouhi, T. Forrester, R. T. Gansevoort, P. V. Gejman, C. Gieger, A. Golay, O. Gottesman, V. Gudnason, U. Gyllensten, D. W. Haas, A. S. Hall, T. B. Harris, A. T. Hattersley, A. C. Heath, C. Hengstenberg, A. A. Hicks, L. A. Hindorff, A. D. Hingorani, A. Hofman, G. K. Hovingh, S. E. Humphries, S. C. Hunt, E. Hypponen, K. B. Jacobs, M. R. Jarvelin, P. Jousilahti, A. M. Jula, J. Kaprio, J. J. Kastelein, M. Kayser, F. Kee, S. M. Keinanen-Kiukaanniemi, L. A. Kiemeney, J. S. Kooner, C. Kooperberg, S. Koskinen, P. Kovacs, A. T. Kraja, M. Kumari, J. Kuusisto, T. A. Lakka, C. Langenberg, L. Le Marchand, T. Lehtimaki, S. Lupoli, P. A. Madden, S. Mannisto, P. Manunta, A. Marette, T. C. Matise, B. McKnight, T. Meitinger, F. L. Moll, G. W. Montgomery, A. D. Morris, A. P. Morris, J. C. Murray, M. Nelis, C. Ohlsson, A. J. Oldehinkel, K. K. Ong, W. H. Ouwehand, G.

Pasterkamp, A. Peters, P. P. Pramstaller, J. F. Price, L. Qi, O. T. Raitakari, T. Rankinen, D. C. Rao, T. K. Rice, M. Ritchie, I. Rudan, V. Salomaa, N. J. Samani, J. Saramies, M. A. Sarzynski, P. E. Schwarz, S. Sebert, P. Sever, A. R. Shuldiner, J. Sinisalo, V. Steinthorsdottir, R. P. Stolk, J. C. Tardif, A. Tonjes, A. Tremblay, E. Tremoli, J. Virtamo, M. C. Vohl, P. Amouyel, F. W. Asselbergs, T. L. Assimes, M. Bochud, B. O. Boehm, E. Boerwinkle, E. P. Bottinger, C. Bouchard, S. Cauchi, J. C. Chambers, S. J. Chanock, R. S. Cooper, P. I. de Bakker, G. Dedoussis, L. Ferrucci, P. W. Franks, P. Froguel, L. C. Groop, C. A. Haiman, A. Hamsten, M. G. Hayes, J. Hui, D. J. Hunter, K. Hveem, J. W. Jukema, R. C. Kaplan, M. Kivimaki, D. Kuh, M. Laakso, Y. Liu, N. G. Martin, W. Marz, M. Melbye, S. Moebus, P. B. Munroe, I. Njolstad, B. A. Oostra, C. N. Palmer, N. L. Pedersen, M. Perola, L. Perusse, U. Peters, J. E. Powell, C. Power, T. Quertermous, R. Rauramaa, E. Reinmaa, P. M. Ridker, F. Rivadeneira, J. I. Rotter, T. E. Saaristo, D. Saleheen, D. Schlessinger, P. E. Slagboom, H. Snieder, T. D. Spector, K. Strauch, M. Stumvoll, J. Tuomilehto, M. Uusitupa, P. van der Harst, H. Volzke, M. Walker, N. J. Wareham, H. Watkins, H. E. Wichmann, J. F. Wilson, P. Zanen, P. Deloukas, I. M. Heid, C. M. Lindgren, K. L. Mohlke, E. K. Speliotes, U. Thorsteinsdottir, I. Barroso, C. S. Fox, K. E. North, D. P. Strachan, J. S. Beckmann, S. I. Berndt, M. Boehnke, I. B. Borecki, M. I. McCarthy, A. Metspalu, K. Stefansson, A. G. Uitterlinden, C. M. van Duijn, L. Franke, C. J. Willer, A. L. Price, G. Lettre, R. J. Loos, M. N. Weedon, E. Ingelsson, J. R. O'Connell, G. R. Abecasis, D. I. Chasman, M. E. Goddard, P. M. Visscher, J. N. Hirschhorn and T. M. Frayling (2014). "Defining the role of common variation in

the genomic and biological architecture of adult human height." (1546-1718 (Electronic)).

Wood, L. D., D. W. Parsons, S. Jones, J. Lin, T. Sjoblom, R. J. Leary, D. Shen, S. M. Boca, T. Barber, J. Ptak, N. Silliman, S. Szabo, Z. Dezso, V. Ustyanksky, T. Nikolskaya, Y. Nikolsky, R. Karchin, P. A. Wilson, J. S. Kaminker, Z. Zhang, R. Croshaw, J. Willis, D. Dawson, M. Shipitsin, J. K. Willson, S. Sukumar, K. Polyak, B. H. Park, C. L. Pethiyagoda, P. V. Pant, D. G. Ballinger, A. B. Sparks, J. Hartigan, D. R. Smith, E. Suh, N. Papadopoulos, P. Buckhaults, S. D. Markowitz, G. Parmigiani, K. W. Kinzler, V. E. Velculescu and B. Vogelstein (2007). "The genomic landscapes of human breast and colorectal cancers." Science **318**(5853): 1108-1113.

Yates, L. R. and P. J. Campbell (2012). "Evolution of the cancer genome." Nat Rev Genet **13**(11): 795-806.

Ye, C. J., G. Liu, S. W. Bremer and H. H. Heng (2007). "The dynamics of cancer chromosomes and genomes." Cytogenet Genome Res **118**(2-4): 237-246.

Ye, C. J., J. B. Stevens, G. Liu, S. W. Bremer, A. S. Jaiswal, K. J. Ye, M. F. Lin, L. Lawrenson, W. D. Lancaster, M. Kurkinen, J. D. Liao, C. G. Gairola, M. P. Shekhar, S. Narayan, F. R. Miller and H. H. Heng (2009). "Genome based cell population heterogeneity promotes tumorigenicity: the evolutionary mechanism of cancer." J Cell Physiol **219**(2): 288-300.

Yoon, M. S., K. U. Suh Ds Fau - Choi, M. Y. Choi Ku Fau - Sol, D. H. Sol My Fau - Shin, W. Y. Shin Dh Fau - Park, J. H. Park Wy Fau - Lee, S. M. Lee Jh Fau - Jeong, W. G. Jeong Sm Fau - Kim, N. R. Kim Wg Fau - Shin and N. R. Shin

(2010). "High-throughput DNA hypermethylation profiling in different ovarian epithelial cancer subtypes using universal bead array." (1791-2431 (Electronic)).

Yoshihara, K., A. Tajima, T. Yahata, S. Kodama, H. Fujiwara, M. Suzuki, Y. Onishi, M. Hatae, K. Sueyoshi, Y. Kudo, K. Kotera, H. Masuzaki, H. Tashiro, H. Katabuchi, I. Inoue and K. Tanaka (2010). "Gene expression profile for predicting survival in advanced-stage serous ovarian cancer across two independent datasets." PLoS One **5**(3): e9615.

Yousef, G. M., M. E. Polymeris, G. M. Yacoub, A. Scorilas, A. Soosaipillai, C. Popalis, S. Fracchioli, D. Katsaros and E. P. Diamandis (2003). "Parallel overexpression of seven kallikrein genes in ovarian cancer." Cancer Res **63**(9): 2223-2227.

Zhang, P., C. Wang, L. Cheng, L. Guo, W. Liu, Z. Zhang, Y. Huang, Q. Ou, X. Wen and Y. Tian (2015). "Comparison of HE4, CA125, and ROMA Diagnostic Accuracy: A Prospective and Multicenter Study for Chinese Women With Epithelial Ovarian Cancer." Medicine (Baltimore) **94**(52): e2402.

Zhang, W., Y. Liu, N. Sun, D. Wang, J. Boyd-Kirkup, X. Dou and J. D. Han (2013). "Integrating genomic, epigenomic, and transcriptomic features reveals modular signatures underlying poor prognosis in ovarian cancer." Cell Rep **4**(3): 542-553.

Zhu, J., D. Heinecke, W. A. Mulla, W. D. Bradford, B. Rubinstein, A. Box, J. S. Haug and R. Li (2015). "Single-Cell Based Quantitative Assay of Chromosome Transmission Fidelity." G3 (Bethesda) **5**(6): 1043-1056.

ABSTRACT

FUZZY INHERITANCE: A NOVEL FORM OF SOMATIC CELL INHERITANCE THAT REGULATES CELL POPULATION HETEROGENEITY

by

BATOUL Y. ABDALLAH

August 2016

Advisor: Dr. Henry Heng

Major: Molecular Biology and Genetics

Degree: Doctor of Philosophy

Multi-level heterogeneity is a characteristic feature of cancer cell populations. However, how a cell population regulates and maintains its cell population heterogeneity is not well understood. Based on conventional theories of genetic inheritance, cell division is precise, where a daughter cell inherits an identical karyotype, more or less, from its mother cell. Errors that are generated during cell division occur at low frequencies and accumulate over prolonged time periods to accumulate. However, the overwhelming heterogeneity found in unstable cancers is largely inconsistent with current models of genetic inheritance. In order to determine the mechanism of how heterogeneity is regulated, the pattern of inherited traits, including karyotype and growth rate, are compared in cell lines with different degrees of genome instability. Single cell and population-based assays were conducted and illustrate the following: 1) single unstable cells cannot pass a specific karyotype or growth rate and instead pass a heterogeneous array of karyotypes and growth rates; 2) genome heterogeneity is linked to other heterogeneous features of the system, like growth heterogeneity; 3) cells that are outliers dominate cell population dynamics when the cell population is unstable; and 4) the statistical average does not give an accurate portrayal of unstable cell populations. Altogether, this suggests that genome instability leads to

genome replacement-mediated macro-cellular evolution that precludes the clonal expansion of a few abnormal cells; and 2) a given degree of heterogeneity can be inherited from a single cell. Because a given degree of heterogeneity is inherited, and the specific variants change between cell passages, this inheritance is termed fuzzy inheritance. According to fuzzy inheritance, rather than passing specific changes, the potential to generate genomic variation is passed. Fuzzy inheritance provides a cell population with the necessary evolvability and explains how heterogeneity is regulated and maintained in normal tissue and in cancer cells.

AUTOBIOGRAPHICAL STATEMENT

Education

PhD, Molecular Biology and Genetics — Summer 2016
Wayne State University School of Medicine — Detroit, MI

Masters in Public Administration — Summer 2009
University of Michigan — Dearborn, MI

Bachelor of Arts, Integrated Science & History — Winter 2004
University of Michigan — Dearborn, MI

Selected Publications

Abdallah BY, Horne SD, Liu G, Stevens JB, Ye CJ, Bremer SW, Gorelick,R, Krawetz SA, Heng HH. "Fuzzy inheritance: a novel form of inheritance that regulates tumor heterogeneity." (submitted)

Liu G, Stevens JB, Horne SD, **Abdallah BY,** Ye KJ, Bremer SW, Ye CJ, Chen, DJ, Heng HH. (2015) "Genome Chaos: survival strategy during crisis." Cell Cycle. 13(4): 528-537.

Stevens JB, Liu G, **Abdallah BY**, Horne SD, Ye KJ, Bremer SW, Ye CJ, Krawetz SA, Heng HH. (2015) "Unstable genomes elevate transcriptome dynamics." Int J Cancer. 134(9): 2074-2087.

Abdallah BY, Horne SD, Kurkinen M, Stevens JB, Liu G, Ye KJ, Barbat J, Bremer SW Heng HH. (2014) "Ovarian cancer evolution through stochastic genome alterations: defining the genomic role in ovarian cancer." Systems Biology in Reproductive Medicine. 60(1): 2-13.

Abdallah BY, Horne SD, Stevens JB, Liu G, Ying AY, Vanderyhyden B, Krawetz SA, Heng HH. (2013) "Single cell heterogeneity: why unstable genomes are incompatible with average profiles." Cell Cycle. 12(23): 3640-3649.

Selected Oral Presentations

Inherited heterogeneity: a novel from of inheritance that regulates tumor heterogeneity — Oct 2014
Center for Molecular Medicine Seminar Series
Wayne State University School of Medicine. Detroit, MI

Single cell heterogeneity of cancer cell populations: why unstable genomes are incompatible with average profiles — Sept 2013
Graduate Student Research Day
Wayne State University School of Medicine. Detroit, MI

Nonclonal chromosomal aberrations as genomic markers of ovarian cancer — May 2012
Michigan Alliance for Reproductive Technologies research symposium Detroit, MI