1-1-2010

Tracking profiles of genomic instability in spontaneous transformation and tumorigenesis

Lesley Lawrenson
Wayne State University,

TRACKING PROFILES OF GENOMIC INSTABILITY IN SPONTANEOUS TRANSFORMATION AND TUMORIGENESIS

by

LESLEY EILEEN LAWRENSON

DISSERTATION

Submitted to the Graduate School

of Wayne State University

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2010

MAJOR: MOLECULAR MEDICINE AND GENETICS

Approved by:

Advisor	Date

DEDICATION

This work is dedicated to my husband, family, friends, mentors, and colleagues with deepest gratitude for their guidance and support.

ACKNOWLEDGEMENTS

The submission of this dissertation brings to an end a wonderful period in which I was a graduate student in Molecular Medicine and Genetics at Wayne State University School of Medicine. Along this path, my mentors, friends and family have helped me grow in sharing with me the many joyous moments as well as the challenges presented throughout the development of this work. I am forever indebted to those who encouraged me to continue to pursue my education. I am deeply grateful to my advisors, Prof. Henry Heng and Prof. Wayne Lancaster for their guidance, encouragement, and faith. Both have served as wonderful role models for me. Their knowledge, reasoning, and experience will continue to serve me well in my future career.

I wish to thank the members of my committee for their patience, wisdom, and good humor. I am extremely grateful for the contributions of Dr. Lucie Gregoire for sharing her knowledge of cell culture of this model and for her guidance throughout this time; Dr. Raja Rabah for evaluations of the tumor histopathology for the animal portion of this work; Dr. Josh Stevens for his technical assistance, his willingness to help, and for our numerous discussions in the development of this manuscript; and Guo Liu for his contributions with cell culture and SKY analysis. Additionally, I am grateful to Jayson Field, MD for his clinical expertise and for providing me the opportunity to participate in the surgical and clinical care of ovarian cancer patients.

I would like to thank my family. Thank you to my loving parents, for instilling in me the desire for knowledge and the determination to succeed. Finally, I thank my husband, Gavin, for his unending support and for always believing in me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. BACKGROUND AND SIGNIFICANCE

1.1 Ovarian Cancer Etiology

Ovarian cancer is the most prevalent and deadly gynecologic malignancy in the United States, and, as the fifth most common cause of cancer death in women, accounts for approximately 16,000 deaths per year [1]. Nearly all ovarian cancers originate from the epithelial cells comprising the outermost surface of the ovary. There are four major subtypes of epithelial ovarian cancers (EOC) which are categorized by their histopathologic characteristics [2]. EOC variability is present not only in the morphology of histopathologic samples, but also in the wide variety of unique karyotypes, and the numerous low frequency genetic abnormalities associated with this disease. Specific genetic alterations tend to be inconsistent between studies, but include changes on many levels including DNA sequence, copy number, methylation status, and miRNA levels [3-4]. Features on each level have been correlated with resistance to chemotherapy and with differing patient survival rates [5]. Despite an increased knowledge of the etiology of this disease, improvements in surgical techniques, advancements in chemotherapeutic treatments, and the characterization of ovarian cancer genomes at many levels [5-6], the morbidity and mortality associated with EOC overall has remained largely unchanged (0.39 in 1980–1989 to 0.43 in 1990–1997) [7]. EOC is often asymptomatic in its earliest stages and strategies with sufficient sensitivity and specificity to detect early-stage disease are currently not available. As a result, most patients are diagnosed with late-stage disseminated EOC which typically becomes resistant to both standard and combination chemotherapies [7-9] and carries a prognosis of only 20% survival over five years [7]. Therefore, the specific challenges presented by the heterogeneous and clinically insidious nature of EOC underlie the rationale for continued focus on these areas of research to improve upon the high morbidity and mortality currently associated with this disease.

Due to the paucity of early-stage clinical samples, the initiating events in ovarian cancer transformation are not well understood. It is widely believed that the vast majority of ovarian

cancers originate from the single layer of simple cuboidal to low pseudostratified columnar epithelial cells encasing the mammalian ovary [10-11]. The genetic and morphologic heterogeneity seen in transformed EOC cells has been postulated as a mechanism by which the early homogeneous surface epithelial cells could and differentiate from their common origin to form the multiple malignant ovarian cancer subtypes. The resulting cancers are clinically categorized into four distinct histopathologic categories with the following distribution of incidence: 1) papillary serous (50-60%), 2) endometrioid (25%), 3) mucinous (4%), 4) and clear cell (4%) [12-14]. These subtypes are distinguished by their morphologic features, which resemble those of the specialized epithelia of the reproductive tract that derive from the Müllerian ducts. Specifically, papillary serous cancers resemble cancers of the fallopian tube, endometrioid ovarian cancers are characterized by endometrial-like glands, and mucinous ovarian cancer types resemble endocervical and intestinal epithelial cells [15-16]. The ability of the ovarian surface epithelium (OSE) to assume these varying Müllerian-like features has recently been linked to the inappropriate activation of homeobox (HOX) genes that control patterning of the reproductive tract [17-19]. This finding adds yet another regulatory level by which ovarian cancer may generate heterogeneity by phenotype. The dynamic interactions between the multiple regulatory systems that govern differentiation and spontaneous neoplastic transformation from a normal epithelial cell to one of many tumorigenic ovarian cancer phenotypes are beginning to be explored. As improvements in prevention and early detection will likely have a significant impact on morbidity and mortality of this disease, emphasis is placed on expanding the knowledge bases of the early transformative events.

The specific mechanisms underlying spontaneous transformation in EOC are not well understood. Epidemiologic data have provided valuable clues as well as data supporting several seemingly paradoxical theories that explain precisely how ovarian cancer might form. Three major interrelated theories have been proposed to explain the epidemiologic data associated with ovarian cancer susceptibility [20-21]. The first of these is based on the link

between ovarian cancer risk and the number of ovulatory cycles and is thus termed the incessant ovulation hypothesis. It is based on the concept that repetitive wounding and cell proliferation in postovulatory repair of the ovarian surface epithelium results in an accumulation of genetic abnormalities that promote cellular transformation [20]. If these damaged epithelial cells are invaginated into the tumor-promoting stromal environment, aberrant autocrine and paracrine stimulation by trophic hormones, phospholipids, and vascular endothelial growth factor are thought to accelerate cellular transformation during post-ovulatory growth and epithelial regeneration [15-16]. These putative mechanisms are incongruent in a temporal sense with the findings that stoppage of ovulation for a disproportionately short time interval significantly reduces the lifetime risk of ovarian cancer. Short-term stoppges in the ovulatory cycle reflect less than 1/40 of a woman's total lifetime ovulations and include the carriage of a single full-term pregnancy, less than one year of breast feeding, or as little as 6 months of oral contraceptive usage [22]. An alternative transformation mechanism has been hypothesized that centers on the reduction in pituitary gonadotropin levels occuring with these same protective events. This theory implicates surges of pituitary gonadotropins occurring at ovulation and persistent high gonadotropin levels following menopause in the activation of mitogenic pathways and the resulting accumulation of genetic changes leading to carcinogenesis [23-25]. The third theory for ovarian cancer causation states that inflammation and changes in redox potential in the setting of ovulation and OSE repair account for the increased risk of ovarian cancer associated with pro-inflammatory conditions. These include talc/asbestos exposure, endometriosis, pelvic inflammatory disease, and mumps infection [26-27]. Together, these theories are largely based on altered incidence rates generated from physiologic and epidemiologic data with some molecular and cellular correlates. Each of these theories explains a portion of disease causation in terms of modifiable risk factors, however, optimal minimization or elimination of these factors does not prevent EOC completely. Therefore these variables likely serve as cofactors for underlying events resulting in cellular transformation.

At the molecular level, studies in cell biology demonstrate that the epithelial cells adjacent to the site of follicular rupture sustain sub-lethal oxidative DNA damage and are more likely in healing to give rise to cells with genetic alterations and a transformed phenotype [28]. This is supported by clinical reports that have documented cellular atypia including metaplasia, hyperplasia, and ovarian intraepithelial neoplasia of the surface epithelium adjacent to invasive carcinoma [29-31]. Despite this knowledge, our current understanding of the mechanisms of surface epithelial cell transformation is limited by the paucity of early stage clinical samples [32]. This is particularly true in the characterization of how early-stage transformative events lead to progression and eventually tumorigenesis. Increased understanding of how the initiating events underlying the transformation process and the inherent plasticity of the ovarian surface epithelium could link to later transformative stages will likely reduce the morbidity and mortality currently associated with ovarian cancer through the identification of new strategies for detection, diagnosis, and treatment.

1.2 Animal Models of Ovarian Cancer

A number of *in vitro* and *in vivo* murine models have been developed to study the characteristics and behavior of EOC, each with its own set of strengths and limitations. Chemically induced cellular transformation models mark some of the first successful attempts to induce EOC in laboratory mice and rats and, with varying success and timelines, linked DNA damage by 1,3-butadiene [33], and dimethlybenz(a)anthracene (DMBA) 20-methylcholanthrene [34], to EOC. The applicability of data from carcinogenic induction models is unclear, as evidence supporting the association of chemical carcinogenesis and ovarian cancer etiology in humans is largely absent [3].

In addition to chemical models of DNA damage, genetically induced ovarian epithelial tumor models have also provided valuable information regarding gene alterations linked to tumorigenesis. One of the first genetically induced models was created through the cross breeding of transgenic mice expressing the avian retroviral receptor (TVA) with $p53^{-/-}$ mice. The

induction of transformation by oncogenic alteration of any two of three (*c-myc, Kras, and Akt*) oncogenes on the $p53^{-/-}$ background was sufficient to increase the rate of tumorigenesis [35]. A second approach demonstrated that injection of the simian virus 40 T-antigen (SV40 T-Ag) into the male pronucleus of half-day old embryos was sufficient to generate ovarian tumors.

However, tumor outcomes differed significantly from human EOC because, despite linking of the SV40 Tag to the epithelial ovarian surface with the Müllerian Inhibitory Substance Type II Receptor (*MISIIR*) promoter, uterine masses and polycystic kidneys also developed [36]. In a third model, the targeted expression of murine *PIK3CA* under the *MISIIR* promoter resulted in mouse ovarian surface epithelial cell (MOSEC) hyperplasia, but cells were not tumorigenic after 18 months. However, both *PIK3CA* and mutant *K-ras* were still thought hold promise for spontaneous transformation models as they were shown to increase anchorage-independent growth of cultured MOSEC [37]. A fourth model used transient (21 day) inactivation of *p53* and *Rb* by adenoviral infection within the ovarian bursa. These experiments demonstrated that early causative events in ovarian cancer transformation need not be maintained for tumorigenesis as early transient inactivation resulted in ovarian tumors in 33/34 mice at ~225 days [38]. Adenoviral administration was also utilized for in a fifth genetic model that showed that alteration of P*ten* when paired with *K-ras* [39] or *Apc* [40] generates invasive primary ovarian endometrioid adenocarcinomas. Because of its strong association with inherited risk for ovarian cancer [41] the conditional inactivation of *Brca1* in the MOSEC was speculated to increase ovarian cancer formation. Additionally, the *Brca1* protein is involved in diverse cellular events and functions related to genomic instability, including homology-directed DNA repair [42], transcriptional regulation [43-44], chromatin remodeling [45] ubiquitin ligation [46], and centrosome amplification [47]. However in this seventh genetic model, conditional inactivation of *Brca1* introns 5-13 ($Brca1^{\Delta 55\text{-}13}$) using AdCre administration in the ovarian bursa resulted in surface epithelium hyperplasia and the formation of inclusion cysts but not in tumor formation after 240 days *in vivo*. Conditional deletion of *Brca1 in vitro* resulted in slowed growth that was

reversed by *p53* deletion, as well as increased sensitivity to the DNA damaging agent cisplatin [48]. The last relevant genetic model showed that *Brca1* inactivation coupled with either *p53* or *Rb* inactivation also failed to yield epithelial tumors. Unpredictably, leiomyosarcomas formed in adjacent cells and showed more rapid progression with concomitant inactivation of *Brca1* and *p53*, but not *Rb* and *p53* [38]. Considering the relative successes of these eight MOSEC models for genetically induced transformation, the unpredictability of success and the limitations of each, coupled with the fact that most human EOCs are not predictably linked to a specific genetic alteration, brings into question the utility and significance of genetic models for comparison with human EOC.

Animal models that most closely mimic the human disease are vital not only in furthering our understanding of the biological and genetic factors that influence disease phenotype, but also serve as the basis for the development and of detection and treatment strategies. As early-stage human ovarian cancer samples are rarely available, and human OSE do not spontaneously transform in culture, animal models are heavily relied upon as investigative tools. These models afford investigators the opportunity to study the mechanisms underlying cellular transformation longitudinally and from its earliest events. Considering the various chemical and genetic animal models currently available for ovarian cancer, very few develop ovarian tumors spontaneously. Low tumor prevalence and extended time to first occurrence of tumors also renders many of these models impractical for experimental studies [49-51]. Each MOSEC model that incorporates genetic alteration in the induction of tumorigenesis has proven somewhat paradoxical, as the action of gene alteration using MOSEC *in vitro* does not equate to EOC causation *in vivo*. Together, these models suggest that either *ex vivo* manipulation or embryonic transgene expression do not accurately represent human epithelial ovarian cancer (EOCs), which arise spontaneously and in a vast majority of cases are linked to age and to the total number of ovulatory cycles. Additionally the significance of data generated from studies where transformed human cells are injected into immune-compromised mice is confounded by

the absence of immune regulatory mechanisms known to block tumor growth and immune mediated tumor-stromal interactions.

1.3 Utility of MOSEC Model

More recently, syngeneic rodent models have been developed where ovarian surface epithelial cells were transformed *in vitro*. Transfused cells were then injected into the peritoneal cavity of immune competent animals *in vivo* to study tumorigenesis. Initially, spontaneous transformation of OSE *in vitro* was shown to occur in a matter of weeks in rat [52-53], with multiple divergent cytogenetic changes in each transformed cell line. A syngeneic mouse model was developed soon after by Roby et al. [54] with significantly slower transformative properties that facilitated the characterization of key transformative stages. Similar to the rat model, in the mouse ovarian surface epithelial cell (MOSEC) transformation occurred spontaneously *in vitro* with repeated passages, and the cells proved tumorigenic in C57BL6 mice. Variations on this model have since been used by numerous investigators to study the transformation process and the tumorigenic features of ovarian cancer, including cytoskeletal and adhesion protein alterations, tumor stromal interactions, proliferation, differential activation of survival and apoptosis pathways, and endothelial growth factor expression [55-59].

The similarity of genetic alterations between late stage transformed cells of the MOSEC model and genetic alterations previously characterized from human ovarian cancers provides further support for the appropriateness of this model. For example, array comparative genomic hybridization (array-CGH) of late-stage transformed MOSEC revealed 80% conservation of synteny between the MOSEC and human ovarian malignancies of epithelial origin [57]. Parallel pathways of karyotypic change were also reported between species [57]. Additionally, genome-wide transcriptional profiles of late-stage tumorigenic MOSEC lines revealed a similar gene expression pattern to human carcinomas of varying cell morphological, behavioral, and prognoses [60]. These include a high percentage of shared, differentially expressed genes, similar alterations in major signal transduction pathways, and common alterations in pathways

of cellular metabolism [61]. The MOSEC model mimics human ovarian cancer in its abnormal distribution of the actin cytoskeleton and of the adhesion proteins necessary for dissemination of the primary tumor from the surface of the ovary into the peritoneal cavity [55]. Together, these morphologic and genomic analyses provide evidence that the MOSEC disease model as an accurate parallel *in vitro* and *in vivo* system with which to study of human ovarian cancer initiation and progression.

The MOSEC timeline also closely mimics human EOC because transformation arises over relatively a long period of time and requires repeated long-term growth and passaging. Continuous culture of the MOSEC is intended to mimic the repeated growth and repair cycles theorized to contribute to ovarian cancer causation in humans [54-55]. This model is similar to human EOC because significant variability between and within MOSEC cell lines is well established [54-55]. Additionally, histopathologic heterogeneity is also seen in tumors formed from late stage injected MOSEC *in vivo* [54-55]. Until a mechanism for induction of EOC within the mouse ovary *in vivo* is established, the MOSEC model remains a highly relevant syngeneic animal model for hypothesis testing in transformation and tumorigenesis of the ovarian surface epithelium.

Taken together, MOSEC model is particularly well suited for modeling cellular transformation in ovarian cancer because it offers the following advantages: 1) A long timeline (12-18 months) for the analysis of neoplastic transformation, 2) spontaneous transformation without viral or genetic alteration, 3) the generation of late stage cells with transcriptomic and behavioral similarity to the human disease, 4) tumor establishment in immune competent syngeneic mice. Additionally, the MOSEC model is superior to previously established immortalized human ovarian epithelial cell lines as it allows for continuous tracking of the transformation process across the gamut of clearly defined stages without viral alteration or carcinogenic induction of DNA damage. Additionally, this model allows for *in vivo* biomarker detection of tumors following transformed MOSEC xenografting and subsequently for efficacy studies of both

chemotherapeutic treatment regimens and chemo preventive strategies in syngeneic immune competent mice.

1.4 Genomic Heterogeneity in Ovarian Cancer

Like the vast majority of solid tumors, human ovarian carcinomas have near-diploid to highly aneuploid karyotypes and many contain complex combinations of structural and numeric chromosomal aberrations that parallel tumor grade [62]. Additional cytogenetic analyses of five human ovarian surface epithelial cell lines immortalized by HPV16-E6E7 viral oncogenes revealed a high number of chromosomal imbalances including +1q, +3q, +8q, and +20q, -4q, -5q, -8p, -17p, -18q, and -22q, only one of which was universally present among the lines (+19q (5/5 lines) [63]. Additional reported non-clonal chromosome imbalances with variable prevalence marked heterogeneity among the cell lines and included -13q (4/5 lines), +5q, +20q -22q (3/5 lines each), +1q +11q, -2p, -4q, -8p, -10p and -11q (2/5 lines each). This level of variability was documented despite the selective omission of chromosomal imbalances present in only one line. Most important is the realization that each transformed cell line was characterized by a unique set of imbalances, and that only 1/12 (+19q) of the imbalances was universally detected across all transformed lines [64]. A second, independent study found different karyotype alterations when forming a second set of transformed ovarian cancer cell lines [65]. This result demonstrates diversity in the karyotypic alterations is permissive to the tumorigenic phenotype of human OSE, and supports the concept that multiple combinations of genomic alterations can result phenotypically similar cellular transformations.

Heterogeneity among tumors is commonly reported, and ranges from a high degree of structural aberrations to only low frequency of relatively simple rearrangements [66-70]. This is significant to cancer transformation as karyotype heterogeneity can have many gene based and non-gene causes including dysregulation of cell cycle checkpoints [71-75], the presence of structural aberrations caused by errors in homologous and nonhomologous end-joining of double-stranded DNA, and telomere shortening [76-78]. Independent of the mechanism, the

variability that arises from karyotype level change likely provides the basis for selection and evolution of cancer cell populations. Previously, intra-tumor "heterogeneity" has been defined as metaphase-to-metaphase variations among the 50 quantified for each transitional stage, but the complexity of such variations within a single tumor or cell line has eluded quantification [79]. Recently, a method developed by Castro et. Al. [80] showed how inter-tumor diversity among patient samples could be quantified for non-normally distributed data sets by applying the Shannon Index to quantify the heterogeneity of chromosome count data for various tumor types [80]. As a measure of genomic instability, we have adapted this method and applied the Shannon Index to measure the diversity contributed by each whole and aberrant chromosome to the cell population. The indices for each chromosome are summed to generate a summed index for the karyotype of a given cell.

Chromosome instability and aneuploidy result in large scale DNA copy number variations by whole chromosome that can significantly influence the transcription profiles of cancer cells. Several investigations into cancer genomes suggest that nearly all cellular genes may be expressed in proportion to the dosage of the corresponding chromosomes [81-83]. These studies support the notion that cancer karyotypes have significant impact on gene expression and therefore for cell phenotype. Oppositely, in human disease of inherited autosomal trisomies, matching of whole chromosome copy number proportional to the cellular transcriptome does not usually occur or is cell type specific [84-85]. Understanding how this mechanism of gene regulation may function in cancer cell macro evolution and selection by genome system would further define the chromosome-transcriptome relationship of the cancer cell genome. The inherent property of spontaneous transformation in the MOSEC model is important as it eliminates the artificial induction of immortalization by the use of viral or other methods required to transform human cells [63]. Although these methods have become commonplace tools for molecular cell biologists, the generation and usage of genetically altered clones to study cellular transformation presents several significant challenges that warrant

further discussion. The first challenge lies in the determination of whether the change in cell phenotype should be attributed to the intended gene-level alteration in transcription, translation, and protein synthesis of the desired gene product, or, alternatively, reflects the inevitable disruption of the host genome sequence, structure, and spatial relationship of all other genes on that chromosome. Secondly, the selection pressure needed to isolate stably expressing cells creates an "evolutionary bottleneck" in which individual cellular clones are picked and are individually subcultured *in vitro*. The use of virally transformed cells as a model for this disease does not align well with our current understanding of human ovarian cancer formation *in situ* as there is no known viral link to this disease. Depending on the site of insertion, it is also possible that disruption of the host genome may increase genomic instability, subsequently increasing the chance for transformation by this mechanism [86].

High-degrees of karyotypic heterogeneity have consistently marked genomic instability in OECs and cancer cell lines from both humans and mice [87]. Currently, neither the timing nor the extent of this genomic instability has been characterized. Genetic models for MOSEC tumor induction may not mimic spontaneous disease and have limited ability to form epithelial ovarian cancer *in vivo*. Additionally, previous attempts at long-term culture of human ovarian epithelial cells have failed to spontaneously generate tumorigenic cells, and tumorigenesis of transformed human lines can only be tested in immune compromised mice. Therefore, the MOSEC model is well suited in this regard to investigate several aspects of genomic instability as it relates to the key transformative stages of transformation and tumorigenesis.

1.5 Specific Aims

Using the MOSEC model of spontaneous transformation and tumorigenesis in syngeneic C57BL6 mice together with a longitudinal study design, genome-wide multivariate analysis was facilitated at several key transformative stages. With the understanding that cancer cell populations have the properties of a complex adaptive system, our experimental design

accounted for the possibility that these cell populations may work as a dynamic network comprised of many parallel agents that act in a non-linear fashion. Therefore, the following study emphasizes a holistic approach to experimental design that encompasses the possibility of stochastic, non-linear change. To this end, data collection and analysis was performed on cell populations such that the degree of karyotypic variability from cell to cell is characterized at each transitional stage. Using these means under the genome centered paradigm, the relationships among gene expression profiles, cellular phenotype cellular behavior and karyotype heterogeneity can be elucidated. The current study thereby provides evidence supporting an alternate hypothesis of cancer formation that emphasizes genome level variability and cellular selection over the current widely held gene-based theory for cancer causation and stepwise clonal evolutionary theory of tumorigenesis.

These specific aims are designed to provide insight into current major challenges in the diagnosis and treatment of epithelial ovarian cancer which stem from the paucity of early stage samples, the heterogeneity of the disease, and the ability of OSEC to differentiate to a number of heterogeneous subtypes: 1) characterization of genomic instability and the onset of macro-evolutionary change from the earliest stages of spontaneous transformation from the diploid genome 2) definition of the relationship between karyotype heterogeneity and the cell phenotype of each key transformative stage 3) comparing phenotypes and karyotype heterogeneity between spontaneously transformed cells and genetically induced models of cellular transformation 4) quantifying the differences in karyotype profiles between late-stage injected cells and the tumors which evolved *in vivo* 5) identification and profiling of differentially expressed genes between stages and between tumorigenic and non-tumorigenic cell lines 6) defining the role of whole chromosome copy number change in gene regulation by evaluating the genome-transcriptome relationship at each key transformative stage.

2. METHODS

2.1 Experimental Overview

To determine the pattern of karyotype variability and its relationship to spontaneous cellular transformation over time, mouse ovarian surface epithelial cells (MOSEC) were grown continuously in culture from primary cells to their eventual transformation to the malignant phenotype. Population karyotype, senescence, and morphologic data were collected at the earliest transformative stages. Subsequent key transformative stages were determined by cell phenotype and by repeat testing of the MOSEC *in vitro* using traditional and organotypic models. Genome level change was measured at several levels including spectral karyotype analysis for MOSEC populations, array comparative genomic hybridization (aCGH) for DNA copy number, and RNA transcript abundance for transcriptional change (Figure 1). To confirm the tumorigenicity of MOSEC, late stage day 528 cells were compared to day 245 cells *in vivo* by allograft followed by histopahologic and karyotype evaluation of harvested tumors

All analyses were performed real-time with the exception of the microarray analyses. The selection of key transitional stages was based on previously established patterns of *in vitro* phenotype change in this model [54-55]. These included ranges of passage number, size, density, growth rate, anchorage-independent growth efficiency in soft agar, invasiveness in matrigel culture, and depth of invasion on collagen rafts.

2.2 Cell Culture

2.2.1 Establishment of Spontaneously Transforming Primary Cells

Primary MOSEC were generated from a single ovary from a 6-week-old C57BL6 female mouse. To avoid the possibility that heterogeneity was generated from the establishment of cell lines from multiple mice, primary cultures were attempted using cells from the single ovary, but viability was low using this method (1:20). Cultures were discarded if they failed to show viability by day 30. The ovaries were resected, and residual remnants of the oviducts and bursa were

removed. Each ovary was individually incubated for 20 minutes in a 20 mm dish with 50 µl of 0.25% trypsin and was rolled gently with a toothpick to help dislodge the surface epithelial cells. For each ovary, cells were resuspended in MOSEC growth medium (low glucose DMEM with 4% fetal bovine serum, 100 µg/ml each of penicillin and streptomycin, 5 µg/ml insulin, 5 µg/ml transferrin, and 5 ng/ml sodium selenite (Invitrogen, Carlsbad, CA) were plated in one well each of standard 6-well tissue culture plates (Becton-Dickinson, Oakville, ON, Canada).

2.2.2 Conditional Inactivation of Brca1

To determine cell phenotype and the extent of karyotype heterogeneity in genetically induced versus spontaneously arising transformation models, MOSEC cells from mice bearing loxP sites in introns 4 and 13 of the *Brca1* gene (*Brca1*$^{loxP/loxP}$ [FVB;129-*Brca1*^{tm2Brn}]) were provided as a generous gift from investigators Clark-Knowles and Vanderhyden [48]. Following the initial establishment of the primary cells MOSEC as above (30 days), targeted inactivation of *Brca1* was achieved via adenoviral delivery of Cre recombinase to MOSEC using recombinant adenoviruses Ad5CMVCre (AdCre), (Vector Development Laboratory, Houston, TX, USA). Mitotic cells were passaged only enough times to assure mitotic activity sufficient for SKY analysis of metaphase spreads (20 days post AdCre administration).

2.2.3 Routine Cell Culture Conditions

Cells were passaged in MOSEC medium and were split at varying ratios, depending on their growth rate (1:2 - 1:12 ratios). For growth rate analyses, cells were seeded at densities of 1 x 10^4 and 5 x 10^4 cells, and subconfluent cell counts were determined at different times postseeding. Cell doubling times were estimated from the formula: doubling time = $(T)\ln 2/\ln(X_e/X_b)$, where X_e is the cell number determined at the endpoint, X_b is the cell number at the beginning time point, and T is the total elapsed time (in hours). Transitional states of carcinogenesis were captured by preparing frozen cell stocks every three to five passages. Stocks were stored in liquid nitrogen and preservation of cell phenotype was verified and population karyotype analyses were shown to be >95% similar in re-seeded cells from frozen

stocks. Freezing medium contained 50% fetal bovine serum, 40% low glucose DMEM, and 10% DMSO. All analyses were performed on transitional stages real time using never-frozen cells with the exception of the collection of mRNA for analysis of transcript abundance.

2.3 In vitro invasion, proliferation, and migration assays

2.3.1 Soft Agar Proliferation Assay

MOSEC (1.5×10^4) were suspended in 1.0 ml of 0.5% Bactoagar (Difco Laboratories, Detroit, MI) in the MOSEC growth medium at 42°C, layered over 1 ml of 0.8% Bactoagar in DMEM in 6-well dishes, and cultured in 5% CO_2 in a humidified chamber. After 15 days of growth monitoring, colonies were stained with cresyl violet, sized using a 10x calibrated eyepiece and were counted if they exceeded >10 μm in diameter.

2.3.2 Migration Assay

MOSEC were trypsinized and resuspended at 10^5 cells per 0.3 ml serum-free media and migratory capacity was tested using 6-well invasion chambers with 8 μl Matrigel coated trans-well filter inserts (BD Biosciences) and 8 μl uncoated controls. Cell suspension was added to the upper chambers and the lower chamber contained 0.8 ml serum-containing media. After 4 h in a humidified 37°C incubator, the inserts were removed, the upper surface of the filters was cleaned thoroughly with cotton swabs, and the lower surface was fixed with 4% paraformaldehyde for 10 min and stained with 0.5% crystal violet in 20% methanol for 5 min. Membranes were mounted on glass slides and images were captured using a light microscope with 20x objective and were counted in a blinded fashion. Values reported are normalized to controls as follows: Migration = migratory cells matrigel/ migratory cells control.

2.3.3 Three dimensional growth in Matrigel

MOSEC (2×10^4 cells/well) were embedded directly in Matrigel (BD Biosciences, Franklin Lakes, NJ) diluted 1:1 with the MOSEC growth medium. Following solidification of the Matrigel, cultures were incubated at 37°C and monitored for up to 15 days. Phase contrast images of

colonies grown in Matrigel were captured with a Nikon Coolpix 990 digital camera attached to a Nikon Diaphot microscope using 20x objectives. Invasiveness was scored by branching morphology as 1) number of branches/ cell and 2) branching cells/ total cells. Two hundred cells were counted per assay.

2.3.4 Organotypic Collagen Raft Cultures

Organotypic collagen raft cultures were prepared essentially as described by Gregoire et al. [63], with minor modifications. Briefly, collagen plugs (2.5 ml of 1.63 mg/ml rat tail collagen; BD Biosciences) were prepared in trans-well filter inserts (BD Biosciences) and were equilibrated with the MOSEC growth medium prior to seeding the cells. Cells (5×10^5) were seeded onto collagen plugs and allowed to grow to confluency at 37°C in a humidified incubator with 5% CO_2 for 1-2 days. After removing the medium from the top chamber, raft cultures were maintained with sufficient medium for cells to grow at the air/ liquid interface for 14-21 days. Cell growth was monitored and cells were fed by changing the medium every 2 to 3 days from the bottom chamber. Rafts and their associated cultures were carefully transferred into 10% neutral buffered zinc formalin (Thermo *Fisher* Scientific Inc.) for processing by routine histopathologic analysis.

2.4 Senescence Associated Beta Galactosidase Activity

Senescence Associated Beta Galactosidase Activity (SA-βgal) was measured from MOSEC cells grown on glass slides under standard culture conditions. Slides were removed from culture dishes, rinsed twice with phosphate buffered saline (PBS), and were fixed in 2% formaldehyde, 0.2% glutaraldehyde in PBS at room temperature for 10 minutes. Cells were rinsed twice again with PBS, covered with fresh staining solution (400 nM citric acid, sodium phosphate pH 6.0 1.5 mM NaCl, 200 mM $MgCl_2$, 50 nM potassium ferrocyanide, 50 nM potassium fericyanide, 20 mg/ ml X-gal in DMF), and incubated without light for 12-24 hours at 37°C. SA-βgal images were captured with an Olympus BX41 microscope at 40x with DP72 camera. Cells were counted with

the investigator was blinded to their identity using NIH Image (U.S. National Institutes of Health and available at http://rsb.info.nih.gov/nih-image/).

2.5. Population Karyotype Analysis

2.5.1 Spectral Karyotype Analysis (SKY)

Painting and detection of metaphase cells from key transitional stages was performed following overnight treatment with colcemid and collection using mitotic shake-off. Chromosomes were prepared using standard hypotonic treatment, fixation, and air drying [88]. After pepsin treatment and fixation with formaldehyde followed by dehydration, the slides were denatured in 70% formamide and 2x SSC and hybridized with denatured mouse painting probes (Sky Paint, Applied Spectral Imaging, Vista, CA) for over 48 hours at 37°C. Following two additional washes in 2x SSC, the slides were mounted in an anti-fade solution (Vectashield; Vector Laboratories, Burlingame, CA) with 4,6-diamidino-2-phenylindole (DAPI). Image acquisition was performed using an SD200 Spectracube (Applied Spectral Imaging, Carlsbad, CA) mounted on a Leica DMRXA microscope (Leica, Wetzlar, Germany). Applied Spectral Imaging software (Spectral Imaging and SkyView) was used for image acquisition and analysis of cell populations at each key transitional stage. Typically, multi-color SKY karyotype descriptions are based on only five to ten mitotic figures with the emphasis on tracing clonality of the cell line [79]. The current data set was generated by capture of 50 mitotic figures with clearly defined boundaries and minimal chromosome overlap [89] for each population of interest. Chromosomes were organized into a karyotype table according to their color and size using SKY viewer software [90]. Results were verified by a second investigator who was blinded to the sample identity and concordance was ~97% for whole and aberrant chromosome identification.

Results were tabulated for all whole chromosome copy number change and derivative chromosomes. Chromosomal aberrations were considered clonal if present in >20% of the metaphases, (> 30% of metaphases for chromosome loss), according to established

International System for Chromosome Nomenclature (ISCN) conventions. Aberrations found in <20% of metaphases were termed, "nonclonal."

2.5.2 Quantifying karyotype heterogeneity within a key transformative stage

To obtain a quantitative measure for karyotypic diversity of cells at each key transitional stage, an adaptation of the Shannon Index was calculated using methods of adapted systematic cytogenetic analysis adapted from Castro et. al. [80]. This method has been widely used by ecologists to calculate population diversity and has been validated for tumor heterogeneity by total chromosome count by Castro et. al. [80]. Population diversity at each key transitional stage was quantified by summation of Shannon indices (*H'*) [91] for each whole (1-19) and aberrant (der) chromosome according to the following equations:

equation 1 $$H = \sum_{i=1}^{S} \left(p_i \ln \frac{1}{p_i} \right)$$

equation 2 $$H_{max} = \sum_{i=1}^{S} \left(p_i \ln \frac{1}{p_i} \right) = \ln S$$

equation 3 $$H' = H / H_{max}$$

Where n_i is the number of cells with chromosome count *i* (the abundance of chromosome count *i*), S the number if unique chromosome counts in the population, p_i the relative abundance of each chromosome count (calculated as the proportion of cells with that count). It can be shown that for any population, there is a maximum possible *H* based on the number of different chromosome counts, where H_{max} = ln *S*. This occurs when all counts for a given chromosome

are equally distributed within the population. The X chromosome was omitted from all Shannon Index calculations as Barr body inactivation is likely dysregulated and its impact on X chromosome variability is an unknown potentially confounding variable for the current data and to future comparisons with male genomes.

2.6 Whole Genome Analysis of mRNA Transcript Abundance

2.6.1 Sample Collection, Labeling, Hybridization, and Scanning

Total RNA was isolated from biological triplicates of cells grown to 75% confluency in T-75 flasks at key transformative stages days 170, 245, 450, and 528. RNA extraction was performed from with Trizol reagent (Invitrogen, Carlsbad, CA) followed by clean-up using an RNeasy mini kit (Qiagen, Valencia, CA) in accordance with the manufacturers' instructions. First strand cDNA synthesis, second strand synthesis, clean-up, and *in vitro* amplification of biotin-labeled cRNA fragments were performed using an original starting amount of 1µg RNA for each biological replicate. The hybridization cocktail containing 10µg fragmented RNA and probe array controls was bound to Illumina MouseRef-8 v2.0 expression bead chips. By design, one array on each chip served as a technical replicate to account for between-chip variation and was determined to have a correlation coefficient of >0.996. Beadchips were scanned on the Illumina BeadArray Reader confocal scanner for 90 min and initial quality assessments were performed according to the manufacturer's instructions. Procedures were carried out by personnel in the Applied Genomics Technology Center (AGTC, Wayne State University).

2.6.2 Preprocessing and Normalization

High density microarray image files were interpreted and quality was assessed of both the internal controls and the between-chip technical replicate. Microarray data pre-processing including background subtraction, cubic-splines based normalization, and quantification of differential expression compared to negative controls were performed using Illumina Genome

Studio Software. Probesets were excluded from the outputs of time course and differential expression analyses if they failed to meet the following criteria for at least one of the arrays: significant differential expression above background p >0.1, and signal intensity greater than 3x background (signal intensity >420).

2.6.3 Time Course Analysis of Differential Gene Expression

To identify differentially expressed genes throughout spontaneous transformation, time-course microarray data analysis was performed using Extraction of Differential Gene Expression (EDGE) software [92]. EDGE utilizes a series of cubic splines to model within class temporal gene expression. A statistic is calculated for each gene that quantifies the goodness of fit between the null hypothesis (H_o is defined as constant expression over time, where the sum of squares is minimized among all possible flat lines) and the alternative hypothesis (a curve that minimizes the sum of squares among a general class of cubic-spline curves). A significance cut-off was applied to the statistics by using a false discovery rate criterion where the null distribution of the statistic is based on genes with no differential expression. These analyses were implemented using the Optimal Discovery Procedure for simultaneous significance testing of differential expression [93] using C++ extensions with the open source R version 2.7.1 Statistical Environment (www.r-project.org) [94]. Differential gene expression was considered significant where $q < 0.01$.

2.6.4 Profile Fitting for Differentially Expressed Genes

Time-course based profile analysis of the differentially expressed genes was performed using Short Time Expression Miner (STEM) [95]. Genes were assigned to model temporal expression profiles based on maximization of the correlation coefficient between the model profile and the actual gene expression. The number of genes expected to be assigned to a profile was estimated by random permutation of the original time point values and assignment of

genes most closely matching model profiles. The process was repeated for a large number of permutations. The statistical significance of the number of assigned versus expected genes for each profile was computed, and model profiles were grouped to form clusters of profiles. Profiles were assigned based on unfiltered relative expression levels from the previously generated list of differentially expressed genes using EDGE. Settings were kept at the original defaults with the exception that maximums were set to 5 unit change between time points, 50 theoretical profiles, and correlation of 0.9 among profiles and among clusters. Genes were included in a given profile at a significance level of p <0.05. The Benjamini and Hochberg method [96] was used for multiple testing correction of the false discovery rate. The biological, molecular, and cellular functions for each cluster of genes were interpreted using Gene Ontologies [97] as indicated below.

2.6.5 Biological Significance of Microarray Data

The biological significance of sets of genes assigned to lists or expression profiles of interest was assessed using Gene Ontologies (GO) enrichment analysis. Differentially expressed genes were classified by their enrichment for corresponding GO categories, and the observed number of genes in each of these GO categories was recorded. Genes represented on the Illumina MouseRef-8 v2.0 expression bead chip served as the reference gene list. The expected number of genes in each GO category corresponds to the number of genes falling into that GO category in the reference gene list. A given GO category was considered enriched when the observed number of genes in that category was significantly greater than the expected number. A Bonferroni corrected p-value, p <0.05 was used for lists generated by SAM analyses. For STEM data, annotations for biological processes, molecular function and cellular componentry were determined at a depth of three annotation levels for Gene Ontologies, requiring a minimum of five genes per annotation and annotation significance was determined where p <0.05 after correction for multiple hypothesis testing by randomization.

2.7 Validation of Spectral Karyotype Data by Array Comparative Genomic Hybridization (aCGH)

2.7.1 DNA Isolation, Labeling, and Hybridization

DNA copy number analyses for interphase and mitotic-enriched MOSEC at key transitional stages (days 245, 450 and 528) were measured on Agilent Mouse 4 x 44K arrays. Interphase cells were cultured using standard MOSEC conditions, mitotic enriched cells were harvested after 8 hours treatment with Colcemid and were shaken off. Genomic DNA (gDNA) was isolated using the DNeasy Tissue kit (Qiagen, Inc., Valencia, CA) following the supplier's protocol. DNA quantity and quality were assessed by UV/Vis spectrophotometry using the ND-1000 Spectrophotometer (NanoDrop Technologies, Rockland, DE). CGH labeling was performed in accordance with the chip manufacturer's instructions (Protocol v3.1, 2009, Agilent Technologies, Santa Clara, CA). Briefly, gDNA (3 µg/ sample) from the reference (Promega Mouse Genomic (XY/XX pooled) DNA, Madison, WI) and from day 245, 450 and 528 cells were heat fragmented and denatured at 95° C for 10 minutes. The DNAs were labeled with ULS- Cy3 (reference) or ULS- Cy5 (test) using the Agilent Genomic DNA Labeling Kit Plus (Agilent Technologies). Labeled DNA products were purified with Microcon YM-30 filtration devices (Millipore, Inc., Bedford, MA). For hybridization, the appropriate Cy3 and Cy5-labeled DNA sample pairs were combined and then mixed with mouse Cot-1 DNA (Invitrogen Carlsbad, CA), Agilent Blocking Agent, and Agilent Hybridization Buffer. Prior to hybridization, the samples were heated at 95°C for 3 minutes and then incubated for 30 minutes at 37°C. The labeled target solution was then hybridized to the arrays using Agilent SureHyb chambers. The hybridization chambers were placed in a 65°C rotisserie oven set at 15 rotations per minute for 40 hours. Post-hybridization, the microarrays were washed and dried according to the manufacturer's protocol and were scanned immediately using an Agilent microarray scanner.

2.7.2 Normalization and Pre-processing

Microarray data for the individual features including $\log_2$ ratios and quality metrics (median signal intensity, background noise, signal-to-noise ratio, reproducibility of replicate probes, and probe-to-probe log ratio noise were extracted from the scan image using Agilent Feature Extraction software (v 9.1.3.2).

As cancer genomes are universally known for aneuploidy and variable chromosome copy number imbalances, the prerequisites for conventional normalization methods (that nearly all of the DNA copy numbers are homogeneous and nearly equal to the diploid genome) were not met. As chromosome copy number correlated with intensity, a novel strategy for array-CGH data was implemented. Data normalization using a custom adaptation of PopLowess [98] with the open source R Statistical Environment (www.r-project.org) was performed using libraries from the Bioconductor Project (www.bioconductor.org). This normalization method accounts for the presence of variable whole chromosome copy number change prior to segmentation and calling of small amplified or deleted regions. A maximum of five subpopulations with a merge cluster criteria for the subpopulations of 0.3 was utilized. The k-means clustering algorithm was adapted to incorporate the weighted means for each cluster. Following zero-centering to the largest subpopulation, the median value for the cluster with the lowest chromosome copy number was added as a constant to the M value for each probe. With this correction, data were centered with the assumption that the cluster with the lowest median M was assumed to have a chromosome copy number count of two.

2.7.3 Analysis of DNA Copy Number Change

Whole chromosome copy number change for validation of SKY analyses and detection of genomic gains and losses was determined using PopLowess to rank chromosomes by cluster median for comparison with median chromosome counts from SKY analyses. Copy number for sub-chromosomal regions was assessed after normalization using PopLowess and was determined using Agilent's CGH Analytics v3.5 software with statistical algorithm ADM-1, sensitivity threshold 6.0, and a moving average window of 1 Mb. Copy number change for

chromosomal sub-regions was called by locus if a positive call was made by ADM-1 testing and ≥10 consecutive probes were seen with ≥1.5-fold change.

2.8 In vivo Tumorigenesis

2.8.1 Determination of Tumorigenicity in vivo

To determine the relative tumorigenicity of day 245 versus day 528 cells *in vivo*, six-week-old female C57BL6 mice were purchased from syngeneic immune competent C57BL6 mice and in nod-SCID-gamma (NSG) immune-compromised mice (Charles River Laboratories, Wilmington, MA) and from JAX® Mice strain NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ (Jackson Laboratories Bar Harbor, ME).

Following five days of acclimatization, each mouse received a total of three injections. Intraperitoneal injections consisted of $5x10^6$ MOSEC in 0.2 ml DMEM with 0.25% Matrigel, and bilateral subcutaneous injections of $1x10^5$ in 0.1 ml DMEM with 25% matrigel were administered bilaterally on the back. Cell types were divided as follows so that each mouse received only one cell type (n=7 day 245, n=7 day 528, n=2 positive control (ID8) cells from one of 10 previously established tumorigenic MOSEC late stage clones from the discoverer of this model system [54], n=2 negative control acellular Matrigel mixed 1:1 with DMEM). Mice were weighed and tumor size was approximated by diameter using calipers once per week. Mice were sacrificed at day 70 when tumor locations, weights, and sizes were documented. All animal investigations were conducted in accordance within the guidelines and with approval from the Wayne State University Institutional Animal Care and Use Committee.

2.8.2 Tumor Histopathology

The location and amount of intraperitoneal tumor formation was scored where 1 point was given for positivity in each of the following: pancreas, liver, peritoneal fat, reproductive organs, lymph node, spleen, small intestine, large intestine, diaphragm, hemorrhagic ascites (max=10). Subcutaneous findings were scored based on invasiveness and severity where 1

point was given for positivity: dermis, subcutaneous fat, muscle, diffuse, necrotic (max= 5 each side, 10 total). Tumor implants and the surrounding tissues were collected from multiple sites, fixed overnight in 4% paraformaldehyde and embedded in paraffin for routine hematoxylin and eosin staining (Beaumont Hospital Tissue Biobank, Royal Oak, MI). After de-identifying the tumor and tissue specimens, each was evaluated by a clinical pathologist (RR).

2.9 Statistical Analyses

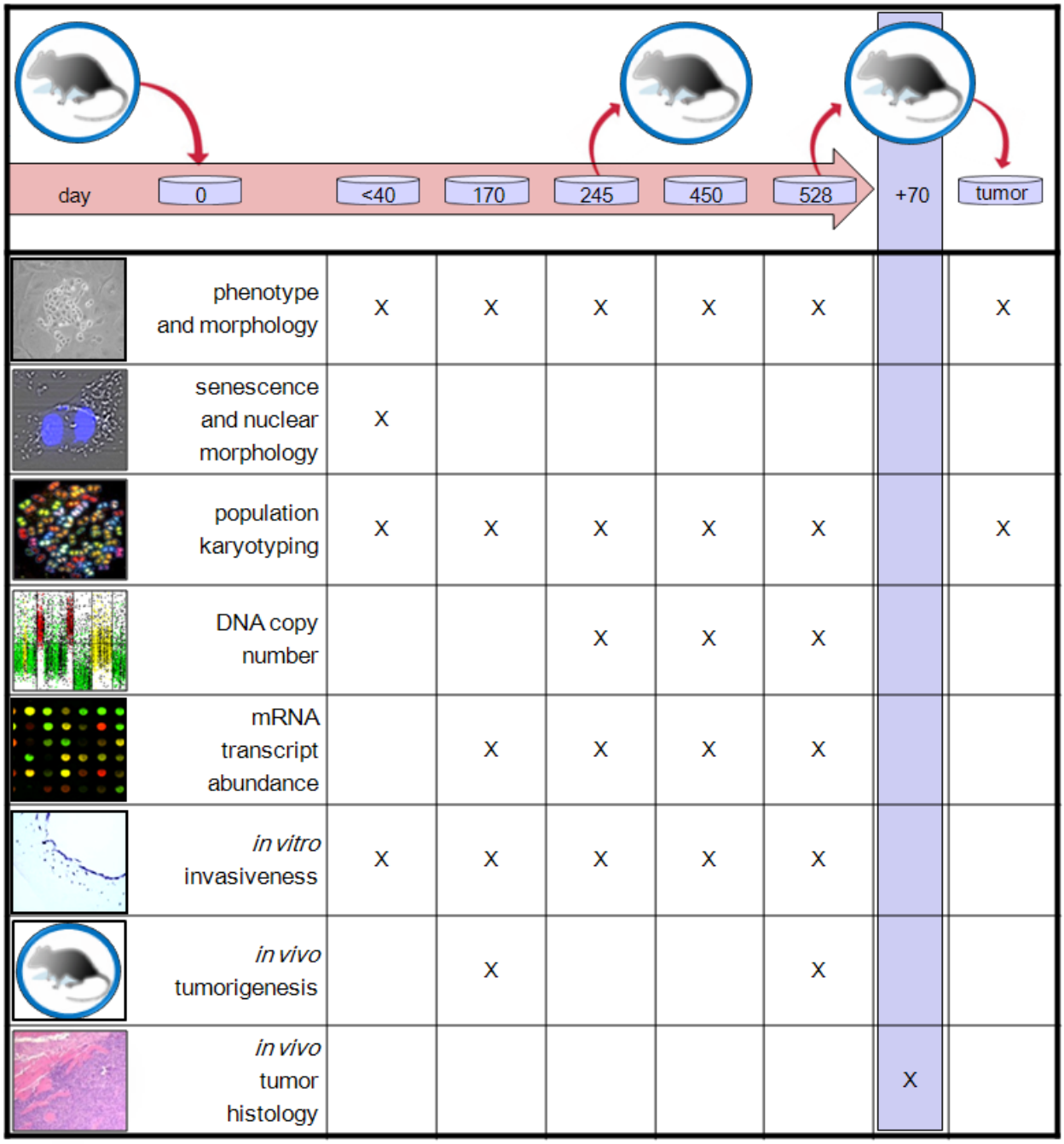
Counts of morphological features in the ovarian surface epithelial are expressed as the mean ± SEM (standard error of the mean) of the number of morphological features per section in five non-consecutive ovarian sections for n ovaries, where n is the number of ovaries examined. *In vitro* cell counts are expressed as the mean ± SEM of three independent experiments performed in triplicate. The predictive value of chromosome ID, time and Shannon Index was determined using general linear mixed model. For each sub-population, chromosome count was compared using a paired t-test. The probability of significant differences upon comparison of only two groups was determined by Student's *t* test. When multiple groups were analyzed, statistical comparisons were made by analysis of variance (ANOVA). Bonferroni's post hoc correction for multiple comparisons was applied in the determination of significant differences between specific groups when whole group differences were detected by ANOVA. Alpha was set at a value of 0.05 and all *p*-values are two-sided. The proportion of variance in SA-βgal accounted for by time was estimated by calculation of R-squared. The link between senescence by SA-βgal and nuclear abnormality was tested using the chi-squared statistic. Statistical analyses were performed using NCSS (Kaysville, Utah).

3. RESULTS

3.1 Timeline and Experimental Overview

3.1.1 Whole Genome Analysis of Key Stages of Transformation and Tumorigenesis

Mouse ovarian surface epithelial cells (MOSEC) were seeded from a single ovary from a healthy 6-week old C57BL6 mouse and were repeatedly passaged until reaching a tumorigenic phenotype similar to positive control cells (ID8) *in vitro*. Tumorigenicity testing of day 245 versus day 528 MOSEC *in vivo* was performed by allograft injection of select transformative stages and tracking of tumor formation for 70 days *in vivo*. Tumorigenicity *in vivo* correlated well with phenotype analyses *in vitro* and confirmed the ability for this cell type to transform spontaneously and without out viral transfection or carcinogenic insult [54-55]. Using this model, data was acquired longitudinally from five key transformative stages, beginning with primary cells in culture until malignant characteristics were eventually acquired at day 528. Cell culture was continued until the characteristics of the malignant phenotype were acquired and the MOSEC performed similarly to positive control ID8 MOSEC in several types of *in vitro* and *in vivo* measures of proliferation, invasion, migration, and tumorigenicity (Figure1, Table1). Real-time analyses of MOSEC were performed for all assays except for measuring tumorigenesis *in vivo* in which cells were retrieved from prepared frozen stocks. This experimental design is unique as it successfully facilitated the longitudinal study of ovarian cancer transformation real-time and allowed for the comparison of cellular phenotype with global genomic profiles at multiple levels including cytogenetic abnormalities, population karyotype analysis, mRNA transcript abundance, and array comparative genomic hybridization (Figure 1).



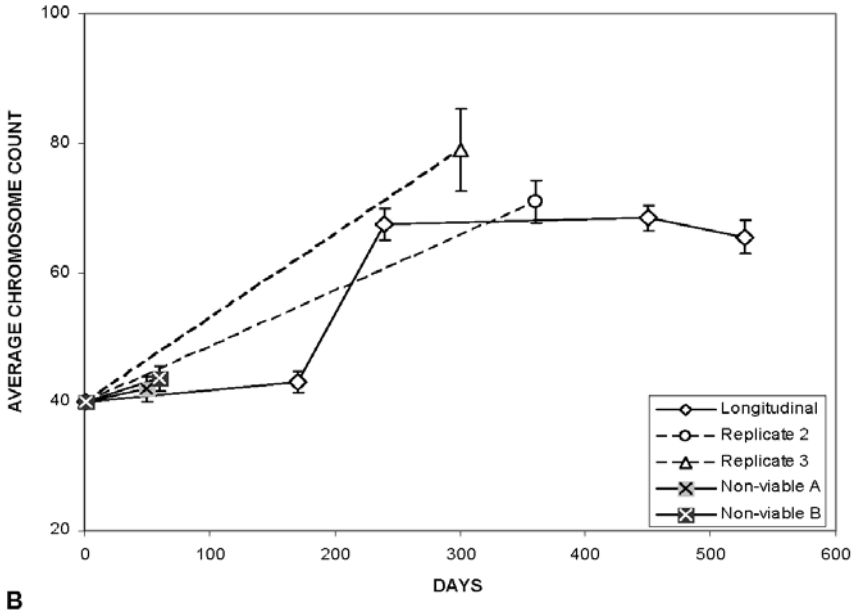
	phenotype and morphology	X	X	X	X	X		X
	senescence and nuclear morphology	X						
	population karyotyping	X	X	X	X	X		X
	DNA copy number			X	X	X		
	mRNA transcript abundance		X	X	X	X		
	in vitro invasiveness	X	X	X	X	X		
	in vivo tumorigenesis		X			X		
	in vivo tumor histology						X	

Figure 1. Experimental overview and project timeline throughout the spontaneous transformation and tumorigenesis of mouse ovarian surface epithelial cells. Cells were seeded in primary culture at day zero from the surface of the mouse ovary and were passed continuously until day 528. Various cell phenotypic and genome wide molecular measures of chromosome copy number, DNA copy number, and mRNA transcript abundance were performed at key transitional stages (X). Tumorigenesis in vivo was tested by subcutaneous and intraperitoneal injection of re-seeded cell stocks from day 245 and day 528 cells into 6-week old C57/BL6 mice. Tumor histopathology of injected day 245 vs. 528 cells was evaluated and harvested tumors from day 528 cells were reseeded in culture for 2-3 days for phenotype and population karyotype analyses.

3.1.2 Establishment and Selection of Primary Ovarian Surface Epithelial Cells

Primary cells were difficult to establish in culture and the survival rate was approximately 5%. The high prevalence of cellular senescence and the low frequency of mitotic events in these cells likely accounts for their low viability and the difficulty incurred in establishing these lines in culture from small numbers of seeded cells. Five lines were established past day 30, two of these lost viability before day 100 ("non-viable A", "non-viable B") and three of these progressed to transformation ("Longitudinal", "Replicate 1", "Replicate 2"). Whole-genome data was acquired from the "Longitudinal" line including repeated measures of spectral karyotype data, microarray analyses, and tumorigenicity testing and is presented from here forward as MOSEC. Importantly, "Replicate" lines displayed similarly elevated chromosome counts, cellular variability and altered cellular morphology throughout transformation. The frequency of aneuploid cells was not different among lines from the viable and non-viable groups, but each was different from a the normal diploid karyotype at first measure (day 76,82, and 170) where chi-squared <0.01 for each of three lines sampled before day 200, (Figure 2).

A



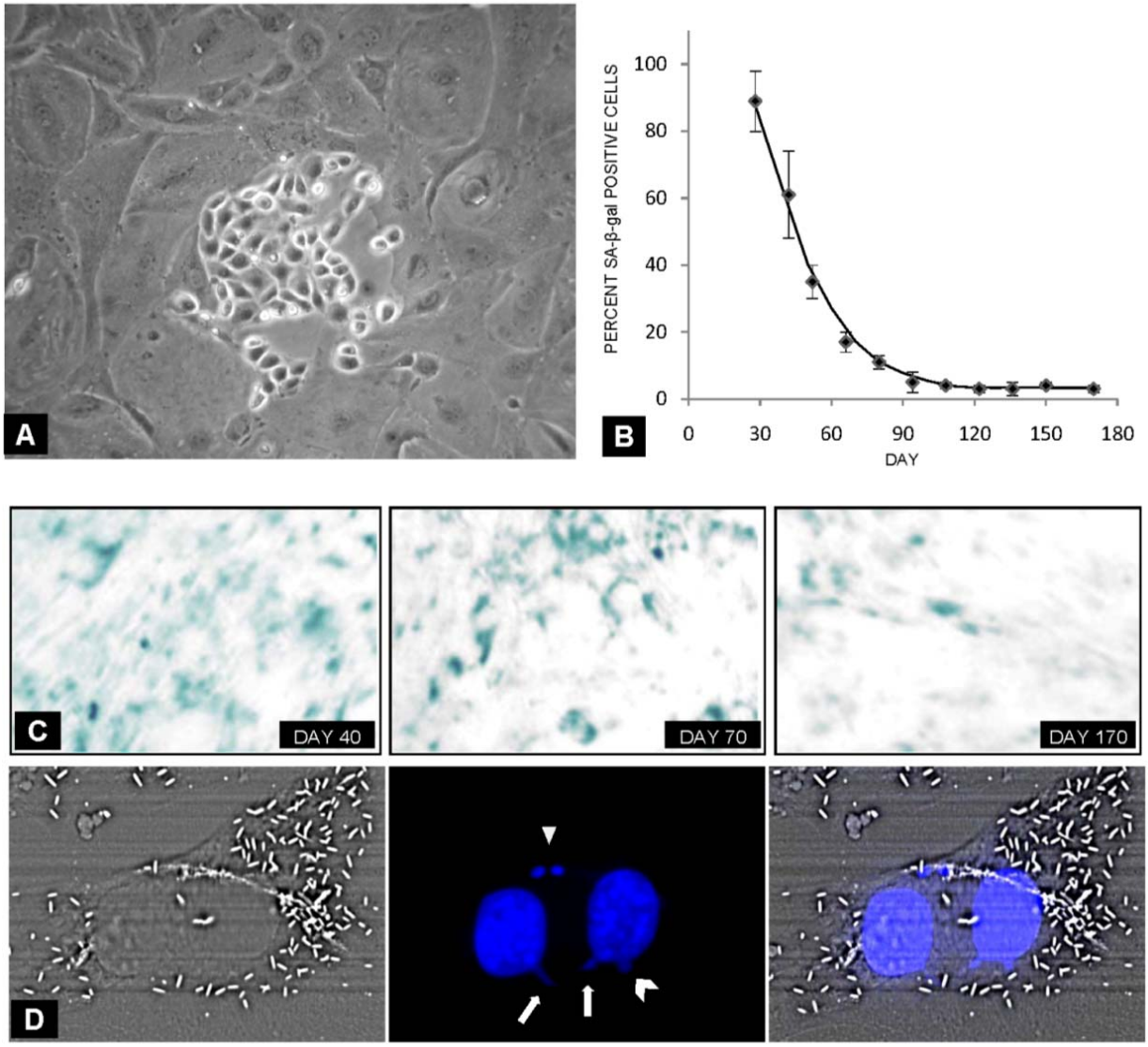
B

	Non-Viable A	Non-Viable B	Longitudinal
aneuploid cells (%)	29	28	31
chi-squared	*9.4*	*7*	*13*
p-value	*<0.01*	*<0.01*	*0.002*

Figure 2. Chromosome count data for replicate and non-viable primary cell lines, each seeded from a single C57BL6 mouse ovary. Chromosome count data taken from metaphase preparations by counting of DAPI stained centromeres. Replicates are shown for three viable lines (Longitudinal, Replicate 2, Replicate 3) and for two non-viable lines (Non-viable A, Non-viable B). Viability of primary cultures reaching day 30 was approximately 5% and frequency of aneuploid cells was different from a normal diploid karyotype at the first time point evaluated for each line. Aneuploidy frequency was not different between non-viable and viable lines. Tabulated data for two non-viable versus longitudinal line aneuploidy frequency at <200 days was significantly different from diploid karyotype, p<0.01, (B).

3.2 Senescence and Nuclear Abnormalities of Early Stage Interphase Cells

The earliest observations of the MOSEC *in vitro* revealed a heterogeneous pattern of cellular morphologies. At day 40, large, ballooned cells with diverse morphotypes and low cell densities typical of senescent cell populations occupied the majority of the surface of the culture dish. Within this population lay pockets of mixed cuboidal and refractile cells. These pockets are more typical of the phenotype of MOSEC at later transformative stages (Figure 3A). Further evidence for the senescence phenotype of these large ballooned cells was generated by the bi-weekly assaying for senescence-associated β-galactosidase (SA β-gal) activity from day 28 to day 170. The percentage of SA β-gal positive cells averaged 94 ± 3%) at day 28 and fell logarithmically with routine passaging over time to reach <5% from day 100 onwards, $R^2 = 0.83$ (Figure 3B, C). Nuclear abnormalities including blebs, micronuclei, and lagging chromosomes were noted with increased frequency at day 28, 40, and 70 in SA β-gal positive versus negative cells (chi-square analysis: *p=0.04, 0.02, and 0.02* respectively, Figure 3D, 3E). From day 28-70, the frequency of abnormal nuclear morphology was unchanged in normal and SA β-gal positive cell populations, despite the largest decline in the prevalence of SA β-gal positive cells.



E

	DAY 28		DAY 40		DAY 70	
SA β -gal	-	+	-	+	-	+
nuclear abnormality	5 ± 1	14 ± 2	5 ± 2	15 ± 3	4 ± 2	14 ± 4
normal	45 ± 3	36 ± 5	45 ± 3	35 ± 4	46 ± 3	36 ± 3
chi-square (*p-value*)	4.2 (*0.04*)		5.1 (*0.02*)		5.5 (*0.02*)	

Figure 3. Morphologic and behavioral characteristics of mouse ovarian surface epithelial cell senescent and mitotic sub-populations. Cellular morphology of day 20 cells included both large flattened balloon-shaped cells (background) and pockets of more typical cobblestone type cells with some refractile cells (A). Ninety-four percent of day 20 cells were senescence associated β-galactosidase (SA β -gal) positive. Senescence fell logarithmically with routine passaging over time reaching <5% from day 100 onwards; loess smoothed best fit line with bandwith of 0.4, $R^2 = 0.83$ (B). SA β -gal activity of cells at day 40, 70, and 170 (C). Images of a SA β -gal positive binucleate cell with micronuclei (arrowhead, D), laging chromatin between nuclei (arrows, D), and a nuclear bleb (chevron, D). Images were captured with light and fluorescence microscopes, where SA β -gal crystals are refractile white, and DAPI, stained nuclei are blue, merged image (D). Nuclear abnormalities occurred with increased frequency in SA β -gal positive versus negative cells at day 28, 40, and 70, data shown as mean ± standard error, persons chi-squared $p < 0.05$ (E).

3.3 Early Stage Abnormalities of Cell Division

The early stage (day<40) MOSEC were slow-growing and cytogenetic preparation for spectral karyotype analysis revealed the presence of several types of nuclear and cytogenetic abnormalities. Typical examples from classes of nuclear and cytogenetic abnormality are depicted from colcemid treated SKY painted cells that provide evidence for early genomic instability and cell cycle checkpoint insufficiency (Figure 4). When late-stage metaphase cells enter mitosis in the presence of colcemid, unpairing of centromeres can occur without spindle formation and is termed late c-metaphase [99]. During cellular division when the centromeres are visible, discondensed chromosomes are seen that comprise individual circular blebs with the aforementioned split centromere morphology (Figure 4A, B). SKY analyses of these cells clearly shows that many of the individual blebs contain DNA from a single chromosome with separated centromeres, whereas other images show coalescence of further discondensed chromosomes with unclear centromere morphology (Figure 4C). Images of nuclear outpouchings and macro-nuclear formations were additionally captured where DNA from at least three different chromosomes is seen to be contained within the rounded outpouchings protruding from the normal nuclear domain (~4% of interphase cells, Figure 4C). Metaphase cells with well condensed chromosomes and clearly defined single centromeres also show chromosomal abnormalities including multiple breaks, fusions, and other structural abnormalities (Figure 4D-F). In total, early signs of genomic instability and cell cycle checkpoint insufficiency including defects in chromosome condensation, segregation, and nuclear morphologic abnormalities such as those depicted occurred in day <40 cells at a frequency of 19 ± 6%.

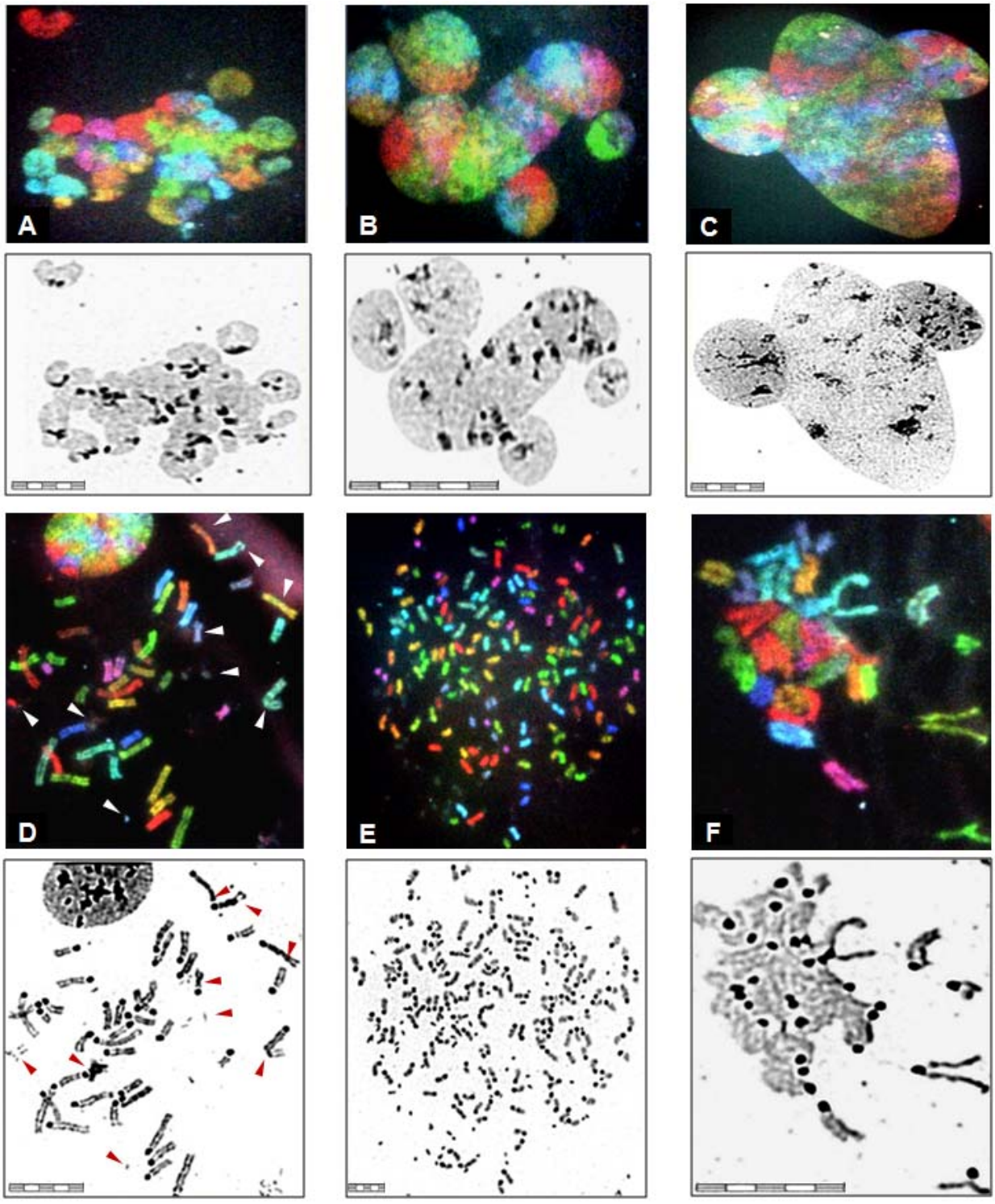


Figure 4. Spectral karyotype analysis of nuclear and cytogenetic abnormalities from primary mouse ovarian surface epithelial cells before day 40. Early SKY analyses (upper, A-F) and counterstained DAPI image (lower, A-F). Nuclear blebbing, where entire chromosomes or chromosome pairs can be seen as individual blebs (A,B), individual giant nuclei with abnormal morphology (C). Chromosomal abnormalities of metaphase cells including breaks, fusions, minute chromosomes and other structural abnormalities (D arrowheads, E). Functional defects in chromosome condensation and morphologic abnormalities (F). Images were captured with a 63x objective lens with post acquisition sizing to 1x, 2x, or 4x is indicated by 30μm scale bar (lower left, A-F).

3.4 Morphologic and Phenotypic Features Defining the Key Transformative Stages in vitro

Several distinct phenotypic changes characterizing the MOSEC key transformative stages were captured in two-dimensional, three-dimensional, and organotypic culture scenarios (Table 1, Figure 5). Day 170 MOSEC divide more slowly than all subsequent stages, maintain a large surface area, low cell density, and slightly elongated morphology compared to day 245 cells. The day 245 cells are characterized by increased cell density compared to day 170 and have acquired the typical cobble-stone appearance of epithelial cells grown *in vitro*. Day 450 and day 528 MOSEC continue to increase cell density until transitioning to a phenotype of tightly packed cells with regions of overlap (Figure 5A). Three-dimensional cell growth in matrigel shows parallel morphologic changes and a phenotype of increasing branching that was significantly greater at each transformative stage (170;2%-528;96%, $p=0.01$). Day 170 cells have limited branching and invasion capabilities as seen by their growth limitation to one plane on the substrate surface. In comparison, the day 245 and later stage cells show a progressive increase in proliferation, branching, and invasion into the deeper layers of the three-dimensional substrate. Day 528 cells show increased frequency of branching and superior ability to grow within the three-dimensional medium (93 ± 6% branching cells, Figure 5B, Table 1). Organotypic collagen raft assays demonstrate the contact inhibited monolayer growth on the raft surface without invasion at day 170. The acquisition of invasive capabilities began with day 245 cells, which demonstrated focal invasion of cells viable cells with atypical and heterogeneous nuclei (height 1, depth 1, pattern focal). Day 450 cells showed a diffuse pattern invasion to a depth of ~6 cells in most areas and increased cell density at the raft surface (Figure 5C). Day 528 cells additionally acquired the capacity for multilayered upward growth and stratification of cell layers, typically resulting in a cell height of ~6 cells or more above the raft surface. Invasion morphology was changed in the day 450 cells to now include both focal and diffuse patterns with atypical nuclear morphologies (Figure 5C). Cell phenotype alterations were compared with previous reports and, based on the total information gleaned from the *in vitro* studies, were

considered weakly positive for the transformed phenotype at day 245. MOSEC were considered certainly positive for the transformed phenotype at day 528, having acquired all characteristics previously linked to strong tumorigenicity *in vivo*.

Table 1. Characteristics of mouse ovarian surface epithelial cell key transformative stages

	170	245	450	528	ID8
PROLIFERATION					
doubling time [1] (hours)	36 ± 6	24 ± 4*	17 ± 2*	11 ± 1*	13 ± 2
colony forming units in agar [2] (%)	0.44 ± 0.21	2.6 ± 1.1*	8.5 ± 1.7*	16.8 ± 2.9*	14 ±3.5
INVASION					
‡ raft culture [3] (height/ depth/ pattern)	1/ 0/ 0	1/4 / F	1/ 6/ D	6/ 6/ FD	6/ 6.5/ D
branching cells [4] (%)	2 ± 1	43 ± 4*	63 ± 7*	93 ± 6*	96 ± 8
MIGRATION					
trans-well filter inserts [5]	2.1 ± 1.2	6.6 ± 2.8*	42 ± 17*	54 ± 9*	89 ± 16*
TUMORIGENICITY					
subcutaneous tumors [6] (% positive)	N	25%	N	50%	100%
subcutaneous tumors (mm)	N	1	N	10	10
subcutaneous tumors [7] (score/10)	N	1/10	N	6/10	7/10
intraperitoneal tumors (% positive)	N	-	N	50%	100%
intraperitoneal tumor [8] (score/10)	N	-	N	4/10	2/10

1) Doubling time $=(T)\ln 2/\ln(X_e/X_b)$, where X_n is the cell number determined at the beginning (X_b) and end (X_e) time poinst, and T is the total elapsed time (hours), seeding of 1 x10^4 and 5 x10^4 cells were harvested at 70% confluency. 2) Colony forming units in soft agar = mean colony count / number of cells seeded x 100%. 3) Organotypic collagen raft culture: height= number of cells above, depth= number of cells below the collagen. Pattern: D=diffuse, F=focal, 0=none. 4) Three-dimensional Matrigel branching = branched cells/ total cells. 5) Tran swell filter insert migration =(cells crossing matrigel / total cells) / (cells crossing open pores / total cells) 6) Tumor positive/ injected sites in eighteen syngeneic C57BL6 mice (245:n=7, 528:n=7, ID8: n=2, vehicle: n=2). Each mouse at each of 3 sites (left and right subcutaneous flank (1E^5) and intraperitoneal 1E^6) at day 70. 7) Average bilateral subcutaneous tumor invasion (dermis, muscle, subcutaneous fat, diffuse: 1point each/ 10 total). 8) Intraperitoneal tumorigenicity organ involvement and pattern (pancreas, liver, peritoneum, reproductive organs, lymph node, spleen, intestine, lung, heart, infiltrative, 1 point each/ 10 total).

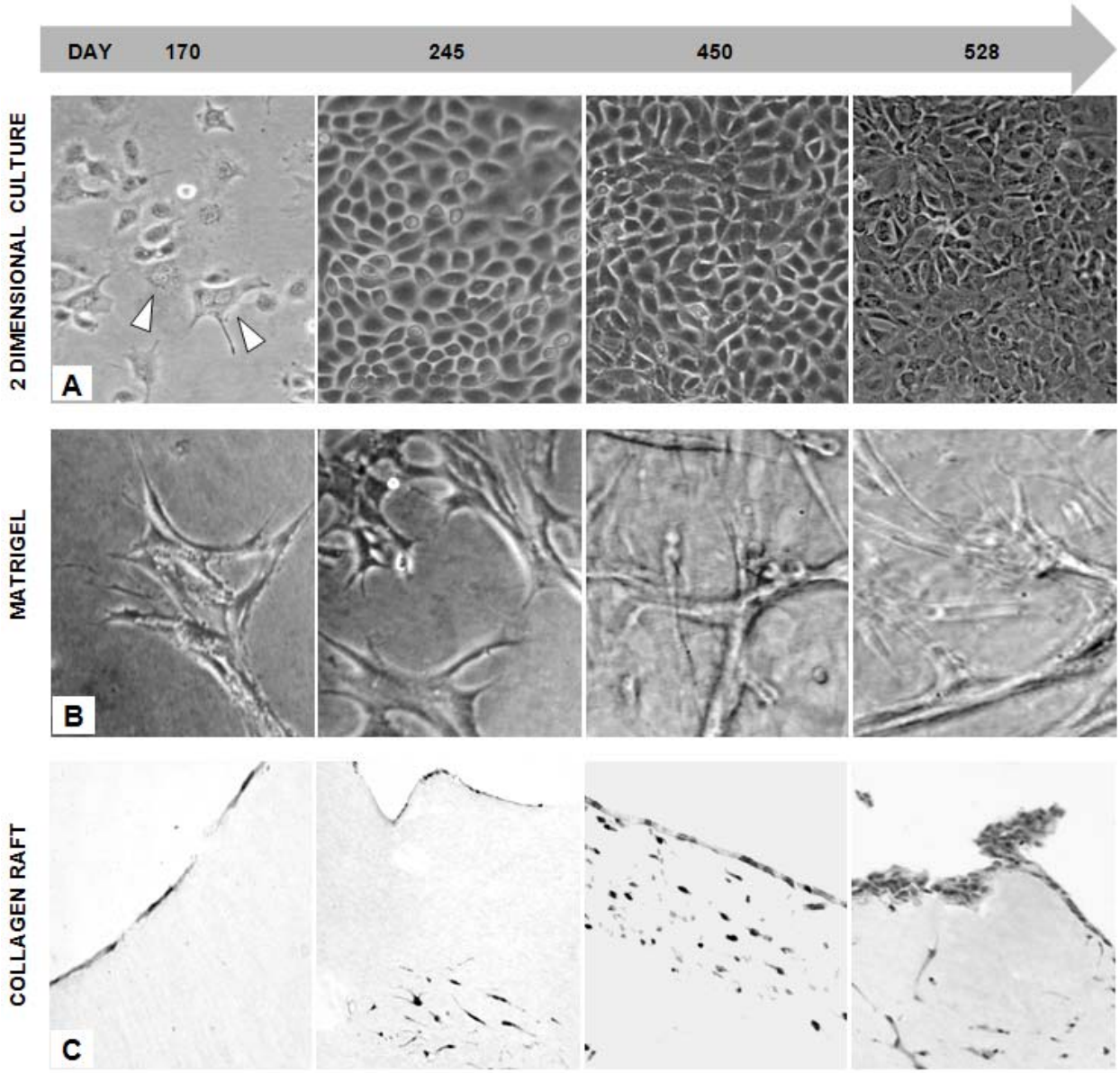


Figure 5. Cellular morphology and phenotype of key transformative stages during spontaneous transformation in vitro. Longitudinal analysis of cellular morphology shows multinucleate cells (arrowheads, A) with large surface area in day 170 cells, and the progressive decrease in cell surface area with the transition from cobblestone morphology at day 245 to nodular focal pattern at day 450, to spindle shaped tightly packed morphology at day 528 (A). Three dimensional cell growth in matrigel shows parallel morphologic changes and a phenotype of increasing invasiveness from day 170 cells with little to no invasion, day 245 cells and higher passages show a progressive increase in proliferation and depth of invasion into the matrigel (B). Cells grown on organotypic collagen rafts are characterized by contact inhibited monolayer growth with no invasion at day 170, focal invasion with atypical nuclei at day 245, diffuse invasion with loss of contact inhibition at the raft surface at day 450, and multilayered upward stratification above the raft surface with invasive, branched atypical cells at day 528 (C).

3.5 Population Based Karyotype Analysis of Key Transformative Stages

3.5.1 Summary Characteristics of Cell Populations by Transformative Stage

Spectral karyotype analysis of MOSEC key transformative stages was meticulously performed on 30-50 metaphase cells, resulting in ~1000 data points per transformative stage. These raw karyotype data are presented for comparative evaluation among MOSEC key transformative stages, tumor cells, and *Brca1*$^{\Delta 5-13}$ MOSEC to characterize the variability, as well as more standard cytogenetic reports including cell ploidy and the incidence of clonal and non-clonal chromosome aberrations (Table 2). Histogram reporting of raw data shows the expansion of variability acquired by the MOSEC from day 170-528 (Table 2). The first clonal chromosome aberration (CCA, ≥20% of cells) was seen at day 245 and these small pieces of chromosome 11 were without centromere (min 11) and persisted until day 528. Additional clonal chromosome aberrations arose at day 450 t4;3 and t6;15 (~30% of cells) and several (t4;3 t4;5 t5;4, min19) that were common to tumor 1 and 2. Interestingly neither t4;3 nor t5;16 were present at day 528. Non-clonal chromosome aberrations (<20% of cells) were present at each transformative stage and consisted predominantly of chromosomes linked by centromere at day 170, but expanded to include all types of aberrations from day 245 onwards.

3.5.2 Spectral Karyotype Analysis for Population Based Karyotype Heterogeneity

Karyotype heterogeneity at each transformative stage was measured using an adaptation of concepts published by Castro et. al. [80] for measuring tumor diversity among patient samples and applies the Shannon Index to chromosome count data. The method allows diversity to be quantified for normal or abnormally distributed data at each whole or aberrant chromosome and can be summed to generate an index for the entire karyotype (methods). This measure was applied to karyotypes of varying heterogeneity and served as a surrogate measure of genomic instability for each MOSEC transitional stage. The Shannon Index of diversity was calculated for each chromosome and is reported as the sum for whole

Table 2. Karyotype and variability of key transformative stages, Brca1 Δ5-13, and for harvested tumors

CHROM ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	X	CCA	NCCA
DAY 170	1(1) 2(27) 4(3)	2(28) 3(2) 4(1)	2(29) 3(1) 4(1)	1(2) 2(27) 4(2)	2(29) 3(1) 4(1)	2(28) 4(3)	1(1) 2(28) 4(2)	0(1) 2(27) 3(1) 4(2)	1(1) 2(29) 4(1)	1(2) 2(26) 3(1) 4(2)	0(1) 2(27) 3(1) 4(2)	2(28) 4(3)	1(1) 2(26) 3(2) 4(2)	2(27) 3(2) 4(2)	1(1) 2(26) 3(2) 4(2)	1(2) 2(26) 3(1) 4(2)	1(3) 2(25) 4(3)	1(1) 2(27) 4(3)	1(2) 2(25) 3(2) 4(2)	2(28) 3(2) 4(1)	-	t2, t13, t19, t7, t12;16, t12;19, t10;14, t10, t3=3, t5=5, t8=8, t9=9, t11=11, t13=13, , t14=14, tX=X, t1=1, t2=2, min(4, 8, 10, 15, 17, 19)
∑H'= 12=	0.5	0.4	0.3	0.5	0.4	0.6	0.5	0.6	0.4	0.7	0.5	0.5	0.6	0.4	0.5	0.5	0.5	0.5	0.7	0.5	+ 4 H'der	
DAY 245	0(1) 1(3) 2(11) 3(22) 4(3) 6(1)	1(1) 2(8) 3(12) 4(13) 5(3) 6(2) 7(1) 8(1)	1(3) 2(7) 3(9) 4(18) 5(2) 6(2)	1(1) 2(9) 3(20) 4(5) 5(3) 6(3)	2(8) 3(10) 4(12) 5(6) 6(5)	1(2) 2(9) 3(15) 4(9) 5(3) 6(2) 7(1)	1(1) 2(10) 3(18) 4(7) 5(4) 7(1)	3(7) 4(20) 5(10) 6(1) 7(1) 11(1) 16(1)	1(2) 2(6) 3(19) 4(10) 5(3) 6(1)	1(1) 2(12) 3(11) 4(10) 5(6) 6(1)	1(2) 2(3) 3(15) 4(10) 5(9) 6(2)	1(5) 2(27) 3(4) 4(4) 6(1)	1(3) 2(7) 3(21) 4(8) 5(1) 6(1)	0(1) 1(2) 2(11) 3(13) 4(9) 5(1) 6(3) 7(1)	1(1) 2(7) 3(11) 4(9) 5(12) 6(1)	1(1) 2(7) 3(17) 4(8) 5(6) 6(1) 7(1)	1(1) 2(3) 3(4) 4(20) 5(10) 6(1) 7(1) 9(1)	1(2) 2(11) 3(20) 4(7) 6(1)	1(2) 2(10) 3(18) 4(8) 5(2) 6(1)	0(2) 2(3) 3(8) 4(11) 5(10) 6(4) 7(3)	min 11	der10*, der11*, der16*, der5*, der10*, der12*, der5*, t12;19, t13;19, t14;17, t16;3, t17;15, t19;18, t7;14, t8;20, t9;12, t9;15, t3=3, t4=16, 4, 4*, min(5, 7, 10, 12, 15, 16, 18, 19)
∑H'= 23=	1.3	1.7	1.5	1.4	1.6	1.6	1.4	1.4	1.4	1.5	1.5	1.1	1.3	1.7	1.5	1.5	1.5	1.2	1.4	1.8	+ 7 H'der	
DAY 450	1(1) 2(7) 3(38) 4(3) 5(1)	1(1) 2(9) 3(13) 4(18) 5(9)	1(2) 2(2) 3(13) 4(27) 5(6)	2(12) 3(22) 4(15) 7(1)	2(3) 3(9) 4(22) 5(16)	2(3) 3(10) 4(25) 5(9) 6(3)	1(2) 2(9) 3(23) 4(13) 5(3)	2(1) 3(8) 4(30) 5(9) 6(2)	2(13) 3(19) 4(13) 5(5)	2(10) 3(21) 4(15) 5(4)	1(1) 2(3) 3(17) 4(17) 5(11) 6(1)	0(2) 1(3) 2(39) 3(6)	1(2) 2(15) 3(30) 4(2) 7(1)	1(1) 2(17) 3(21) 4(9) 5(1) 6(1)	1(1) 2(2) 3(8) 4(27) 5(10) 6(2)	2(12) 3(24) 4(11) 5(3)	2(3) 3(20) 4(20) 5(6) 6(1)	1(4) 2(12) 3(30) 4(4)	1(1) 2(5) 3(16) 4(26) 5(2)	1(1) 2(1) 3(4) 4(11) 5(12) 6(15) 7(3) 8(3)	der4;3, der6;15, min11	t12, t12;17, t13=13, t16;10, t17;19, t17=17, t7=12, t8;11, t8;14, t8;19, t9, t7=12, min(2, 1, 7, 10, 12, 13, 14, 15, 16, 17, 19, X)
∑H'= 22=	0.8	1.4	1.2	1.2	1.2	1.3	1.3	1.1	1.3	1.2	1.4	0.7	1.0	1.3	1.3	1.2	1.2	1.1	1.1	1.7	+ 6 H'der	
DAY 528	1(2) 2(28) 3(18) 4(1) 6(1)	0(1) 2(9) 3(13) 4(17) 5(8) 7(1) 8(1)	1(2) 2(5) 3(16) 4(12) 5(12) 6(2) 10(1)	1(3) 2(16) 3(18) 4(10) 5(2) 6(1)	1(1) 2(7) 3(10) 4(26) 5(3) 6(2) 7(1)	1(1) 2(8) 3(17) 4(20) 5(3) 7(1)	1(4) 2(17) 3(13) 4(15) 9(1)	2(6) 3(2) 4(20) 5(14) 6(5) 7(1) 8(1) 14(1)	1(1) 2(15) 3(17) 4(14) 5(2) 7(1)	0(1) 1(1) 2(12) 3(14) 4(10) 5(9) 6(2) 8(1)	0(1) 1(3) 2(8) 3(11) 4(19) 5(6) 6(1) 11(1)	0(1) 1(8) 2(31) 3(7) 4(2) 5(1)	1(4) 2(17) 3(23) 4(3) 5(2) 6(1)	1(3) 2(15) 3(23) 4(8) 8(1)	1(1) 2(10) 3(10) 4(20) 5(7) 6(1) 8(1)	0(1) 1(1) 2(16) 3(22) 4(6) 5(3) 6(1)	1(3) 2(5) 3(13) 4(24) 5(4) 8(1)	1(5) 2(18) 3(18) 4(7) 5(1) 6(1)	1(3) 2(5) 3(18) 4(13) 5(8) 6(3)	1(1) 2(5) 3(7) 4(13) 5(11) 6(10) 7(2) 8(1)	min11	t14;8, der15, t4;3, t7;15, derX, t15, der5, t15;14, t18*, t18^, t2;1, t2;14, t2;19, t2;5, der4, t4=12, min(3, 4, 12, 17, 18)
∑H'= 19=	1.0	1.6	1.6	1.4	1.4	1.4	1.4	1.6	1.4	1.7	1.7	1.2	1.3	1.3	1.5	1.4	1.4	1.4	1.6	1.8	+ 5 H'der	
Brca1 Δ4-13	1(3) 2(4) 3(37) 4(1) 6(3)	2(2) 3(3) 4(16) 5(18) 6(5) 7(2) 10(1) 12(1)	2(4) 3(5) 4(17) 5(12) 6(5) 7(2) 9(2) 13(1)	1(1) 2(5) 3(34) 4(5) 6(2) 7(1)	2(1) 3(8) 4(19) 5(14) 6(2) 7(3) 14(1)	2(1) 3(7) 4(17) 5(15) 6(3) 7(2) 8(1) 9(1) 11(1)	1(3) 2(7) 3(32) 4(4) 6(1) 7(1)	2(2) 3(3) 4(33) 5(5) 6(2) 7(2) 9(1)	2(5) 3(21) 4(17) 5(2) 7(2) 10(1)	3(2) 4(19) 5(17) 6(5) 7(4) 8(1)	3(1) 4(5) 5(16) 6(15) 7(5) 8(4) 10(1) 13(1)	1(1) 2(4) 3(23) 4(15) 5(2) 6(2) 8(1)	2(1) 3(7) 4(9) 5(21) 6(6) 7(1) 10(2) 11(1)	2(6) 3(14) 4(19) 5(5) 6(1) 7(1) 8(2)	0(1) 2(1) 3(2) 4(7) 5(17) 6(11) 7(6) 9(1) 12(1) 14(1)	3(8) 4(26) 5(6) 6(3) 7(1) 8(1) 10(2) 12(1)	2(1) 3(6) 4(21) 5(15) 6(2) 8(2) 9(1)	2(3) 3(13) 4(21) 5(8) 7(2) 11(1)	1(1) 2(1) 3(1) 4(4) 5(22) 6(13) 7(2) 8(2) 11(1) 12(1)	1(1) 2(12) 3(12) 4(12) 5(5) 6(4) 9(1) 10(1)	-	der10, der11, der14, der16, der1, der5, der7, der4*, derX*, der12*,der13, der^7, t7;10, t4=4, t19=19t12;9, t13;15, t15;10, t18;8, t3;4, t3;4, t4=4;X, min(4, 7, 10)
∑H'= 15=	0.8	1.2	1.4	0.4	1.2	1.3	0.9	1.3	1.1	1.4	1.4	1.0	1.1	1.5	1.3	1.6	0.9	1.6	1.2	1.5	+ 3 H'der	
TUMOR 1	0(1) 1(1) 2(28) 3(15) 4(2) 5(3)	2(6) 3(16) 4(16) 5(1) 6(7) 8(3) 10(1)	0(1) 1(1) 2(14) 3(18) 4(7) 5(4) 6(2) 7(2) 8(1)	2(18) 3(21) 4(6) 5(3) 6(2)	1(3) 2(8) 3(17) 4(15) 5(4) 6(1) 7(2)	2(7) 3(17) 4(18) 5(6) 7(1) 8(1)	2(14) 3(20) 4(8) 5(4) 6(3) 7(1)	2(1) 3(2) 4(22) 5(11) 6(3) 7(1) 8(5) 9(2) 10(1) 11(1) 13(1)	1(4) 2(20) 3(18) 4(3) 5(2) 6(3)	0(1) 1(1) 2(7) 3(18) 4(12) 5(5) 6(3) 7(1) 10(1) 14(1)	1(1) 2(9) 3(15) 4(10) 5(7) 6(2) 7(4) 8(2)	1(4) 2(25) 3(7) 4(9) 5(2) 6(1) 8(2)	2(23) 3(11) 4(9) 5(1) 6(3) 8(1) 10(1) 14(1)	1(6) 2(18) 3(11) 4(13) 5(1) 6(1)	1(2) 2(4) 3(18) 4(18) 5(3) 6(2) 7(2) 10(1)	2(3) 3(19) 4(16) 5(5) 6(3) 7(2) 8(1) 9(1)	3(12) 4(25) 5(3) 6(1) 7(6) 8(2) 9(1)	1(1) 2(9) 3(19) 4(10) 5(4) 6(2) 7(4) 11(1)	1(5) 2(7) 3(7) 4(20) 5(4) 6(2) 7(5)	2(4) 3(6) 4(18) 5(8) 6(5) 7(2) 8(4) 9(2) 10(1)	t4;5 t5;4 t4;3* t11;13 min 19	t14;8, der15, t7;15, derX, der5, t15;14, t18*, t18^, t2;1, t2;14, t2;19, t2;5, der4, t4=12, min(3, 4, 12, 17, 18)
∑H'= 23=	1.1	1.6	1.7	1.3	1.6	1.4	1.5	1.7	1.4	1.8	1.8	1.5	1.5	1.5	1.6	1.6	1.4	1.7	1.7	1.9	+ 7 H'der	
TUMOR 2	1(1) 2(6) 3(30) 4(8) 5(3) 6(1) 10(1)	2(7) 3(11) 4(14) 5(10) 6(3) 7(2) 8(1) 10(2)	1(1) 2(3) 3(21) 4(14) 5(6) 6(1) 7(1) 8(2) 10(1)	0(1) 1(3) 2(14) 3(13) 4(13) 5(3) 6(1) 7(1) 8(1)	2(12) 3(20) 4(13) 5(2) 6(2) 7(1)	2(8) 3(22) 4(9) 5(6) 7(3) 8(1) 9(1)	0(1) 1(1) 2(8) 3(23) 4(9) 5(3) 6(1) 7(2) 8(2)	1(1) 2(6) 3(8) 4(17) 5(9) 6(5) 8(3) 9(1)	0(1) 1(4) 2(15) 3(11) 4(11) 5(4) 6(3) 7(1)	1(2) 2(12) 3(12) 4(13) 5(3) 6(3) 7(3) 9(1) 12(1)	0(1) 1(1) 2(5) 3(15) 4(12) 5(7) 6(1) 7(6) 9(1) 10(1)	0(1) 1(5) 2(25) 3(8) 4(5) 5(2) 6(3) 8(1)	0(1) 1(3) 2(24) 3(13) 4(6) 5(1) 6(1) 8(1)	1(2) 2(23) 3(12) 4(7) 5(1) 6(3) 7(1) 8(1)	2(4) 3(10) 4(26) 5(6) 8(3) 11(1)	1(2) 2(2) 3(24) 4(9) 5(7) 6(3) 8(2) 9(1)	0(1) 1(1) 2(5) 3(12) 4(19) 5(5) 6(1) 7(5) 10(1)	1(2) 2(12) 3(23) 4(8) 5(3) 8(1) 12(1)	0(1) 1(2) 2(13) 3(16) 4(8) 5(4) 6(4) 8(1) 12(1)	1(1) 2(8) 3(8) 4(11) 5(9) 6(3) 7(5) 8(3) 10(1) 12(1)	t4;3, t5;4, t4;5, min(10, 11, 13, 19, X)	der10, der11, der14, der16, der1, der5, der7, der4*, derX*, der12*,der13, der^7, t7;10, t4=4, t19=19t12;9, t13;15, t15;10, t18;8, t3;4, t4=4;X, min(4, 7)
∑H'= 29=	1.3	1.8	1.6	1.7	1.4	1.5	1.6	1.8	1.8	1.8	1.9	1.6	1.4	1.5	1.4	1.6	1.7	1.4	1.8	2.0	+ 13 H'der	

*Chromosome ID by column header, key transformative stage or cell type by row end (left). Chromosome count data for: count (frequency), derivative chromosomes: * telomeric aberration, ^ centromeric aberration; translocations (t, where ";" for chromosome arm, "=" centromeric linkage), double minute (min), clonal chromosome aberration (CCA≥20% of cells), non-clonal chromosome aberration (NCCA<20% of cells, Normalized Shannon Index of Variability (H').*

chromosomes, sum for aberrant chromosomes, and the total of both indices. Population karyotype characteristics are summarized for chromosome counts, ploidy distribution Shannon indices of variability, CCAs and NCCAs in Table 3. Day 170 cells were the only stage characterized by 2n ploidy status. Ploidy increased to double this (4n) in day 245, 450, and was 3n in day 528 cells and tumor derived MOSEC. *Brca1*$^{\Delta 5-13}$ also displayed high ploidy level (4n) despite their short time in culture (50 days). Karyotype abnormalities were detected in early stage day 170 cells at a frequency of 48% abnormal karyotypes (Table 3). The median count for each chromosome at day 170 was 1.9 ($\bar{x}$ = 2.1; range = 2.1-2.2), with 32% of the cells containing at least one whole chromosome imbalance. This increased to day 450, and decreased slightly at day 528. In contrast to the early passage MOSEC, whole chromosome analyses at subsequent time points revealed aneuploidy in 98-100% of cells. The Shannon index values (H' whole chromosome, ∑H') were elevated in day 245, 450, and 528 cells compared to day 170 cells, and followed a similar pattern of elevation until day 450, with a small but significant reduction in day 528 cells.

The Shannon Index for whole chromosome diversity ranged from 10 to 29, climbing to reach a maximum of H=29 at day 245 and 528, falling slightly at day 450. The total Shannon Index (H) follows a similar trend, as it is largely driven by whole chromosome heterogeneity. The Shannon Index for derivative chromosomes begins at 2 at day 170, rises to 4 at day 450, and falls again to 2 at day 528 with a similar pattern to ∑H' values. The Shannon Index value (H') is corrected for maximum variability (H'=H/ lnS). In descriptive terms, H' is a measure randomness within the constraints of the variability inherent in the population at that time with a limited range ($0 \leq H' \leq 1$ for each whole and derivative chromosome). At day 170, ∑H' was at a minimum (∑H'=13), and showed significant increase over day170 values at each subsequent stage (245, ∑H'=35; 450, ∑H'= 31, 528, ∑H'=29, Table 2 and Figure 6).

3.5.3 Characterizing the extent of cell-to-cell heterogeneity by karyotype

A visual impression of the relationship between cellular heterogeneity and key transformative stage during spontaneous MOSEC transformation is seen in the karyograph plots (Figure 6). By using three-dimensional plots of complete karyotype data for each transitional stage, the dynamic changes in the karyotype profiles and the extent of the karyotypic heterogeneity at each stage are readily apparent. Each line on the karyograph shows chromosome count data for normal and derivative chromosomes, thus presenting data in a format that preserves its relationship to the individual cell. Normal and abarrent chromosomes are identified along the x-axis, corresponding counts for each chromosome are plotted on the y-axis, and metaphase number (M_1,M_{50}) on the z-axis. The predominant karyotype at day 170 is diploid (2n) and, with the exception of some rare peaks and infrequent tetraploid cells, is remarkably similar to the normal mouse genome. Rare changes in whole chromosome copy number and in occasional derivative chromosomes are seen in several cells (Figure 6A). There is little to no inherent noise in the SKY data at this stage. Metaphase spreads from day 245 cells showed an increased median chromosome number and a significant increase in variability of chromosome count at each position (Figure 6B). Total un-normalized karyotype heterogeneity fell slightly in day 450 cells ($\sum H=35$, Figure 6C), when the occurrence of clonal 4;3 and 6;15 translocations appeared in addition to the clonal min11 double minute noted in the previous stage. These translocations were found at a frequency of ~30% (Figure 6C). Neither clonal chromosome aberration was seen in the late stage day 450 MOSEC ($\sum H=32$, Figure 6D).

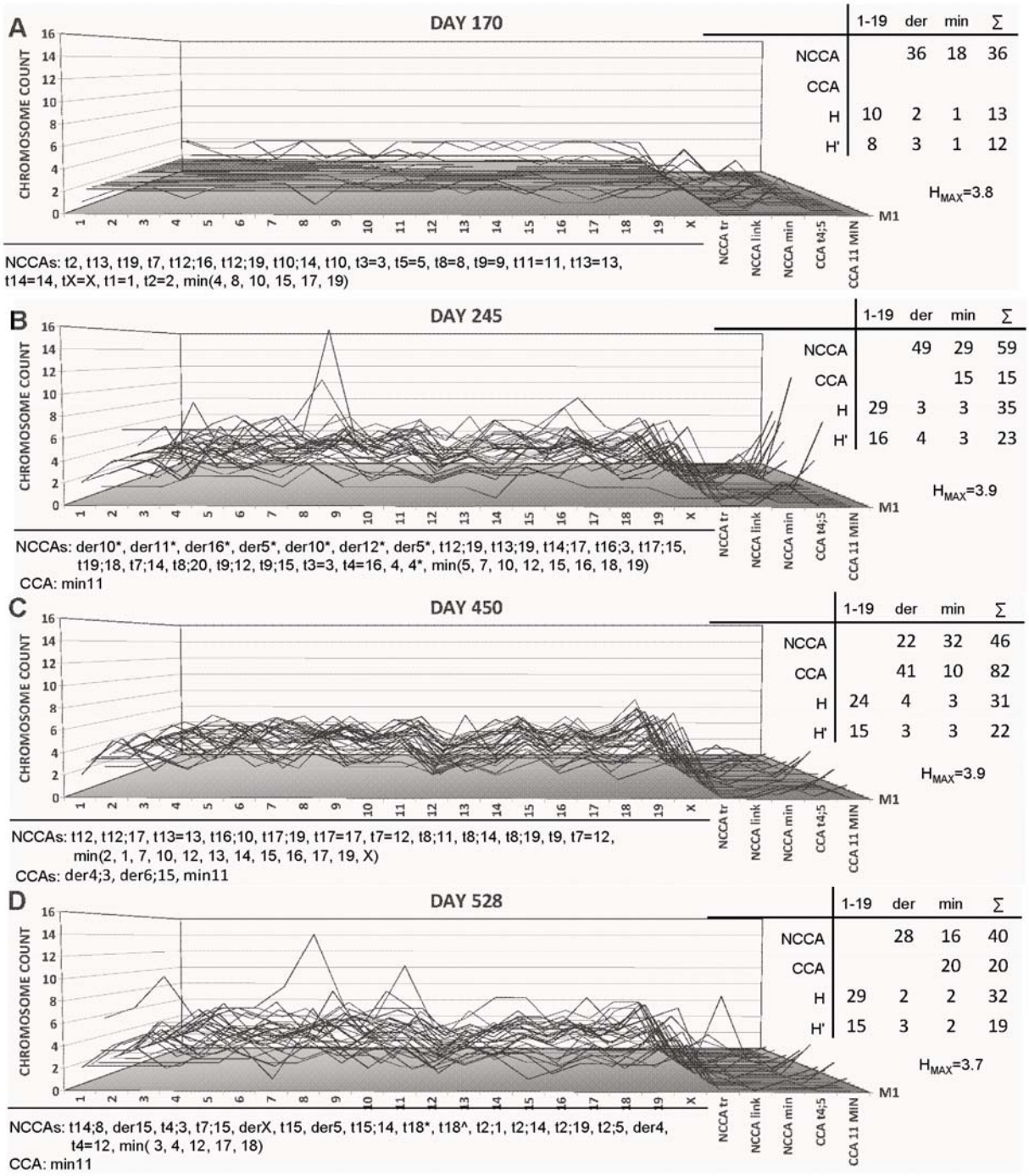


Figure 6. Karyographs showing the extent of karyotype diversity throughout the spontaneous transformation of mouse ovarian surface epithelial cells. Day 170 cells had near-normal karyotypes with few non-clonal chromosome aberrations (A). Cells at day 245 had significantly increased median chromosome number at each position, and showed a great deal of heterogeneity among cells. Day 450 cells further increased their heterogeneity which continued until day 450, where a 4;3 clonal translocation is present in 34% of the cells (C). Shannon Index (SI) for population diversity was calculated as the sum of the SI for each chromosome and was summed at each timepoint and significantly increased between 170:245, and between 245:528 ($p<0.05$). Ploidy (y-axis) of clonal and non-clonal whole and derivative chromosomes (x-axis) is shown for 50 metaphases per time point (M1-M50, z-axis).

3.5.4 Comparing the distribution of count data by chromosome across the key transformative stages

To demonstrate the variability contributed by varying distributions of each chromosome measured in the population karyotype data, violin plots of chromosome counts and the distribution for each are presented across the transitional stages (Figure 7A-D). The median chromosome count (y-axis, right) for each whole chromosome and the sum of all derivative chromosomes is shown (gray circles x-axis), with 95% confidence intervals for each (black lines within curve), maxima and minima by the thin lines projecting from each curve, and distribution of each chromosome count by the width of the curve at that count. From these analyses, certain patterns can be recognized over time that are not clear from the karyograph plots. For example, the uniformity of chromosome count distribution is better seen in this plot. Additionally, the findings that certain chromosomes (chromosome 12) and chromosome 1 approached significance as chromosomes that did not undergo significant change in count over time can also be seen (Figure 7, B-D). Chromosomes 1 and 12, showing little change over time additionally have the lowest Shannon Index at each key transitional stage, indicating stability of that chromosome over time is related to the diversity of chromosome count measured at each transitional stage. Oppositely the X chromosome has the largest and most persistent increase in chromosome counts and a high Shannon Index (Figure 7A-D, Table 2). The difference in distribution profiles between the Brca1 and the spontaneously derived line can also be seen. Finally, the increase in the maximum chromosome count for day 528, Brca1, Tumor 1 and Tumor 2 is seen by the height of these upward going lines from the body of the violin plot as compared to the earlier stage (day 170, 245, 450) spontaneously transforming lines.

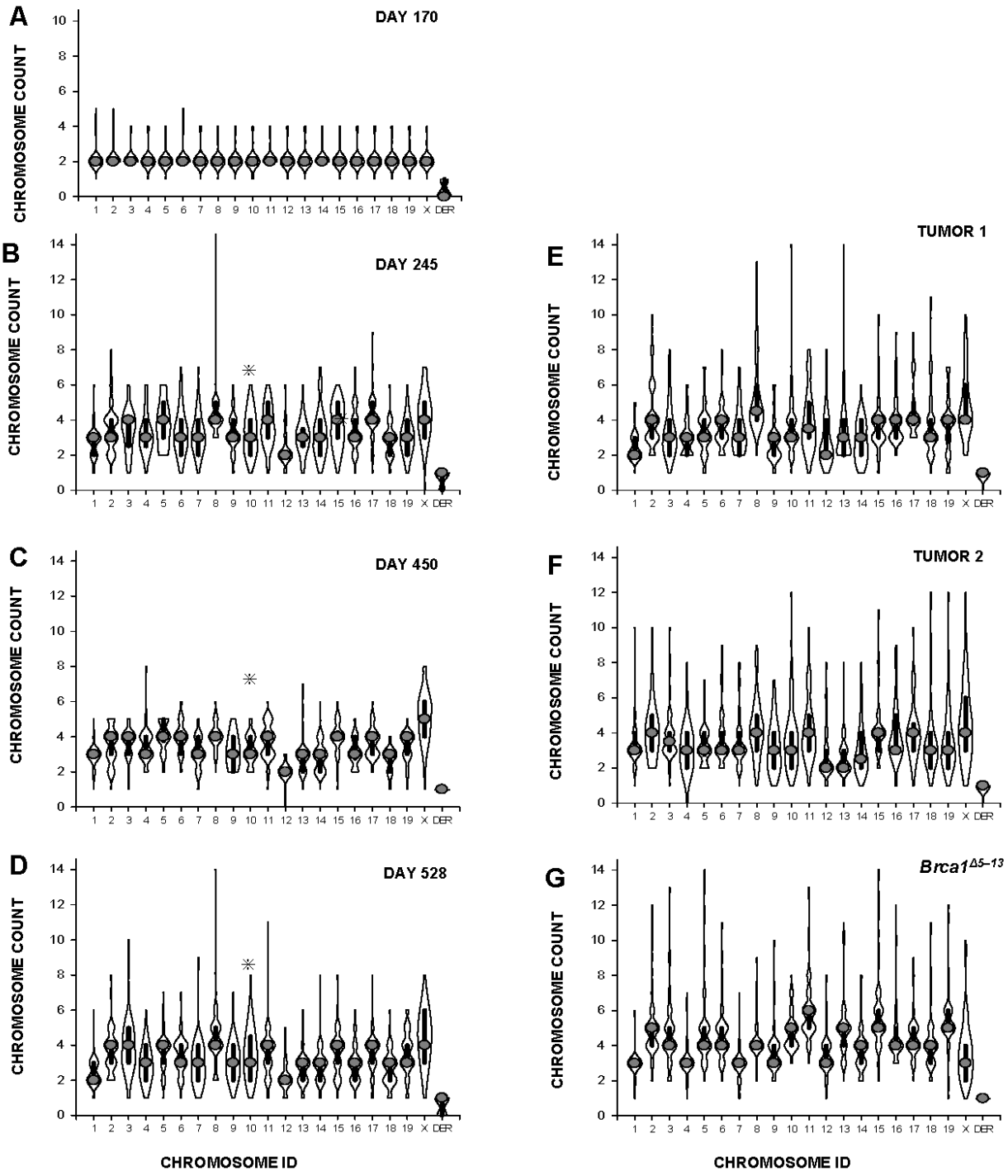


Figure 7. Violin plots and Shannon Index characterizing karyotype diversity of key transformative stages, harvested tumors, and Brca1$^{\Delta 5-13}$ line. Day 170 cells had diploid near-normal karyotypes with few derivative chromosomes (A). Day 245, and day 450 cells had increased median chromsome numbers (B, C). Shannon Index continued to increase in day 450 cells (1.12-1.69, C), and decreased slightly in the last stage, day 528, (0.98-1.57). Counts (y-axis) of the intact whole and all derivative chromsomes, (x-axis) for 50 individual metaphases per transitional stage. Variability was increased in both tumors above the cell lines (E, F) and was matched between day 528 and Brca1$^{\Delta 5-13}$ cells. For violin plots, gray circles are median chromosome number, black lines are 95% confidence interval, width of plot is proportional to the frequency of chromosomes with a given count. Chromosome 12, without significant change over time (*).

Table 3. Population karyotype characteristics for mouse ovarian surface epithelial cell stages and lines.

		170	245	450	528	Brca1$^{\Delta 5-13}$	TUMOR 1	TUMOR 2
chromosome count	mode	40	75†	71	67	82‡	62	61
	min-max	36 - 77	26 - 120†	45 - 84†	31 - 148†	43 - 165	42 - 129	37 - 159
	median	1.9	3.2	3.6	3.4	4.2	3.2	3.3
ploidy range		2.1 - 2.2	3.3 - 3.4	3.4 - 3.4	3.2 - 3.3	4.3 - 4.4	3.6 - 3.7	3.6 - 3.6
ploidy best fit		2n	4n	4n	3n	4n	3n	3n
whole chromosome	loss	0	1	13	9	5	6	4
	gain	0	12	4	14	10	13	11
	change	0	13	16	20	15	18	14
2n (%)		90	7	2	14	4	2	4
3n (%)		3	49	36	42	4	74	66
4n (%)		6	37	62	42	81	4	18
5n (%)		0	5	0	0	2	6	4
>5n (%)		0	2	0	2	8	14	8
aneuploid cells (%)		48	100	100	98	100	100	100
H	(1-19)	10	29†	24†	29	29	31	33
(der)		2	3	4	2	3	3	6
(min)		1	3	3	2	0	4	6
∑ H		13	35	31	32	32	38	44
H'	(1-19)	8	16†	15†	15†	14	16	14
(der)		3	4	3	3	4	7‡	1‡
(min)		1	3	3	2	1	6	3
∑ H'		12	23†	22†	19†	19	29‡	18‡
CCA	translocation (%)	0	0	30	0	0	96	98
linked (%)		0	0	0	0	0	0	0
minute (%)		0	44	30	34	0	90	0
∑CCA (%)		0	44	54	34	0	98	98
NCCA	translocation (%)	19	41	20	24	19	6	58
linked (%)		19	5	2	2	2	22	0
minute (%)		0	29	32	16	3	22	0
∑NCCA (%)		32	59	46	40	20	42	58

Percentages indicate the percentage of metaphase cells belonging to that category.
Ploidy best fit defined by modal chromosome number (m), "n" is defined such that "m" falls within 20n ±10.
Whole chromosome loss and gain defined as the difference between the predicted (best fit ploidy) count and the actual chromosome count. CCA, clonal chromosome aberration present in ≥20% of metaphase cells; NCCA, non-clonal chromosome aberration present in <20% of metaphase cells. H, Shannon Index of karyotypic heterogeneity; H' Shannon index of karyotypic heterogeneity normalized for maximum population variability.
† transformative stage significantly different from adjacent time point, ‡ different from day 528.

3.6 Variability within sub-populations identified by clonal chromosome aberrations

To determine if the presence or absence of a marker clonal chromosome translocation was linked to variability of whole chromosome counts, data from day 450 cells were divided into subpopulations based on the presence or absence of the clonal 4;3 translocation (34% of day 450 cells, Figure 8A). Variability of whole chromosome copy number is shown to exist in karyotype tables showing the genomes of four CCA t4;3 cells (Figure 8A). Among these four karyotypes, 18/20 possible chromosome positions demonstrate variability of whole chromosome copy number. Additionally, an example of a non-clonal chromosome translocation is provided which additionally contributes to the genomic heterogeneity of the CCA t 4;3 line (Figure 8A). Presence of the clonal chromosome marker t 4;3 was not linked to variability of chromosome count as measured by standard deviation of the mean for each chromosome ID. Additionally, H' value for each group by total chromosome count approached the theoretical maximum (1 for each chromosome) and indicated slightly higher variability in the subpopulation carrying the clonal marker compared to the remaining day 450 cells (H' CCA t4;3 = 0.96 versus H' CCA none = 0.91, Figure 8B).

3.7 Analyses of differential mRNA transcript abundance from key transformative stages

3.7.1 Time course analysis and biological significance for mRNA temporal expression profiles of key transformative stages

To determine the gene expression corresponding to each key transformative stage and its karyotype profile, genome wide analysis of mRNA transcript abundance was performed. Time course analysis was implemented with EDGE software as a robust method for determining differential gene expression over time and discovered 599 of the 18,000 candidates genes represented on the Illumina MouseRef-8 v2.0 chip. The following exclusionary criteria were used: maximum signal >3x background, >2.5 fold change during the time course, and statistical significance $q < 0.01$ to reject the null hypothesis of constant gene expression over time. Rather

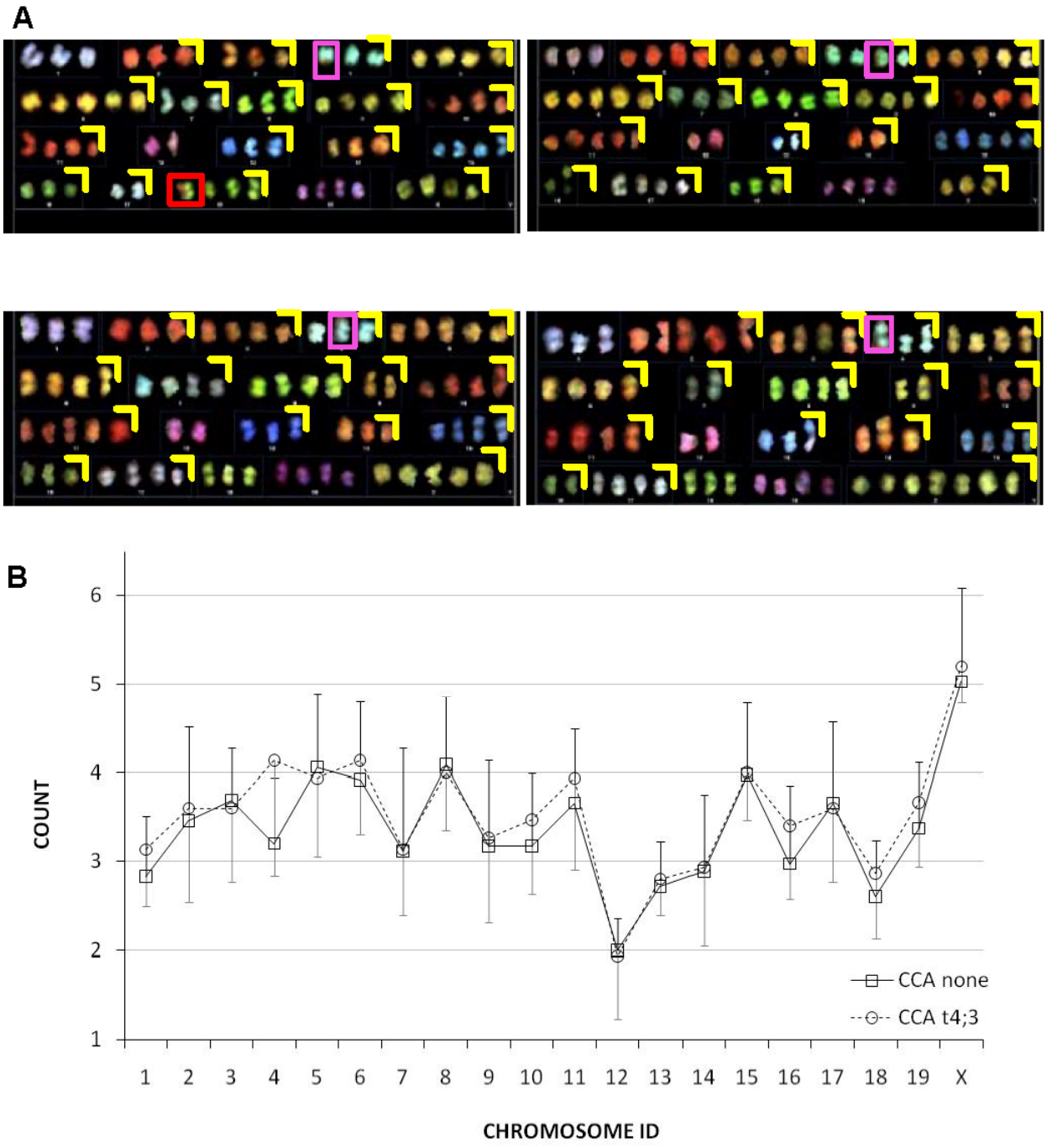
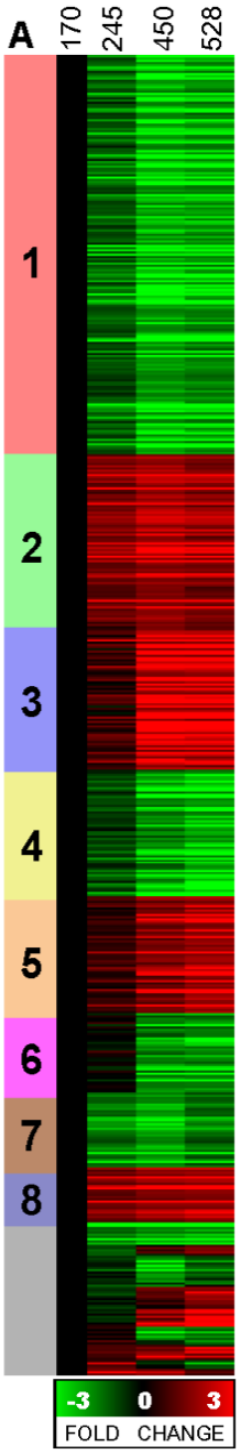


Figure 8. Karyotypic variability in day 450 subpopulations by presence or absence of clonal 4;3 translocation. SKY images of four unique karyotypes showing variability of whole chromosome copy number in 18/20 positions (yellow corners) among cells within the CCA t4;3 subpopulation (pink boxes, A). A non-clonal contributor to genomic heterogeneity is additionally present (red box, upper left panel, A). Average chromosome count and standard deviation are plotted for cells with the clonal 4;3 chromosome translocation (CCA t4;3: open circles, grey downward going error bars) versus cells without this marker (CCA none: black squares, black upward going error bars, B). Average count was higher for cells with the 4;3 translocation (CCA 4;3=3.56 versus CCA none=3.41, p=0.02). Standard deviation of chromosome count and average H' approached 1 (SD:CCA t4;3=0.86 versus CCA none=0.81, p~0.4; H':CCA t4;3=0.96 versus CCA none=0.91).

than arbitrarily grouping the 599 significant genes based increase *versus* decrease between consecutive time points, the gene list and corresponding expression values were input into short-time course-expression-miner (STEM) software which tests gene expression profile for each gene to determine the relative goodness of fit to a randomly generated profiles of fold change over time versus constant expression. STEM software matched 530/599 differentially expressed genes to 8 unique profiles, where the maximum allowable correlation coefficient between profiles was <0.9. Unmatched gene sets comprise the remaining 69 genes at the bottom of the heat map (white block, *Figure 9*A). Fold change and enrichment for cluster membership and the defining characteristics of each color-coded cluster and its corresponding heat-map intensity plot are summarized (*Figure 9*B). *Figure 9*C shows plots (y-axis-fold change, x-axis=days) for the genes matched to each profile (red lines), which ranged in number from (12-181) and are plotted as red lines against the normalized median profile (black line) as a function of days in culture (Figure 9C). The biological significance of enriched genes from each profile was determined using GO enrichment analysis for biological process, molecular function, and cellular component. These were significant for functional gene enrichment in profile 2 and profile 5 where corrected $p < 0.05$ (Figure 9D). Profile 2 enrichment lists categories relating to chromatin, chromosome, protein-DNA complex assembly, nucleosome organization and assembly, cellular macromolecular complex assembly and subunit organization, RNA processing and others. Profile 5 is enriched for cholesterol biosynthetic processes, metabolic processes, microsome and vesicular formation. Both profiles with functional gene enrichment show a net increase in gene expression during transformation, but differ in the magnitude of their slopes (450-528). Additionally, profile 2 genes slightly decrease expression between days 450 and 528, whereas profile 5 genes continue to increase, (Figure 9C, D). A compliation of matched and unmatched genes lists genes by membership in each fold change by profile (Table 4).

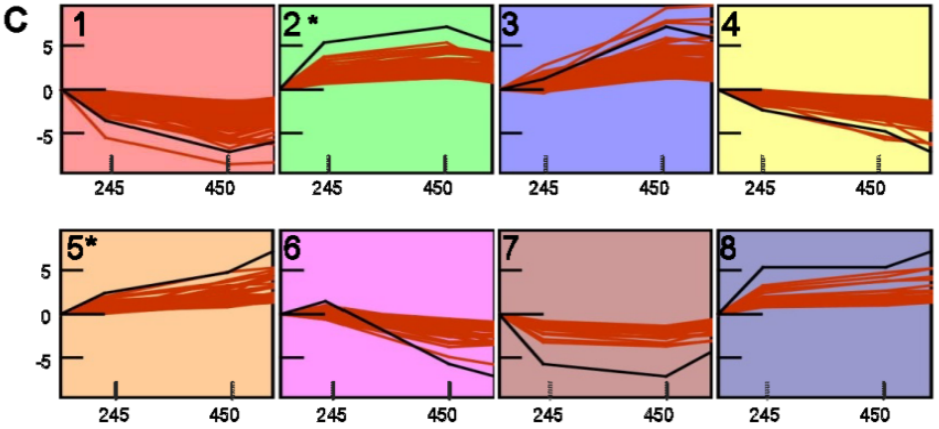
A



B

Profile ID	Profile Fold Change 170	245	450	528	Genes Expected	Genes Assigned	Profile Enrichment p-value	Significant GO Enrichment: Number of Categories	p-value
1	0	-3	-6	-5	17	181	2_E^{129}		
* 2	0	3	4	3	12	78	5_E^{39}	21	<0.04
3	0	1	6	5	23	65	1_E^{25}		
4	0	-1	-2	-3	10	58	9_E^{22}		
* 5	0	1	2	3	10	53	5_E^{16}	5	<0.05
6	0	1	-4	5	20	36	2_E^{7}		
7	0	-4	-5	-3	18	34	6_E^{4}		
8	0	3	3	4	7	25	9_E^{4}		

C



D

Profile ID	GO ID	GO Category Name	Genes Expected	Genes Assigned	GO Enrichment p-value
2	0000785	chromatin	2.9	10.1	<0.001
	0006323	DNA packaging	2.9	10.1	<0.001
	0071103	DNA conformation change	2.9	10.1	<0.001
	0051276	chromosome organization	3	10	<0.001
	0065004	protein-DNA complex assembly	2.7	9.3	<0.001
	0034728	nucleosome organization	2.7	9.3	<0.001
	0006325	chromatin organization	2.7	9.3	<0.001
	0006334	nucleosome assembly	2.7	9.3	<0.001
	0006333	chromatin assembly or disassembly	2.7	9.3	<0.001
	0031497	chromatin assembly	2.7	9.3	<0.001
	0006996	organelle organization	6.6	12.4	<0.001
	0034621	cellular macromolecular complex subunit org.	3.5	9.5	<0.001
	0034622	cellular macromolecular complex assembly	3.5	9.5	<0.001
	0000786	nucleosome	2.6	8.4	<0.001
	0032993	protein-DNA complex	2.6	8.4	<0.001
	0006396	RNA processing	1.2	5.8	<0.001
	0044428	nuclear part	2.3	7.7	<0.001
	0065003	macromolecular complex assembly	4	9	0.006
	0043933	macromolecular complex subunit org.	4	9	0.006
	0005681	spliceosomal complex	0.5	3.5	0.04
	0022607	cellular component assembly	4.8	8.2	0.04
5	0006695	cholesterol biosynthetic process	0.5	4.5	0.002
	0016126	sterol biosynthetic process	0.7	4.3	0.012
	0008203	cholesterol metabolic process	0.8	4.2	0.046
	0005792	microsome	0.8	4.2	0.046
	0042598	vesicular fraction	0.8	4.2	0.046

Figure 9. Timecourse profiling, cluster analysis, and biological significance for each of eight significant temporal expression profiles during tumorigenesis. Color-coded clusters match heat-map for each profile ID 1-8 comprising 530 profile-matched and 69 unmatched genes differentially expressed during transformation (A). Tabulated characteristics of fold changes, significance of genes for each profile, and Gene Ontologies (GO) functional enrichment analysis (B). Eight corresponding color-coded tiles showing fold change in gene expression for individual genes matched to that profile (red lines) and the profile median (black line) as a function of days in culture (x-axis), profile ID (upper left corner, (C)). GO ID number, category name, and functional enrichment data for biological process, molecular function, and cellular component of profiles with significant biological enrichment. GO results are listed in the table according for each cluster where corrected p-value <0.05 (D).

Table 4. Differentially expressed genes during tumorigenesis

(1) genes=181 ($2e^{-129}$)	170	245	450	528	(2) genes=78($5e^{-39}$)	170	245	450	528
1110033j19RIK	0	-1.9	-4.0	-3.8	*1110067d22RIK*	0	0.8	1.4	1.3
1200013b22RIK	0	-0.6	-2.2	-1.9	*2410008j05RIK*	0	1.9	2.2	2.0
1700011h14RIK	0	-0.9	-2.9	-2.3	*2610318c08RIK*	0	1.5	1.8	1.4
1700019e19RIK	0	-1.1	-1.6	-1.3	*2610510j17RIK*	0	1.4	1.7	1.5
1700088e04RIK	0	-1.2	-2.2	-1.6	*2610524h06RIK*	0	1.2	1.5	1.1
1810015a11RIK	0	-2.5	-3.6	-3.3	*4832406c22*	0	2.0	2.2	1.4
2300002d11RIK	0	-2.0	-2.6	-2.3	*Asf1b*	0	1.3	1.8	1.3
2310044g17RIK	0	-0.5	-1.4	-1.2	*Birc5*	0	2.1	2.2	1.6
2700055k07RIK	0	-1.1	-3.7	-3.6	*Bub1b*	0	1.5	1.8	1.7
2810003c17RIK	0	-1.6	-3.2	-2.4	*Ccl7*	0	2.8	4.4	4.0
2810417j12RIK	0	-2.1	-3.2	-3.2	*Cdc42ep2*	0	1.7	2.3	2.3
3110004l20RIK	0	-0.8	-1.7	-1.6	*Cdkn2a*	0	1.3	2.2	2.0
4732481h14RIK	0	-1.0	-1.7	-1.8	*Cldn1*	0	2.0	2.5	2.4
4930402h24RIK	0	-1.4	-2.2	-1.8	*Col5a1*	0	1.0	1.9	1.5
4933405a16RIK	0	-1.5	-2.5	-1.8	*Copg*	0	1.1	1.4	1.3
5133400g04RIK	0	-0.8	-1.8	-1.4	*Cyp7b1*	0	3.2	4.6	4.0
5730469m10RIK	0	-0.8	-1.3	-1.1	*Drctnnb1a*	0	1.0	1.8	1.6
5830467p10RIK	0	-2.6	-4.2	-4.0	*Eg433923*	0	0.7	1.4	1.0
6330580j24RIK	0	-1.5	-2.0	-2.0	*Emp3*	0	1.4	2.0	1.8
9130213b05RIK	0	-0.6	-3.1	-2.6	*Espl1*	0	1.5	2.2	1.8
9830002i17RIK	0	-1.4	-3.1	-2.9	*Exosc8*	0	1.1	1.6	1.4
Acvr2b	0	-0.2	-1.5	-1.1	*Fkbp2*	0	1.7	2.0	1.9
Agtrap	0	-0.9	-1.5	-1.6	*Gap43*	0	3.7	5.3	2.8
Akap12	0	-0.9	-2.2	-1.7	*Gnb4*	0	1.2	1.4	1.0
Aldh1a1	0	-2.1	-6.0	-4.5	*H2afz*	0	1.4	2.0	1.8
Ankrd1	0	-1.9	-2.8	-3.1	*Hey1*	0	1.2	2.3	1.8
Anxa11	0	-0.6	-1.5	-1.0	*Hist1h2ad*	0	1.8	2.6	2.3
Anxa6	0	-0.7	-1.5	-1.5	*Hist1h2af*	0	1.5	2.2	2.0
Apg10l	0	-0.7	-1.5	-1.3	*Hist1h2ag*	0	2.0	2.7	2.6
Aplp1	0	-1.9	-3.5	-3.6	*Hist1h2ah*	0	1.7	2.5	2.3
Atp6v0e2	0	-0.8	-3.1	-2.2	*Hist1h2ak*	0	1.8	2.6	2.3
Aw146242	0	-0.9	-1.6	-1.6	*Hist1h2an*	0	1.5	2.2	2.1
Aw555464	0	-0.8	-2.0	-1.6	*Hist1h2ao*	0	1.2	1.5	1.5
Axin2	0	-1.4	-2.8	-2.5	*Hist1h4f*	0	1.5	2.0	1.3
B430104h02RIK	0	-1.1	-2.7	-2.7	*Hist2h2ab*	0	1.9	2.7	2.7
B930041f14RIK	0	-2.7	-3.7	-3.7	*Hist2h2ac*	0	1.3	1.6	1.5
Bc024814	0	-1.2	-1.9	-1.5	*Hmgb2*	0	2.2	3.1	2.4
Bc031181	0	-0.8	-1.6	-1.3	*Hnrpa1*	0	1.1	1.6	0.9
Bc036718	0	-1.7	-2.4	-2.2	*Hnrpa2b1*	0	1.3	1.6	1.2
Bc056929	0	-1.2	-5.3	-3.9	*Irf1*	0	1.3	1.8	1.7
Bicc1	0	-0.8	-1.7	-1.4	*Irs2*	0	2.6	3.5	3.5
Bicd2	0	-0.3	-1.4	-1.3	*Kdelr2*	0	1.5	2.1	1.6
Blcap	0	-1.0	-1.6	-1.5	*Klra18*	0	2.2	3.5	3.3
Bok	0	-1.0	-3.1	-2.8	*Klra33*	0	1.7	3.3	2.8
Btbd6	0	-1.2	-2.1	-1.6	*Klra4*	0	2.3	3.6	3.4
Bzrap1	0	-2.5	-3.8	-4.3	*Kpna2*	0	1.6	1.9	1.3
Camk2n1	0	-1.5	-2.9	-2.8	*Loc381795*	0	3.0	3.5	3.2
Carhsp1	0	-0.3	-1.7	-1.5	*Lsm2*	0	1.3	1.5	1.2
Ccbl1	0	-1.9	-2.4	-2.1	*Ly6a*	0	2.3	4.1	3.9
Cd59a	0	-1.2	-3.6	-3.3	*Ly6e*	0	1.1	1.8	1.6
Cd97	0	-2.0	-3.2	-3.5	*Mcm5*	0	1.1	1.4	1.2
Cdc23	0	-1.1	-1.8	-1.7	*Mettl9*	0	1.1	1.4	1.3
Cdc42bpb	0	-1.2	-1.9	-1.5	*Ncapd2*	0	1.5	1.8	1.6
Ceecam1	0	-1.8	-2.6	-2.5	*Nsbp1*	0	1.1	1.5	1.3
Cib2	0	-0.8	-2.3	-1.9	*Odz3*	0	2.0	3.0	3.0
Ckb	0	-5.5	-8.5	-8.4	*Otx1*	0	2.3	2.8	2.5
Cln2	0	-1.3	-2.1	-1.7	*Pak3*	0	1.2	1.4	1.1

Table 4. Differentially expressed genes during tumorigenesis (continued)

(1) genes=181 ($2e^{-129}$)	170	245	450	528
Col4a1	0	-1.7	-6.7	-5.1
Col4a2	0	-1.4	-5.5	-4.1
Col4a5	0	-1.0	-1.6	-1.2
Creb3	0	-0.7	-1.4	-1.3
Creld1	0	-0.8	-1.8	-1.6
Crim1	0	-0.6	-1.3	-1.4
Crip2	0	-2.0	-4.8	-4.3
Csrp1	0	-1.5	-2.2	-2.3
D14ertd449e	0	-1.3	-2.4	-2.6
D430039n05RIK	0	-0.7	-2.1	-1.9
D4bwg0951e	0	-0.8	-1.3	-1.4
D630003m21RIK	0	-1.0	-3.0	-2.4
D930001i22RIK	0	-0.9	-1.7	-1.5
Dcp1b	0	-1.0	-2.1	-1.9
Ddah1	0	-0.9	-1.9	-2.0
Dnaja4	0	-1.1	-1.6	-1.7
Dos	0	-2.0	-2.9	-2.7
Dtnb	0	-1.0	-2.0	-1.5
Dusp2	0	-1.1	-2.5	-2.6
Dyrk3	0	-1.4	-2.2	-2.4
Edn1	0	-0.5	-1.7	-1.5
Efna5	0	-0.6	-1.7	-1.6
Eno3	0	-1.4	-2.8	-2.5
Enpp5	0	-0.8	-2.4	-1.6
Entpd4	0	-0.7	-1.4	-1.3
Eppk1	0	-0.9	-2.2	-2.3
Etfb	0	-1.1	-1.5	-1.4
Fblim1	0	-1.1	-2.4	-1.9
Fos	0	-0.4	-1.5	-1.5
Fst	0	-3.1	-6.1	-6.0
Fxyd5	0	-1.3	-2.8	-2.3
Galntl4	0	-2.7	-4.2	-3.6
Garnl4	0	-0.8	-2.0	-2.1
Gdf15	0	-3.0	-6.1	-4.6
Ghr	0	-0.5	-1.4	-1.4
Gnaz	0	-1.2	-2.2	-1.8
Golph2	0	-0.6	-1.7	-1.7
Gstk1	0	-2.1	-4.8	-4.1
H13	0	-0.9	-1.4	-1.2
Havcr1	0	-1.2	-3.4	-3.2
Hemk1	0	-1.6	-2.2	-1.8
Hint2	0	-1.2	-2.2	-1.6
Hist1h1c	0	-1.9	-3.5	-3.4
Hist1h2bc	0	-1.5	-3.3	-3.4
Hist1h2bj	0	-1.1	-1.8	-2.0
Hist2h2aa1	0	-2.1	-3.3	-2.8
Hk2	0	-0.5	-1.7	-1.5
Hr	0	-1.0	-3.3	-3.0
Hyal1	0	-2.0	-3.1	-2.5
Igfbp2	0	-1.3	-3.5	-3.0
Igsf4a	0	-1.1	-2.9	-2.7
Il11	0	-3.1	-4.1	-3.7
Impact	0	-1.1	-2.4	-1.8
Krt1-18	0	-1.5	-2.8	-1.6
Lims2	0	-2.6	-4.5	-4.2
Llglh2	0	-2.0	-4.1	-3.7
Loc239102	0	-1.1	-1.9	-1.7

(2) genes=78($5e^{-39}$)	170	245	450	528
Pfn2	0	1.1	1.5	1.4
Prdx4	0	1.7	1.9	1.9
Rab32	0	1.0	1.4	1.4
Rbm3	0	0.9	1.5	0.8
Rfc3	0	0.9	1.5	0.9
Rfc5	0	0.8	1.4	1.1
Ris2	0	0.9	1.6	1.3
Skp2	0	0.8	1.4	1.0
Slc29a1	0	1.3	1.5	1.5
Smoc1	0	3.6	4.9	4.1
Snx8	0	1.0	1.5	1.5
Spc25	0	1.4	1.5	1.2
Sprr2g	0	2.2	4.2	3.6
Ssbp2	0	0.9	1.7	1.5
Thoc4	0	1.3	1.6	1.4
Tk1	0	1.4	1.6	1.2
Twist1	0	1.2	1.8	1.7
Tyms	0	1.7	2.1	1.5
Tyms-ps	0	1.4	1.7	1.4
Uaca	0	0.7	1.4	1.0
Zfp326	0	1.2	1.6	1.3

(3) genes=65 ($5e^{-17}$)	170	245	450	528
1190002h23RIK	0	0.4	1.4	1.0
1810014l12RIK	0	0.6	1.5	1.1
2610019i03RIK	0	1.0	2.2	2.2
5430420c16RIK	0	1.0	2.3	2.1
Aa467197	0	0.1	4.3	4.0
Acsl3	0	0.9	2.1	2.1
Adh7	0	2.7	7.8	8.0
Ank	0	0.1	1.3	1.4
Aqp5	0	1.3	4.1	4.3
Arhgdib	0	1.0	5.9	5.3
Avpi1	0	-0.3	1.4	1.5
Bc029169	0	1.2	4.1	4.0
Bgn	0	0.4	4.2	4.3
Ccl25	0	0.7	1.4	1.3
Col1a1	0	1.1	7.6	7.4
Col3a1	0	0.3	4.9	4.4
Col6a1	0	0.7	2.3	2.3
Col6a2	0	0.3	2.9	2.8
Cxcl12	0	2.1	5.0	4.7
Cyp51	0	1.0	2.6	2.5
Dap	0	0.3	1.8	2.0
E030003n15RIK	0	0.2	1.5	1.7
Elovl6	0	0.6	1.5	1.5
Emilin1	0	0.9	3.4	3.3
Esm1	0	1.4	3.9	4.0
Fdps	0	0.0	1.5	1.8
Fgf7	0	-0.2	3.7	3.2
Fmnl3	0	0.6	1.8	1.7
Foxq1	0	0.2	4.3	3.9
G431001e03RIK	0	0.8	1.9	2.0
Hmgcs1	0	0.3	1.4	1.3
Igf2bp3	0	1.6	4.7	4.4
Lsp1	0	1.1	2.9	2.7
Ly6c	0	0.9	3.8	3.4

Table 4. Differentially expressed genes during tumorigenesis (continued)

(1) genes=181 ($2e^{-129}$)	170	245	450	528
Loc381297	0	-0.8	-2.0	-2.1
Loc386486	0	-0.8	-1.4	-1.5
Lrrc28	0	-0.9	-1.5	-1.4
Lzf	0	-1.0	-1.6	-1.5
Man2b2	0	-1.2	-1.8	-1.4
Mocos	0	-1.4	-2.4	-1.8
Mscp	0	-1.1	-1.4	-1.3
Mt1	0	-1.3	-3.6	-2.6
Mvp	0	-0.9	-1.7	-1.3
Ndfip2	0	-0.9	-1.5	-1.4
Nek9	0	-0.7	-1.5	-1.1
Nipsnap1	0	-0.6	-1.7	-1.6
Nme4	0	-1.1	-2.5	-2.5
Nme5	0	-0.8	-4.0	-3.7
Nppa	0	-2.1	-3.6	-3.9
Nppb	0	-3.1	-5.8	-6.0
Nt5e	0	-0.5	-1.6	-1.1
Nudt7	0	-0.9	-2.2	-1.9
Ostm1	0	-0.9	-1.6	-1.4
Oxr1	0	-0.6	-1.5	-1.0
Pde4dip	0	-2.1	-3.3	-3.6
Pea15	0	-0.5	-1.5	-1.4
Peg3	0	-1.5	-4.1	-3.8
Picalm	0	-0.7	-1.7	-1.3
Pld3	0	-1.3	-1.8	-1.4
Plekhb2	0	-1.0	-1.7	-1.7
Plekhc1	0	-0.9	-2.0	-1.6
Plekhg3	0	-1.1	-1.7	-1.9
Ppp1r9a	0	-0.8	-2.1	-1.6
Ppt1	0	-1.0	-1.5	-1.4
Prnp	0	-0.8	-1.6	-1.3
Ptpn21	0	-0.7	-1.4	-1.1
Ptprk	0	-0.6	-1.5	-1.3
Pwwp2	0	-0.5	-1.5	-1.3
Rab3d	0	-1.0	-2.1	-1.9
Rab3ip	0	-1.0	-1.7	-1.6
Rab7l1	0	-0.5	-1.6	-1.5
Rbpms	0	-0.6	-1.7	-1.2
Rhob	0	-0.9	-2.2	-2.0
Rin3	0	-0.4	-1.5	-1.3
Rras	0	-0.6	-1.5	-1.4
Rras2	0	-1.1	-2.0	-1.9
Rusc2	0	-1.0	-1.7	-1.5
S100a1	0	-1.2	-2.3	-1.4
Scin	0	-3.3	-5.9	-6.3
Slc7a4	0	-2.1	-3.4	-3.3
Slc7a7	0	-1.6	-2.5	-2.5
Slco2a1	0	-1.7	-3.6	-2.8
Sptlc1	0	-0.9	-1.4	-1.4
Srd5a1	0	-1.3	-2.5	-2.6
Tagln	0	-2.3	-5.6	-5.8
Taldo1	0	-1.1	-1.5	-1.2
Tcf2	0	-0.5	-1.4	-1.3
Tgfb1i1	0	-1.4	-4.1	-4.2
Tgfb1i4	0	-1.2	-2.4	-2.6
Tgfb3	0	-0.7	-2.0	-1.6
Timp3	0	-1.5	-2.8	-2.2

(3) genes=65 ($5e^{-17}$)	170	245	450	528
Tinagl	0	-1.2	-3.8	-3.4
Tnc	0	-0.7	-1.6	-1.6
Mglap	0	1.2	9.3	9.6
Mgst1	0	0.7	2.1	2.0
Mgst2	0	-0.4	4.0	4.3
Mnd1	0	0.7	1.5	1.4
Mvd	0	0.1	2.0	2.0
Mxra8	0	-0.3	3.0	3.4
Nr2f1	0	0.2	5.5	5.9
Nrn1	0	1.7	4.0	3.9
Nupr1	0	0.4	1.6	1.3
Ottmusg00000010673	0	1.4	5.0	5.0
Padi1	0	0.2	4.7	4.9
Pcolce	0	1.1	4.2	4.1
Pdgfra	0	0.4	5.4	6.2
Pem	0	0.2	3.3	2.8
Pmvk	0	0.7	1.9	2.0
Ppbp	0	1.4	4.6	4.3
Ptn	0	1.2	3.6	2.5
Rcn3	0	1.1	2.4	2.2
S100a4	0	0.6	3.1	3.7
Scd2	0	1.0	2.7	2.6
Sec24d	0	0.7	1.6	1.6
Sfrp1	0	0.4	2.9	3.0
Siat4a	0	0.3	1.2	1.3
Slc1a3	0	1.6	4.1	4.0
Sqle	0	0.7	2.5	2.4
Stard4	0	0.6	1.5	1.4
Tnfrsf11b	0	1.2	3.0	2.1
Twist2	0	2.0	5.0	4.8
Ugt1a10	0	0.3	2.3	2.8
Vamp5	0	0.6	1.5	1.1
Zfp537	0	0.8	2.3	2.3

(1) continued...

(1) genes=181 ($2e^{-129}$)	170	245	450	528
Tnfaip3	0	-0.6	-2.3	-2.1
Tspyl3	0	-2.2	-3.2	-3.4
Ttc5	0	-0.9	-1.5	-1.2
Unc13b	0	-2.5	-4.6	-3.5
Zdhhc14	0	-1.3	-3.2	-3.2
Zfhx2	0	-1.1	-2.0	-1.7
Zfp219	0	-0.7	-1.8	-1.4
Zfp354a	0	-0.8	-2.1	-1.8

Table 4. Differentially Expressed Genes During Tumorigenesis (continued)

(4) Genes=58(1E-25)	170	245	450	528
0610041g09RIK	0	-0.4	-1.0	-1.3
1300014i06RIK	0	-0.4	-1.4	-1.5
1500031h04RIK	0	-0.7	-2.4	-3.5
2010004a03RIK	0	-1.0	-1.7	-2.2
2310067e08RIK	0	-0.6	-1.9	-2.2
Acta2	0	-1.3	-2.6	-3.2
Actn1	0	-0.8	-1.6	-2.1
Agrn	0	-0.7	-2.3	-2.6
Akr1c19	0	-1.5	-3.2	-4.5
Arhgdig	0	-0.9	-3.0	-3.5
Bc003236	0	-0.3	-1.2	-1.4
Bc011487	0	-0.7	-3.5	-4.5
Bc058638	0	-1.2	-3.5	-4.0
Bin1	0	-0.8	-1.3	-1.5
Cd44	0	-0.9	-1.5	-1.8
Clcf1	0	-0.9	-1.7	-1.9
Cltb	0	-0.6	-1.1	-1.6
Col4a3	0	-1.0	-2.8	-3.9
Col4a4	0	-1.2	-2.3	-3.0
D330037a14RIK	0	-0.5	-1.1	-1.4
Dbn1	0	-0.7	-1.2	-1.6
Dtr	0	-0.7	-1.8	-2.6
Dusp1	0	-0.5	-1.5	-1.7
Fbln2	0	-0.6	-1.8	-2.7
Fhl1	0	-0.6	-1.4	-2.1
Flrt3	0	-0.9	-2.7	-3.2
Gjb4	0	-0.3	-1.5	-2.0
Hes1	0	-0.1	-1.0	-1.7
Hist1h2bf	0	-1.0	-1.7	-2.1
Hist1h2bh	0	-1.1	-1.7	-2.0
Hist1h2bk	0	-0.4	-0.8	-1.3
Hist1h2bm	0	-1.1	-1.8	-2.1
Hist1h2bn	0	-0.7	-1.2	-1.5
Igsf9	0	-0.6	-1.7	-2.3
Inhba	0	-1.2	-1.8	-3.0
Klf5	0	-0.8	-1.4	-1.8
Lad1	0	-0.7	-2.5	-2.6
Lamb3	0	-1.5	-3.0	-4.1
Mcam	0	-0.7	-2.6	-3.3
Mett11d1	0	-0.6	-1.3	-1.4
Nes	0	-0.9	-1.3	-1.9
Ngfb	0	-0.8	-2.7	-2.9
Panx1	0	-0.3	-0.9	-2.0
Pbp2	0	-0.3	-1.0	-1.7
Pdlim7	0	-0.7	-1.1	-1.8
Ptprb	0	-1.5	-2.7	-3.8
Raet1b	0	-0.9	-1.7	-1.9
Reprimo	0	-1.7	-4.1	-4.7
S100a15	0	-1.1	-2.2	-2.7
Scamp5	0	-0.9	-1.3	-2.1
Sdc3	0	-0.2	-1.5	-3.0
Slc7a3	0	-1.7	-2.5	-3.2
Snrpn	0	-1.3	-3.0	-4.1
Taf9b	0	-0.6	-2.3	-3.1
Ugcg	0	-0.5	-1.9	-2.3
Wnt7a	0	-2.3	-5.4	-6.2
Wnt7b	0	-1.3	-5.7	-6.1

(5) Genes=53(9E-22)	170	245	450	528
1190005i06RIK	0	0.9	2.0	2.2
2310022b05RIK	0	1.0	1.9	1.9
2610027c15RIK	0	0.5	1.0	1.3
2810428i15RIK	0	0.9	1.6	1.7
2810471m23RIK	0	1.3	2.0	3.1
4930504e06RIK	0	0.1	1.0	1.4
9230117n10RIK	0	0.5	1.4	2.4
Acat2	0	0.2	1.3	1.6
Adamts2	0	0.3	3.1	4.1
Ai467484	0	0.9	1.5	2.2
Aldh3a1	0	0.4	3.8	5.0
Antxr1	0	0.9	2.3	3.1
Apbb2	0	0.8	1.5	1.5
Arfgap3	0	0.4	1.0	1.4
Bteb1	0	0.5	1.1	1.6
C3	0	0.6	2.2	4.8
Cacna2d1	0	0.3	1.1	1.5
Card4	0	0.4	1.4	1.7
Casp4	0	1.3	2.6	3.0
Cd14	0	0.5	0.9	1.6
Chst2	0	0.9	1.7	1.7
Efna1	0	0.5	1.5	1.9
Eif4b	0	0.5	1.2	1.4
Emb	0	0.7	1.3	1.9
Fbxo31	0	0.6	1.4	1.6
Figf	0	2.0	4.8	5.3
Gypc	0	0.7	1.4	1.5
Il13ra1	0	0.6	1.5	1.9
Lss	0	0.6	1.7	2.0
Man2a1	0	1.1	1.7	1.9
Mknk2	0	0.8	0.9	1.5
Msln	0	0.4	1.5	2.3
Ntrk3	0	1.6	2.1	3.2
Olfml2b	0	1.2	2.2	2.2
Olfml3	0	0.4	2.8	4.0
Pde1a	0	0.3	3.1	4.2
Pdk1	0	0.3	0.8	1.7
Phka2	0	0.7	1.1	1.4
Prrx1	0	1.4	2.5	3.2
Rev3l	0	0.4	0.8	1.4
Rgl1	0	0.3	1.3	1.9
Sc4mol	0	0.1	1.2	1.5
Scara3	0	1.8	3.3	3.5
Scara5	0	0.7	2.9	4.6
Sdh1	0	0.4	0.9	1.7
Sesn1	0	0.3	1.0	1.9
Socs3	0	0.7	1.7	2.3
Spp1	0	0.6	1.4	1.7
Stard5	0	0.6	0.8	1.6
Tgfbi	0	1.9	3.7	4.6
Tgfbr3	0	0.7	1.5	2.2
Tnfaip8	0	0.8	1.4	1.6
Vdr	0	0.4	1.3	2.1

Table 4. Differentially Expressed Genes During Tumorigenesis (continued)

(6) genes=36(9e-4)	170	245.0	450.0	528.0
0610006i08RIK	0	-0.4	-2.2	-2.4
1110032e23RIK	0	0.4	-0.8	-1.0
2610001e17RIK	0	0.1	-1.5	-1.2
4930422j18RIK	0	0.0	-1.5	-1.7
A230050p20RIK	0	0.8	-1.4	-2.0
AI024069	0	-0.6	-3.8	-3.5
Axud1	0	0.4	-0.5	-0.9
Bdnf	0	-0.1	-1.8	-1.5
Card10	0	-0.3	-1.4	-1.4
Cav1	0	0.2	-0.9	-1.4
Chrnb1	0	0.0	-0.9	-1.4
Clca4	0	0.0	-2.6	-3.5
Ctsw	0	-0.1	-2.0	-3.2
F2rl1	0	-0.5	-2.3	-2.6
Gadd45g	0	-0.2	-1.9	-2.5
Greb1	0	-0.2	-1.5	-1.3
Hspb1	0	-0.4	-1.8	-1.8
Nbl1	0	1.0	-1.2	-1.5
Ndn	0	-1.9	-3.6	-6.3
Nox4	0	0.3	-1.0	-1.4
Pdgfb	0	0.3	-2.3	-2.3
Pdlim4	0	-0.2	-1.4	-1.6
Plk2	0	-0.3	-1.5	-1.5
Pvr	0	0.1	-1.3	-1.0
Pvrl2	0	0.2	-2.0	-1.7
Rasgrp3	0	-0.6	-2.6	-2.7
Rasl12	0	-0.5	-3.4	-3.4
Rassf1	0	0.5	-0.8	-1.1
Rtn1	0	0.1	-3.2	-3.1
S100a3	0	-0.1	-1.2	-1.4
Samd10	0	0.0	-1.8	-1.6
Sema4f	0	-0.6	-4.9	-5.8
Slc39a6	0	0.0	-1.3	-1.4
Smtn	0	0.2	-1.4	-1.5
Sox8	0	0.3	-1.4	-1.3
Tuft1	0	0.1	-1.7	-2.0
Wfs1	0	-0.1	-1.9	-1.5

(7) genes=34(6e-4)	170	245	450	528
1110012d08RIK	0	-1.2	-1.4	-0.8
1500005a01RIK	0	-1.6	-2.0	-1.6
2210008i11RIK	0	-1.1	-1.4	-1.0
2310005e10RIK	0	-1.1	-1.5	-1.0
2610020h15RIK	0	-1.1	-1.5	-1.2
Abhd4	0	-1.6	-2.4	-1.7
Aldh4a1	0	-1.6	-2.2	-1.4
Aldh6a1	0	-1.8	-2.3	-0.9
Atp6v1d	0	-1.1	-1.6	-1.0
Bc031853	0	-1.2	-1.8	-1.2
Ccnd2	0	-1.0	-1.8	-1.0
Ccng1	0	-2.3	-2.7	-2.4
Clu	0	-2.3	-3.3	-2.1
Cox6a2	0	-1.7	-2.2	-1.6
Cyp2d22	0	-1.8	-2.9	-1.3
Cyp4f13	0	-1.7	-2.2	-1.9
Dgka	0	-2.3	-2.6	-2.3
Ganc	0	-1.4	-2.0	-1.5
Ghitm	0	-1.3	-1.8	-1.2
Gm2a	0	-1.2	-2.0	-1.3
Gns	0	-1.0	-1.5	-0.9
Grcc10	0	-0.9	-1.5	-0.8
H2-dmb1	0	-0.8	-1.4	-0.7
Itm2b	0	-1.2	-1.7	-1.2
Lbh	0	-1.4	-2.4	-1.4
Mgc18837	0	-1.9	-2.3	-1.7
Rpl22	0	-1.4	-1.6	-1.1
Sdsl	0	-2.9	-3.5	-3.1
Slc9a3r2	0	-1.2	-1.6	-1.0
Tgm2	0	-1.0	-1.7	-0.8
Tmem141	0	-1.6	-2.0	-1.5
Trp53inp1	0	-3.3	-3.7	-3.1
Trp53inp2	0	-1.1	-1.4	-0.7
Wars	0	-1.3	-1.7	-1.4

(8) genes =25 (9e-4)	170	245	450	528
1810009n02RIK	0	1.9	2.6	2.7
2610002m06RIK	0	1.0	1.6	1.6
Abcb1b	0	1.5	1.8	2.0
Car13	0	1.0	1.3	1.6
Ccl9	0	2.8	3.7	4.3
Cdca3	0	1.7	2.0	2.1
Cdkn2c	0	1.4	1.9	2.0
Cox7b	0	1.0	1.2	1.4
Cxcl1	0	2.4	3.8	3.9
Ednra	0	3.3	4.7	5.2
Frda	0	1.2	1.5	1.7
Ifitm3	0	1.1	1.6	1.7
Lsm6	0	1.0	1.5	1.5
Mad2l1	0	1.3	1.4	1.5
Mcrs1	0	0.8	1.1	1.4
Ogn	0	3.0	4.3	5.2
Pdk3	0	1.4	1.8	1.8
Pold2	0	1.3	1.6	1.6

(8) continued...

(8) genes =25 (9e-4)	170	245	450	528
Rfxap	0	1.1	1.4	1.4
Rgs10	0	1.8	2.2	2.7
Slco1a5	0	2.0	2.6	3.3
Slit2	0	1.3	1.8	2.2
Smpdl3b	0	1.6	2.2	2.6
Snx25	0	1.1	1.4	1.6
Vcam1	0	0.8	1.0	1.3

Table 4. Differentially Expressed Genes During Tumorigenesis (continued)

(Ns) genes=68	170	245	450	528	(ns) genes=68 (ns)	170	245	450	528
1110058a15RIK	0	-4.8	-5.6	-5.8	Gstm2	0	-0.3	2.0	3.0
1700024k14RIK	0	-0.7	0.9	2.9	Igfbp4	0	-0.6	2.3	3.4
2010323f13RIK	0	-1.0	1.1	2.0	Inmt	0	1.2	1.7	5.1
4432405b04RIK	0	-0.6	-1.3	-0.8	Jam4	0	-1.3	-1.7	-2.2
5430435g22RIK	0	-0.7	1.5	1.6	Khdrbs3	0	0.5	1.5	0.7
Ai429612	0	-0.7	-1.6	-0.9	Kif3c	0	-1.0	-1.4	-1.4
Aqp1	0	0.9	1.5	3.2	Lrrfip1	0	-1.7	-2.2	-2.5
Arl4c	0	0.9	0.7	-0.5	Man2c1	0	-1.2	-1.7	-2.0
Arrdc3	0	-0.8	0.1	0.9	Manea	0	0.0	-1.3	-0.8
Bc013481	0	-0.4	1.3	1.7	Mgat3	0	0.1	1.0	1.9
Cd248	0	-0.8	1.3	1.5	Mmp2	0	-0.7	2.3	3.0
Cdkn1a	0	-5.1	-6.0	-5.5	Msc	0	2.6	-0.6	-2.4
Cds1	0	0.6	-1.9	-1.5	Mybl2	0	1.4	1.3	0.7
Cebpb	0	0.0	0.7	1.9	Myd116	0	-0.4	-1.9	-1.3
Cmtm8	0	-0.5	-3.1	-2.0	Nnmt	0	-0.1	-1.9	-1.0
Cnnm2	0	-0.5	0.4	1.5	Palmd	0	3.0	2.4	1.1
Cobl	0	-0.3	-1.7	-0.7	Prelp	0	-0.5	0.8	3.2
Cp	0	0.9	1.1	2.3	Prickle1	0	0.4	-0.4	-1.1
Csrp2	0	0.8	1.9	1.2	Rad52b	0	-0.6	-1.4	-0.3
Ctgf	0	-0.4	-0.5	-1.5	Rasl11b	0	0.5	-3.3	-2.5
Ctsh	0	-0.2	-3.4	-1.0	Rem2	0	-3.5	-4.3	-4.3
Cx3cl1	0	-0.5	1.3	1.4	Rgs17	0	-1.2	0.6	1.4
Cxadr	0	0.0	-1.6	-1.0	Sars1	0	0.0	-1.5	-1.1
D19wsu12e	0	0.0	0.9	1.5	Serpinf1	0	-0.6	1.6	2.9
D330024h06RIK	0	-0.6	0.1	1.0	Stxbp2	0	-1.3	-1.9	-2.1
D430044g18RIK	0	-0.6	-2.1	-1.3	Sytl2	0	-0.5	-1.9	-1.3
Enpp2	0	0.8	4.7	2.6	Tacstd2	0	0.6	-2.3	-1.7
Epb4.1l3	0	0.6	-0.4	-0.8	Tcn2	0	-2.2	-2.8	-2.8
Fbxo2	0	-0.8	0.1	1.0	Tm4sf3	0	0.4	2.4	5.2
Gjb3	0	1.1	-0.4	-1.9	Tm4sf6	0	0.8	-0.1	-0.8
Gnb5	0	-1.1	-1.4	-1.4	Trib3	0	-1.1	-1.7	0.2
Gpx3	0	-1.5	-3.7	-2.1	Tspan7	0	1.2	1.6	0.7
Gstm1	0	-1.1	-0.5	0.6	Uchl1	0	1.1	-0.1	-1.0
					Whrn	0	0.0	-1.6	-1.0

Table 4. Gene symbols, fold change, and profile assignment for each of four key transformative stages during tumorigenesis. Differentially expressed genes over time were assigned to each profile by maximization of the correlation coefficient between the actual and randomly generated model profiles. Fifty theoretical profiles were tested with a maximum correlation of 0.9 among profiles and a maximum allowable change of 5 units. The statistical significance of the number of assigned versus expected genes for each profile is shown and profiles were assigned by relative expression for the list of differentially expressed genes from time course gene discovery analysis. Genes were included in a given profile where $p < 0.05$.

3.7.2 *Traditional gene analysis by paired comparisons of consecutive transformative stages*

In addition to the time course analysis for microarray data, pair wise analyses using Significance Analysis of Microarrays (SAM) on consecutive samples were performed in biological triplicate (Figure 10). Gene lists, definition, location, and fold change are provided for each paired analysis (Table 5). The largest list of differentially expressed genes was generated between Early=170:245= 319 genes. Despite using the same statistical cut-points and fold change of 2.5 for all pair wise analysis, the lists of differentially expressed genes became progressively shorter as transformation progressed, with Mid=245:450=102 genes and Late=450:528= 30 genes. Gene lists between consecutive transformative stages were similarly subjected to GO enrichment analysis to determine the functional significance of the differentially expressed genes (Table 6). Several significant categories were returned for functional enrichment at between each pair of stages. Of note are several categories relating to extracellular matrix and collagen components, as well as several relating to cell cycle, mitosis, and development were enriched for altered gene expression (Table 6).

Table 5. Differentially expressed genes between consecutive transformative stages

170:245 gene	definition	cytoband	fold change
1110031b06RIK	synaptosomal-associated protein, 47	11qB5	-2.9
1110033j19RIK	ribosomal protein S4, Y-linked 2	10qC1	-4.3
1110036o03RIK	RIKEN cDNA 1110036O03 gene	11qB5	-3.5
1110058a15RIK	late cornified envelope 1G	11qC	-109.3
1700088e04RIK	RIKEN cDNA 1700088E04 gene	15qE3	-2.5
1810009n02RIK	RIKEN cDNA 1810009N02 gene	1qC3	4.2
1810015a11RIK	YdjC homolog	5qF	-6.8
1810054o13RIK	transmembrane protein 86A	2qE3	12.8
1810073e21RIK	trafficking protein particle complex 6A	3qE3	-5.0
2300002d11RIK	TMF1-regulated nuclear protein 1	10qD3	-5.4
2310061n23RIK	interferon, alpha-inducible protein 27 like 2A	10qA4	36.2
2410008j05RIK	transmembrane protein 121	1qD	4.2
2810003c17RIK	allograft inflammatory factor 1-like	8qB3.3	-3.0
2810004a10RIK	interleukin 17 receptor D	12qA1.1	-3.1
4833421e05RIK	isoamyl acetate-hydrolyzing esterase 1 homolog	12	-3.3
4930418p06RIK	rhomboid domain containing 1	13qA3.3	-3.3
4933405a16RIK	sphingomyelin synthase 2	3	-2.9
5830467p10RIK	fermitin family homolog 1	16qC4	-10.2
6330406i15RIK	RIKEN cDNA 6330406I15 gene	18qB3	7.8
9830002i17RIK	spinster homolog 3	12qC3	-3.0
Abhd4	abhydrolase domain containing 4	14	-3.2
Ai427138	frizzled homolog 5	4qD3	-3.9
Ai449441	PIF1 5'-to-3' DNA helicase homolog	8qA2	3.9
Ai450948	AHNAK nucleoprotein 2	11qB5	-5.0
Ak1	adenylate kinase 1	4qC7	-19.1
Aldh1a1	aldehyde dehydrogenase family 1, subfamily A1	11qD	-5.3
Amid	apoptosis-inducing factor, mitochondrion-associated 2	7qA3	-2.6
Ankrd1	ankyrin repeat domain 1	1qC1.3	-3.3
Aplp1	amyloid beta precursor-like protein 1	11qB2	-4.7
Arhgap24	Rho GTPase activating protein 24	5qE5	4.4
Ass1	argininosuccinate synthetase 1	1qA5	-47.0
Atp6v0a1	ATPase, H+ transporting, lysosomal V0 subunit A1	5qG2	-3.9
B930041f14RIK	RIKEN cDNA B930041F14 gene	6qB1	-7.1
Bc022687	cDNA sequence BC022687	9qF1	-2.9
Bc024814	cDNA sequence BC024814	7qB2	-2.6
Bc036718	nudix -type motif 18	13qB3	-3.4
Bc046331	cDNA sequence BC046331	4qE2	3.0
Bc049806	family with sequence similarity 126, m B	7qF5	-3.0
Birc5	baculoviral IAP repeat-containing 5	11	4.1
Bmp3	bone morphogenetic protein 3	XqC3	17.9
Bscl2	Bernardinelli-Seip congenital lipodystrophy 2 homolog	9qA5.2	-3.3
Btbd3	BTB domain containing 3	XqA3.3	5.0
Bub1b	budding uninhibited by benzimidazoles 1 homolog, beta	8qB3.1	3.1
Bzrap1	benzodiazapine receptor associated protein 1	17qA3.3	-9.1

Table 5. Differentially expressed genes between consecutive transformative stages

170:245 gene	definition	cytoband	fold change
C630013n10RIK	kelch-like 26	10qC2	-2.6
Cask	calcium/calmodulin-dependent serine protein kinase	XqA1.1	3.1
Ccbl1	cysteine conjugate-beta lyase 1	8qE1	-4.3
Ccnb1	cyclin B1	13qD1	2.9
Ccnd1	cyclin D1	5qF	-3.7
Ccng1	cyclin G1	XqA5	-5.2
Cd97	CD97 antigen	4qA5	-4.5
Cdc20	cell division cycle 20 homolog	8qC5	3.8
Cdca3	cell division cycle associated 3	6qF2	3.2
Cdca8	cell division cycle associated 8	4qD2.2	4.6
Cdkn1a	cyclin-dependent kinase inhibitor 1A	11qB3	-41.7
Cdsn	corneodesmosin	14qC3	-4.2
Ceecam1	cerebral endothelial cell adhesion molecule	9qF4	-3.9
Cenpa	centromere protein A	5	7.2
Cenpi	centromere protein I	13qB3	4.2
Ckb	creatine kinase, brain	19qA	-49.3
Cldn4	claudin 4	5	-75.3
Cln2	tripeptidyl peptidase I	7	-2.7
Clu	clusterin	8qD3	-5.3
Col18a1	collagen, type XVIII, alpha 1	11qB3	-3.0
Col4a1	collagen, type IV, alpha 1	2qE5	-3.3
Crip2	cysteine rich protein 2	15qA1	-4.1
Csrp1	cysteine and glycine-rich protein 1	1	-2.9
Cxcl1	chemokine ligand 1	5qE1	5.6
Cyp2d22	cytochrome P450, family 2, subfamily d, polypeptide 22	15	-4.3
Cyp4a12	cytochrome P450, family 4, subfamily a, polypeptide 12a	6qE3	-71.4
Cyp4f13	cytochrome P450, family 4, subfamily f, polypeptide 13	2qE1	-3.7
Cyp7b1	cytochrome P450, family 7, subfamily b, polypeptide 1	19qC3	10.5
D11ertd18e	solute carrier family 46, m 1	2qB	-7.3
Dab2	disabled homolog 2	12qD2	4.8
Dbf4	DBF4 homolog	11qC	4.1
Dcn	decorin	10qC3	13.2
Dcxr	dicarbonyl L-xylulose reductase	13qA5	-14.1
Ddit4	DNA-damage-inducible transcript 4	1qC1.1	-3.6
Dgka	diacylglycerol kinase, alpha	2qE3	-6.7
Dos	downstream of Stk11	10qD3	-4.6
Dusp6	dual specificity phosphatase 6	14qD3	-3.9
Dyrk3	dual-specificity tyrosine--phosphorylation regulated kinase 3	4qB3	-2.7
E130306d19RIK	RIKEN cDNA E130306D19 gene	4	3.6
Ednra	endothelin receptor type B	7qB1	24.0
Eno3	enolase 3, beta muscle	5qF	-3.0
Ephx1	epoxide hydrolase 1, microsomal	4qA3	-4.4
Ercc5	excision repair cross-complementing rodent repair deficiency,	10qC1	-3.2
Etv4	ets variant gene 4	11qB4	-3.4
Fez1	fasciculation and elongation protein zeta 1	2qA1	-4.6

Table 5. Differentially expressed genes between consecutive transformative stages

170:245 gene	definition	cytoband	fold change
Fibp	fibroblast growth factor intracellular binding protein	2qH1	-5.5
Fignl1	fidgetin-like 1	11qA1	2.8
Foxd1	forkhead box D1	13	6.9
Foxm1	forkhead box M1	6qF3	3.5
Fst	follistatin	13	-8.8
Galntl4	polypeptide N-acetylgalactosaminyltransferase-like 4	10qC1	-8.1
Ganc	glucosidase, alpha; neutral C	2qH3	-3.1
Gcnt2	glucosaminyl transferase 2, I-branching enzyme	17qA3.3	-2.8
Gdf15	growth differentiation factor 15	15qD1	-8.2
Grem1	gremlin 1	7qB1	-18.5
Gspt2	G1 to S phase transition 2	12qC1	-32.9
Gsta4	glutathione S-transferase, mu 2	18qE3	-6.6
H2afz	H2A histone family, m Z	3qG3	2.6
Hes6	hairy and enhancer of split 6	1	-3.3
Hist1h1c	histone cluster 1, H1d	13	-3.6
Hist1h2ad	histone cluster 1, H2ad	13qA3.1	3.5
Hist1h2ag	histone cluster 1, H2ag	13	4.2
Hist1h2bc	histone cluster 1, H2bc	11qB3	-3.1
Hist1h3a	histone cluster 1, H3a	5qB1	6.0
Hist1h3d	histone cluster 1, H2ad	11qB1.3	3.9
Hist1h3e	histone cluster 1, H3e	13qA3.1	4.6
Hist1h4f	histone cluster 1, H4d	13qA3.1	3.3
Hist2h2aa1	histone cluster 2, H2aa1	17qA3.3	-6.1
Hist2h2ab	histone cluster 2, H2ab	11qE2	3.9
Hmgb2	high mobility group box 2	7qF3	4.6
Hmgn2	high mobility group nucleosomal binding domain 2	4qD3	3.0
Hmox1	heme oxygenase 1	7qA3	-3.7
Hoxb7	homeobox B7	5qG2	-2.9
Hs3st3a1	heparan sulfate 3-O-sulfotransferase 3A1	5qB1	-3.2
Hsd3b7	hydroxy-delta-5-steroid dehydrogenase	11qD	-4.0
Htatip2	HIV-1 tat interactive protein 2, homolog	6qG1	-5.8
Idb1	inhibitor of DNA binding 1	17qB1	-2.6
Igf1	insulin-like growth factor 1	2qC1.1	9.4
Igf2bp2	insulin-like growth factor 2 mRNA binding protein 2	5qC3.1	-14.8
Igfbp2	insulin-like growth factor binding protein 2	1qG2	-2.6
Il11	interleukin 11	11qC	-14.3
Ilvbl	ilvB -like	12qA3	-3.1
Immp2l	IMP2 inner mitochondrial membrane peptidase-like	9qE3.3	-3.6
Ing1l	inhibitor of growth family, m 2	8B2	2.6
Iqgap3	IQ motif containing GTPase activating protein 3	19qA	4.6
Irs2	insulin receptor substrate 2	11qB3	12.5
Itga3	integrin alpha 3	4qE2	-3.4
Jam4	immunoglobulin superfamily, m 5	6qD1	-2.7
Kif2c	kinesin family m 2C	4qD1	3.1
Kif4	kinesin family m 4	6qE3	3.9

Table 5. Differentially expressed genes between consecutive transformative stages

170:245 gene	definition	cytoband	fold change
Klra18	killer cell lectin-like receptor, subfamily A, m 18	11qE1	4.5
Klra4	killer cell lectin-like receptor, subfamily A, m 4	6qF3	5.3
Kntc1	kinetochore associated 1	5qF	3.2
Krt1-18	keratin 18	15qE1	-3.0
Lbh	limb-bud and heart	13qD1	-2.8
Lims2	LIM and senescent cell antigen like domains 2	16qC3.3	-8.0
Lin54	lin-54 homolog	5qE4	3.1
Llglh2	Mus musculus lethal giant larvae homolog 2	11E2	-4.3
Loc381795		11qC	9.6
Lrrfip1	leucine rich repeat interacting protein 1	1	-3.5
Ly6a	lymphocyte antigen 6 complex, locus A	10qB4	4.6
Mapkapk3	mitogen-activated protein kinase-activated protein kinase 3	17qA2	-4.0
Mdm2	transformed mouse 3T3 cell double minute 2	2qA3	-7.1
Mgc25972		4qD1	-53.5
Mgmt	O-6-methylguanine-DNA methyltransferase	11qB1.3	-7.0
Mocos	molybdenum cofactor sulfurase	18	-3.1
Msc	musculin	1qA3	6.6
Myo7a	myosin VIIA	16qA1	3.2
Ncapd2	non-SMC condensin I complex, subunit D2	6qF3	2.8
Ndc80	NDC80 homolog, kinetochore complex component	18qD3	3.4
Ndg2	coiled-coil-helix-coiled-coil-helix domain containing 10	3qF1	-27.4
Ndn	necdin	15qD3	-3.9
Nid1	nidogen 1	13	-22.8
Nppa	natriuretic peptide precursor type A	4	-5.1
Nppb	natriuretic peptide precursor type B	11qC	-9.5
Oact1	membrane bound O-acyltransferase domain containing 1	13	3.3
Oasl2	2'-5' oligoadenylate synthetase-like 2	17qA3.3	17.7
Ogn	osteoglycin	15qB2	13.1
Otx1	orthodenticle homolog 1	9qA4	5.8
Palmd	palmdelphin	12qA1.1	18.1
Pbxip1	pre-B-cell leukemia transcription factor interacting protein 1	3qF1	-4.2
Pde4dip	phosphodiesterase 4D interacting protein	2qA3	-5.3
Phf17	PHD finger protein 17	11qB1.3	3.7
Phlda3	pleckstrin homology-like domain, family A, m 3	17qE1.1	-4.8
Pi4k2b	phosphatidylinositol 4-kinase type 2 beta	17qB1	3.2
Pkn2	protein kinase N2	3qH1	2.8
Plk1	polo-like kinase 1	3qH2	4.3
Plk4	polo-like kinase 4	3	3.5
Pole	polymerase , epsilon	5qF	3.4
Prc1	protein regulator of cytokinesis 1	7qD3	2.8
Prdx4	peroxiredoxin 4	XqF3	3.2
Prss19	kallikrein related-peptidase 8	16qB2	-37.5
Psen2	presenilin 2	6qB3	-3.7
Rab27a	RAB27A, m RAS oncogene family	14qD3	-8.7
Rab6b	RAB6B, m RAS oncogene family	XqA4	-4.5

Table 5. Differentially expressed genes between consecutive transformative stages

170:245 gene	definition	cytoband	fold change
Rasl11a	RAS-like, family 11, m A	5qG3	7.1
Rbmx	RNA binding motif protein, X chromosome	16qA3	3.4
Reprimo	reprimo, TP53 dependent G2 arrest mediator candidate	7qA1	-3.6
Rgs10	regulator of G-protein signalling 10	9qB	3.5
Rutbc1	small G protein signaling modulator 2	11	-6.0
Scin	scinderin	16qA1	-10.8
Sdsl	serine dehydratase-like	5qC3.1	-15.2
Serpinb6b	serine peptidase inhibitor, clade B, m 6b	4qB3	10.8
Slc19a2	solute carrier family 19 , m 2	1	-5.6
Slc24a3	solute carrier family 24 , m 3	2qG1	3.3
Slc25a1	solute carrier family 25 , m 1	11qA3.1	-2.5
Slc7a3	solute carrier family 7 , m 3	12qC2	-3.6
Slc7a4	solute carrier family 7 , m 4	19qB	-5.2
Slco1a5	solute carrier organic anion transporter family, m 1a5	6qG2	5.4
Slco2a1	solute carrier organic anion transporter family, m 2a1	7qF3	-3.5
Sprr2g	small proline-rich protein 2K	2qH1	5.1
Srpx2	sushi-repeat-containing protein, X-linked 2	7qB4	11.3
Ssx2ip	synovial sarcoma, X breakpoint 2 interacting protein	3qH2	2.7
Stk6	aurora kinase A	2	3.4
Stmn1	stathmin 1	4	3.4
Tagln	transgelin	9	-5.2
Timp3	tissue inhibitor of metalloproteinase 3	10	-3.0
Tle6	transducin-like enhancer of split 6, homolog of Drosophila E	11qA1	-4.6
Tmem141	transmembrane protein 141	1qG2	-3.3
Tmem150	transmembrane protein 150A	7qE3	-3.3
Tmem205	transmembrane protein 205	9A3	-3.8
Tmem53	transmembrane protein 53	11qD	-3.6
Tob1	transducer of ErbB-21	12qD2	-5.0
Trp53inp1	transformation related protein 53 inducible nuclear protein 1	5qG2	-20.1
Tspyl3	TSPY-like 3	XqA7.1	-6.0
Tyms	thymidylate synthase	5qB1	3.4
Tyms-ps	thymidylate synthase, pseudogene	10qC1	2.6
Unc13b	unc-13 homolog B	7qF1	-7.3
Vgf	VGF nerve growth factor inducible	6qE3	-87.4
Vkorc1l1	vitamin K epoxide reductase complex, subunit 1-like 1	5qG1.3	3.2
Vps25	vacuolar protein sorting 25	11qE2	-3.4
Wig1	zinc finger matrin type 3	3	-4.0
Wisp2	WNT1 inducible signaling pathway protein 2	15qE1	22.2
Wnt7a	wingless-related MMTV integration site 7A	6	-5.5
Zfp365	zinc finger protein 365	17qB1	-6.2

Table 5. Differentially expressed genes between consecutive transformative stages

245:450 gene	definition	location	change
0610006i08RIK	transmembrane protein 223	19qA	3.2
1200013b22RIK	NUAK family, SNF1-like kinase, 2	1qE4	2.9
2310047c17RIK	AHNAK nucleoprotein	19	2.8

Table 5. Differentially expressed genes between consecutive transformative stages

245:450 gene	definition	location	change
2310067e08RIK	endonuclease domain containing 1	9qA1	2.7
2610001e17RIK	coiled-coil domain containing 80	16	2.8
2810003c17RIK	allograft inflammatory factor 1-like	2qB	2.9
4930422j18RIK	splA/ryanodine receptor domain and SOCS box containing 1	4qE2	3.3
9130213b05RIK	prostate androgen-regulated mucin-like protein 1	5	6.1
A230050p20RIK	RIKEN cDNA A230050P20 gene	9qA3	5.8
Aa467197	expressed sequence AA467197	2qE5	-26.4
Adh7	alcohol dehydrogenase 7 , mu or sigma polypeptide	3qG3	-43.7
Agrn	agrin	4	3.4
Aqp5	aquaporin 5	15qF1	-8.7
Avpi1	arginine vasopressin-induced 1	19qC3	-4.0
Bc056929	doublecortin-like kinase 3	9qF3	18.9
Bok	BCL2-related ovarian killer protein	1	5.4
Bst2	bone marrow stromal cell antigen 2	8qB3.3	3.0
Btbd3	BTB domain containing 3	2	3.1
Calr	calreticulin	8qC3	-3.0
Camk2n1	calcium/calmodulin-dependent protein kinase II inhibitor 1	4qD3	2.7
Cd59a	CD59a antigen	2qE2	6.4
Cdh16	cadherin 16	8	5.9
Cds1	CHK2 checkpoint homolog	5qE4	5.5
Cish	cytokine inducible SH2-containing protein	9	3.3
Clca4	chloride channel calcium activated 1	3qH2	6.8
Col1a1	collagen, type I, alpha 1	11	-162.6
Col4a1	collagen, type IV, alpha 1	8qA1.1	31.2
Col4a2	collagen, type IV, alpha 2	8qA1.1	16.3
Col6a1	collagen, type VI, alpha 1	10	-3.2
Col6a2	collagen, type VI, alpha 2	10qC1	-8.7
Crip2	cysteine rich protein 2	12qF1	7.8
Ctsh	cathepsin H	9qE3.1	9.7
Cxadr	coxsackie virus and adenovirus receptor	16qC3.1	3.1
Cxcl1	chemokine ligand 1	5qE1	-2.6
Cyp51	cytochrome P450, family 51	5	-3.5
Cyp7b1	cytochrome P450, family 7, subfamily b, polypeptide 1	3	-2.7
D430044g18RIK	SH3 domain and tetratricopeptide repeats 2	18	3.0
Dap	islet amyloid polypeptide	15qB2	-3.2
Dcn	decorin	10	2.5
Ddit4	DNA-damage-inducible transcript 4	10qB4	-3.9
Dhrs6	3-hydroxybutyrate dehydrogenase, type 2	3qG3	-4.4
Ednra	endothelin receptor type B	14	-2.5
Enpp5	ectonucleotide pyrophosphatase/phosphodiesterase 5	17qB3	2.8
Etv4	ets variant gene 4	11	-4.2
F2rl1	coagulation factor II receptor-like 1	13qD1	2.9
Fdps	farnesyl diphosphate synthetase	3	-3.0
Figf	c-fos induced growth factor	XqF5	-10.4

Table 5. Differentially expressed genes between consecutive transformative stages

245:450 gene	definition	location	change
Fkbp11	FK506 binding protein 11	15	-4.3
Flrt3	fibronectin leucine rich transmembrane protein 3	2qF3	3.2
Gadd45g	growth arrest and DNA-damage-inducible 45 gamma	13qA5	3.6
Gdf15	growth differentiation factor 15	8qB3.3	7.7
Gp38	podoplanin	4	2.8
Gpx3	glutathione peroxidase 3	11qB1.3	4.0
Hist1h1c	histone cluster 1, H1d	13qA3.1	2.9
Hspb1	heat shock protein 1	5	2.6
Igf2bp2	insulin-like growth factor 2 mRNA binding protein 2	16qB1	-9.1
Igf2bp3	insulin-like growth factor 2 mRNA binding protein 3	6qB2.3	-11.1
Igfbp2	insulin-like growth factor binding protein 2	1qC3	3.9
Impact	imprinted and ancient	18	2.5
Kcnk1	potassium channel, subfamily K, member 1	8qE2	22.2
Klra4	killer cell lectin-like receptor, subfamily A, member 4	6qF3	-2.6
Lsp1	lymphocyte specific 1	7qF5	-4.0
Lss	lanosterol synthase	10	-2.6
Ly6a	lymphocyte antigen 6 complex, locus A	15qD3	-3.7
Manea	mannosidase, endo-alpha	4qA3	2.7
Mcc	mutated in colorectal cancers	18qB3	3.0
Mglap	matrix Gla protein	6	-379.7
Mgst2	microsomal glutathione S-transferase 2	3	-23.5
Mmp2	matrix metallopeptidase 2	8	-9.6
Mt1	metallothionein 1	8qC5	4.7
Mvd	mevalonate decarboxylase	8qE1	-4.7
Nbl1	neuroblastoma, suppression of tumorigenicity 1	4qD3	4.6
Ndn	necdin	7	3.6
Neu1	neuraminidase 1	17	3.6
Nipsnap1	4-NPP domain and non-neuronal SNAP25-like protein homolog 1	11	2.6
Nnmt	nicotinamide N-methyltransferase	9qA5.3	3.5
Npr2	natriuretic peptide receptor 2	4	5.5
Npr3	natriuretic peptide receptor 3	15qA1	3.9
Nrn1	neuritin 1	13qA3.3	-5.3
Nup210	IQ motif and Sec7 domain 1	6qD1	5.0
Orf63	open reading frame 63	16qC3.3	7.1
Pcolce	procollagen C-endopeptidase enhancer protein	5	-10.0
Pdgfb	platelet derived growth factor, B polypeptide	15	5.4
Plekhf2	pleckstrin homology domain containing, family F member 2	4qA1	3.0
Plk2	polo-like kinase 2	13qD2.2	2.6
Ppbp	pro-platelet basic protein	5qE1	-11.6
Prkcdbp	protein kinase C, delta binding protein	7qE3	-2.7
Pvrl2	poliovirus receptor-related 2	7qA3	4.5
Rasl11b	RAS-like, family 11, member B	5	17.1
Rasl12	RAS-like, family 12	9qC	8.7
Slpi	secretory leukocyte peptidase inhibitor	2qH3	-33.1
Sprr2g	small proline-rich protein 2K	3qF1	-4.2
Sqle	squalene epoxidase	15	-4.0
Stc2	stanniocalcin 2	11	3.0

Table 5. Differentially expressed genes between consecutive transformative stages

245:450 gene	definition	location	change
Tacstd2	tumor-associated calcium signal transducer 2	6	8.0
Tagln	transgelin	9	11.2
Tgfbi	transforming growth factor, beta induced	13qB1	-3.7
Tinagl		4qD2.2	4.5
Tnfaip3	tumor necrosis factor, alpha-induced protein 3	10qA3	3.3
Twist2	twist homolog 2	1	-8.8
Ugcg	UDP-glucose ceramide glucosyltransferase	4qB3	2.6
Wfs1	Wolfram syndrome 1 homolog	5	3.3

450:528	definition	location	change
Aqp1	aquaporin 1	6qB3	3.3
Arc	activity regulated cytoskeletal-associated protein	15qD3	-3.0
Bc056929	doublecortin-like kinase 3	9qF3	2.9
C3	complement component 3	17E1.3	7.1
Cdh16	cadherin 16	8qD3	3.5
Cdo1	cysteine dioxygenase 1, cytosolic	18qC	12.3
Clip4	CAP-GLY domain containing linker protein family, m 4	17qE2	4.7
Col4a1	procollagen, type IV, alpha 1	8qA1.1	3.4
Col4a2	collagen, type IV, alpha 2	8qA1.1	2.7
Ctsh	cathepsin H	9qE3.1	5.5
Dcn	decorin	10qC3	3.0
Enpp2	ectonucleotide pyrophosphatase/phosphodiesterase 2	15qD1	-5.4
Fmo1	flavin containing monooxygenase 1	1qH2.1	3.4
Gap43	growth associated protein 43	16qB4	-7.8
Gdf15	growth differentiation factor 15	8qB3.3	2.9
Gpx3	glutathione peroxidase 3 , transcript variant 2	11qB1.3	2.8
Gsta4	glutathione S-transferase, alpha 4	9qE1	3.6
Inmt	indolethylamine N-methyltransferase	6qB3	9.9
Kcnk1	potassium channel, subfamily K, m 1	8qE2	9.6
Ndn	necdin	7qC	-20.2
Nid2	nidogen 2	14qA3	2.8
Osmr	oncostatin M receptor	15qA1	2.8
Prelp	proline arginine-rich end leucine-rich repeat	1qE4	6.1
Rasl11a	RAS-like, family 11, m A	5qG3	3.4
Scara5	scavenger receptor class A, m 5	14qD1	3.6
Sdc3	syndecan 3	4qD2.3	-3.3
Slc7a11	solute carrier family 7, m 11	3qC	2.7
Trib3	tribbles homolog 3	2qG3	3.9
Tspan8	tetraspanin 8	10D2	8.2
Wisp2	WNT1 inducible signaling pathway protein 2	2qH3	3.8

Table 5. Gene description, location, and fold change for genes with differential expression of mRNA transcript abundance by pair wise SAM analysis Total RNA was extracted from biological triplicates of each stage and analyzed on Illumina MouseRef-8 v2.0 expression bead chips. Gene analysis by paired comparisons of consecutive time points was performed on genes with significant signal elevation over background p >0.1, exceeding 3x background (>420, 11,817 genes), and yielded 317 genes with differential expression where p<0.05 and fold change >2.5. Cytoband information and gene name obtained from Illumina Mouse Ref-8 v2.0 annotation file or secondarily from Mouse Genome Informatics (MGI, http://www.informatics.jax.org/).

Table 6. Functional enrichment analysis for genes discovered by paired comparisons

170:245	A	B	C	D	E	F
Category	BP_ALL	SPPIR	SPPIR	BP_ALL	SPPIR	BP_ALL
GO:Number	278.000			22402		22403
Term	mitotic cell cycle	cytoplasm	mitosis	cell cycle process	cell division	cell cycle phase
Count	16	43	10	21	12	15
%	0.09	0.25	0.06	0.12	0.07	0.09
PValue	5.40E-09	3.10E-07	1.14E-06	1.36E-07	9.82E-07	3.20E-07
List Total	128	143	143	128	143	128
Pop Hits	260	2169	121	596	192	306
Pop Total	14977	16241	16241	14977	16241	14977
Fold Enrichment	7.2	2.3	9.4	4.1	7.1	5.7
Benjamini	2.80E-05	2.69E-04	3.30E-04	3.53E-04	4.25E-04	5.54E-04
FDR	1.03E-05	4.82E-04	1.77E-03	2.59E-04	1.52E-03	6.11E-04
Genes	*Ndc80*	*Dusp6*	*Kntc1*	*Ak1*	*Kntc1*	*Ndc80*
	Ccng1	*Sprr2g*	*Birc5*	*Ndc80*	*Birc5*	*Bub1b*
	Bub1b	*Pkn2*	*Ccnb1*	*Ccng1*	*Prc1*	*Ccng1*
	Cdca8	*Kif2c*	*Ndc80*	*Bub1b*	*Ccnb1*	*Cdca8*
	Ncapd2	*Ak1*	*Ccng1*	*Trp53inp1*	*Ndc80*	*Ncapd2*
	Cdca3	*Prc1*	*Bub1b*	*Cdca8*	*Ccng1*	*Cdca3*
	Stmn1	*Myo7a*	*Cdca8*	*Ncapd2*	*Bub1b*	*Kntc1*
	Kntc1	*Bub1b*	*Ncapd2*	*Cdca3*	*Cdca8*	*Gspt2*
	Gspt2	*Trp53inp1*	*Cdca3*	*Stmn1*	*Ncapd2*	*Birc5*
	Birc5	*Ndn*	*Cdc20*	*Cxcl1*	*Ccnd1*	*Plk1*
	Plk1	*Cdca8*		*Kntc1*	*Cdca3*	*Ccnb1*
	Ccnb1	*Ncapd2*		*Gspt2*	*Cdc20*	*Htatip2*
	Htatip2	*Tagln*		*Birc5*		*Dbf4*
	Dbf4	*Cdca3*		*Plk1*		*Mdm2*
	Mdm2	*Kntc1*		*Ccnb1*		*Cdc20*
	Cdc20	*Birc5*		*Dbf4*		
		Gsta4		*Htatip2*		
		Ckb		*Cdkn1a*		
		Cask		*Ccnd1*		
		Htatip2		*Mdm2*		
		Pbxip1		*Cdc20*		
		Fez1				
		1810009n02RIK				
		5830467p10RIK				
		Aldh1a1				
		Pde4dip				
		Lrrfip1				
		Stmn1				
		Prdx4				
		Ddit4				
		Vps25				
		Igf2bp2				
		Ccnb1				
		Ccbl1				
		Mapkapk3				
		Aplp1				
		Cdkn1a				
		Unc13b				
		Scin				
		Mdm2				
		Eno3				
		Bzrap1				
		Pi4k2b				

Table 6. Functional enrichment analysis for genes discovered by paired comparisons

170:245	G	H	I	J	K	L	M	N
Category	BP_ALL	BP_ALL	BP_ALL	BP_ALL	BP_ALL	BP_ALL	BP_ALL	BP_ALL
GO:Number	87	32502	7067	7049	48856	74	51726	30154
Term	M phase of mitotic cell cycle	develop-mental process	mitosis	cell cycle	anatomical structure development	regulation of progression cell cycle	regulation of cell cycle	cell differen-tiation
Count	12	51	12	22	39	15	15	36
%	0.07	0.29	0.07	0.13	0.22	0.09	0.09	0.21
PValue	7.42E-07	5.23E-07	7.04E-07	1.02E-06	1.64E-06	3.15E-06	3.46E-06	2.73E-06
List Total	128	128	128	128	128	128	128	128
Pop Hits	193	3028	192	740	2058	371	374	1847
Pop Total	14977	14977	14977	14977	14977	14977	14977	14977
Fold Enrichment	7.3	2.0	7.3	3.5	2.2	4.7	4.7	2.3
Benjamini	6.42E-04	6.79E-04	7.31E-04	7.53E-04	1.06E-03	1.48E-03	1.49E-03	1.58E-03
FDR	1.42E-03	1.00E-03	1.35E-03	1.94E-03	3.13E-03	6.02E-03	6.60E-03	5.22E-03
Genes	KNTC1	HOXB7	KNTC1	AK1	HOXB7	AK1	AK1	SPRR2G
	Birc5	*Ndn*	*Birc5*	*Prc1*	*Sprr2g*	*Bub1b*	*Bub1b*	*Fst*
	Plk1	*Nppb*	*Plk1*	*Ndc80*	*Fst*	*Ccng1*	*Ccng1*	*Myo7a*
	Ccnb1	*Foxm1*	*Ccnb1*	*Ccng1*	*Myo7a*	*Trp53inp1*	*Trp53inp1*	*Bub1b*
	Ndc80	*Tagln*	*Ndc80*	*Bub1b*	*Foxd1*	*Cxcl1*	*Cxcl1*	*Trp53inp1*
	Ccng1	*Hes6*	*Ccng1*	*Trp53inp1*	*Nppb*	*Kntc1*	*Kntc1*	*Foxd1*
	Bub1b	*Dab2*	*Bub1b*	*Cdca8*	*Ndn*	*Gspt2*	*Gspt2*	*Ndn*
	Htatip2	*Dyrk3*	*Htatip2*	*Ncapd2*	*Tagln*	*Birc5*	*Birc5*	*Foxm1*
	Cdca8	*Birc5*	*Cdca8*	*Cdca3*	*Foxm1*	*Plk1*	*Plk1*	*Dab2*
	Ncapd2	*H2afz*	*Ncapd2*	*Stmn1*	*Igfbp2*	*Ccnb1*	*Ccnb1*	*Hes6*
	Cdca3	*Crip2*	*Cdca3*	*Cxcl1*	*Otx1*	*Htatip2*	*Htatip2*	*Birc5*
	Cdc20	*Htatip2*	*Cdc20*	*Kntc1*	*Dab2*	*Dbf4*	*Dbf4*	*Dyrk3*
		Grem1		*Gspt2*	*Hes6*	*Cdkn1a*	*Cdkn1a*	*Phf17*
		Nid1		*Birc5*	*Dyrk3*	*Ccnd1*	*Ccnd1*	*Etv4*
		Irs2		*Plk1*	*Crip2*	*Mdm2*	*Mdm2*	*Htatip2*
		Stmn1		*Ccnb1*	*Etv4*			*Itga3*
		Psen2		*Dbf4*	*Htatip2*			*Ccnd1*
		Aplp1		*Htatip2*	*Itga3*			*Il11*
		Scin		*Cdkn1a*	*Grem1*			*Wnt7a*
		Tob1		*Ccnd1*	*Il11*			*Igf1*
		Clu		*Mdm2*	*Wnt7a*			*Col18a1*
		Sprr2g		*Cdc20*	*Wisp2*			*Aldh1a1*
		Fst			*Igf1*			*Rab27a*
		Bub1b			*Col18a1*			*Stmn1*
		Myo7a			*Aldh1a1*			*Bmp3*
		Trp53inp1			*Nid1*			*Psen2*
		Foxd1			*Irs2*			*Ednra*
		Igfbp2			*Stmn1*			*Ddit4*
		Otx1			*Bmp3*			*Mgmt*
		Phf17			*Psen2*			*Atp6v0a1*
		Etv4			*Ednra*			*Arhgap24*
		Itga3			*Atp6v0a1*			*Cdkn1a*
		Ccnd1			*Arhgap24*			*Aplp1*
		Il11			*Aplp1*			*Scin*
		Wnt7a			*Scin*			*Tob1*
		Wisp2			*Palmd*			*Clu*
		Igf1			*Tob1*			*Scin*
		Col18a1			*Vgf*			*Tob1*
		Aldh1a1			*Nppa*			*Vgf*
		Rab27a						*Nppa*
		Bmp3						
		Ednra						
		Ddit4						
		Mgmt						
		Atp6v0a1						
		Cdkn1a						
		Arhgap24						
		Ercc5						
		Palmd						
		Vgf						
		Nppa						

Table 6. Functional Annotation for Significant Pairwise Gene Analysis

170:245	O	P	Q	R	S	T	U
Category	BP_ALL	BP_ALL	BP_ALL	SPPIR	BP_ALL	BP_ALL	BP_ALL
GO:Number	48869	51301	279		48731	7275	48513
Term	cellular developmental process	cell division	M phase	cell cycle	system development	multicellular organismal development	organ development
Count	36	12	12	14	33	39	29
%	0.21	0.07	0.07	0.08	0.19	0.22	0.17
PValue	2.73E-06	7.81E-06	1.14E-05	1.98E-05	1.73E-05	1.71E-05	1.64E-05
List Total	128	128	128	143	128	128	128
Pop Hits	1847	246	256	366	1749	2270	1416
Pop Total	14977	14977	14977	16241	14977	14977	14977
Fold Enrichment	2.3	5.7	5.5	4.3	2.2	2.0	2.4
Benjamini	1.58E-03	3.11E-03	4.21E-03	4.28E-03	5.27E-03	5.54E-03	5.68E-03
FDR	5.22E-03	1.49E-02	2.17E-02	3.07E-02	3.31E-02	3.27E-02	3.14E-02
Genes	*Sprr2g*	*Kntc1*	*Kntc1*	*Prc1*	*Hoxb7*	*Hoxb7*	*Hoxb7*
	Fst	*Birc5*	*Birc5*	*Ndc80*	*Sprr2g*	*Sprr2g*	*Sprr2g*
	Myo7a	*Prc1*	*Plk1*	*Bub1b*	*Fst*	*Fst*	*Fst*
	Bub1b	*Ccnb1*	*Ccnb1*	*Ccng1*	*Myo7a*	*Myo7a*	*Myo7a*
	Trp53inp1	*Ndc80*	*Ndc80*	*Cdca8*	*Ndn*	*Foxd1*	*Foxd1*
	Foxd1	*Ccng1*	*Ccng1*	*Ncapd2*	*Foxd1*	*Nppb*	*Foxm1*
	Ndn	*Bub1b*	*Bub1b*	*Cdca3*	*Foxm1*	*Ndn*	*Tagln*
	Foxm1	*Cdca8*	*Htatip2*	*Kntc1*	*Tagln*	*Tagln*	*Otx1*
	Dab2	*Ncapd2*	*Cdca8*	*Birc5*	*Hes6*	*Foxm1*	*Dyrk3*
	Hes6	*Ccnd1*	*Ncapd2*	*Ccnb1*	*Otx1*	*Otx1*	*Crip2*
	Birc5	*Cdca3*	*Cdca3*	*Htatip2*	*Dyrk3*	*Dab2*	*Htatip2*
	Dyrk3	*Cdc20*	*Cdc20*	*Cdkn1a*	*Crip2*	*Hes6*	*Etv4*
	Phf17			*Ccnd1*	*Etv4*	*Birc5*	*Grem1*
	Etv4			*Cdc20*	*Htatip2*	*Dyrk3*	*Il11*
	Htatip2				*Itga3*	*H2afz*	*Wnt7a*
	Itga3				*Grem1*	*Crip2*	*Igf1*
	Ccnd1				*Il11*	*Etv4*	*Col18a1*
	Il11				*Wnt7a*	*Htatip2*	*Aldh1a1*
	Wnt7a				*Igf1*	*Itga3*	*Nid1*
	Igf1				*Col18a1*	*Grem1*	*Irs2*
	Col18a1				*Aldh1a1*	*Il11*	*Bmp3*
	Aldh1a1				*Nid1*	*Wnt7a*	*Psen2*
	Rab27a				*Irs2*	*Igf1*	*Ednra*
	Stmn1				*Stmn1*	*Col18a1*	*Atp6v0a1*
	Bmp3				*Bmp3*	*Aldh1a1*	*Arhgap24*
	Psen2				*Psen2*	*Nid1*	*Aplp1*
	Ednra				*Ednra*	*Irs2*	*Scin*
	Ddit4				*Atp6v0a1*	*Stmn1*	*Tob1*
	Mgmt				*Arhgap24*	*Bmp3*	*Vgf*
	Atp6v0a1				*Aplp1*	*Psen2*	
	Arhgap24				*Scin*	*Ednra*	
	Cdkn1a				*Tob1*	*Atp6v0a1*	
	Aplp1				*Vgf*	*Arhgap24*	
	Scin					*Aplp1*	
	Tob1					*Ercc5*	

Table 6. Functional Annotation for Significant Pairwise Gene Analysis

245:450	A	B	C	D	E	F
Category	CC_ALL	CC_ALL	CC_ALL	SPPIR	SPPIR	SPPIR
Number	5615	44421	5576	-	-	-
Term	extracellular space	extracellular region part	extracellular region	signal	Secreted	triple helix
Count	35	35	36	36	21	4
PValue	6.79E-12	3.84E-11	6.82E-11	2.24E-09	9.83E-07	9.46E-06
List Total	76	76	76	81	81	81
Pop Hits	2064	2195	2375	2551	1200	9
Pop Total	15845	15845	15845	16241	16241	16241
Fold Enrichment	3.5	3.3	3.2	2.8	3.5	89.1
Benjamini	5.32E-09	1.50E-08	1.78E-08	1.94E-06	4.26E-04	1.64E-03
FDR	1.04E-08	5.88E-08	1.04E-07	3.47E-06	1.52E-03	1.47E-02
Genes	*Cd59a*	*Cd59a*	*Cd59a*	*Cd59a*	*Enpp5*	*Col4a1*
	Tacstd2	*Tacstd2*	*Tacstd2*	*Tinagl*	*Tinagl*	*Col1a1*
	Nbl1	*Nbl1*	*Nbl1*	*Nbl1*	*Nbl1*	*Col6a1*
	Col4a2	*Col4a2*	*Col4a2*	*Col4a2*	*Col6a1*	*Col4a2*
	Figf	*Figf*	*Figf*	*F2rl1*	*Col4a2*	
	Col6a2	*Col6a2*	*Col6a2*	*Figf*	*Figf*	
	Igfbp2	*Igfbp2*	*Igfbp2*	*Col6a2*	*Col6a2*	
	Npr3	*Npr3*	*Npr3*	*Nup210*	*Cxadr*	
	Cxcl1	*Cxcl1*	*Cxcl1*	*Cxadr*	*Igfbp2*	
	Gdf15	*Gdf15*	*Gdf15*	*Igfbp2*	*Cxcl1*	
	Pcolce	*Pcolce*	*Pcolce*	*Npr3*	*Slpi*	
	Dcn	*Dcn*	*Dcn*	*Cxcl1*	*Stc2*	
	Cyp7b1	*Cyp7b1*	*Cyp7b1*	*Gdf15*	*Gdf15*	
	Aqp5	*Aqp5*	*Aqp5*	*Pcolce*	*Dcn*	
	Col1a1	*Col1a1*	*Col1a1*	*Dcn*	*Pcolce*	
	Calr	*Calr*	*Calr*	*Col1a1*	*Gpx3*	
	Tgfbi	*Tgfbi*	*Ppbp*	*Calr*	*Col4a1*	
	Pdgfb	*Pdgfb*	*Tgfbi*	*Tgfbi*	*Col1a1*	
	Enpp5	*Enpp5*	*Pdgfb*	*Npr2*	*Tgfbi*	
	Pvrl2	*Pvrl2*	*Enpp5*	*Pdgfb*	*Pdgfb*	
	Manea	*Manea*	*Pvrl2*	*Enpp5*	*Mmp2*	
	Col6a1	*Col6a1*	*Manea*	*Pvrl2*		
	Fkbp11	*Fkbp11*	*Col6a1*	*Col6a1*		
	Clca4	*Clca4*	*Fkbp11*	*Fkbp11*		
	Slpi	*Slpi*	*Clca4*	*Slpi*		
	Stc2	*Stc2*	*Slpi*	*Stc2*		
	Ugcg	*Ugcg*	*Stc2*	*Ednra*		
	Sqle	*Sqle*	*Ugcg*	*Nrn1*		
	Gpx3	*Gpx3*	*Sqle*	*Cdh16*		
	Col4a1	*Col4a1*	*Gpx3*	*Gpx3*		
	Agrn	*Agrn*	*Col4a1*	*Col4a1*		
	Cyp51	*Cyp51*	*Agrn*	*Ly6a*		
	Ctsh	*Ctsh*	*Cyp51*	*Ctsh*		
	Mmp2	*Mmp2*	*Ctsh*	*Mmp2*		
	9130213b05RIK	*9130213b05RIK*	*Mmp2*	*Neu1*		
			9130213b05RIK	*9130213b05RIK*		

Table 6. Functional Annotation for Significant Pairwise Gene Analysis

245:450	G	H	I	J	K	L	M	N
SPPIR	SPPIR	SPPIR	SPPIR	CC_ALL	SPPIR	CC_ALL	CC_ALL	CC_ALL
-	-	-	-	5581	-	5578	31012	44420
						proteinaceous		
cell			extracellular			extracellular	extracellular	extracellular
binding	hydroxyproline	glycoprotein	matrix	collagen	collagen	matrix	matrix	matrix part
4	4	32	8	5	6	9	9	6
9.46E-06	6.33E-06	1.84E-05	4.15E-05	3.03E-05	6.51E-05	7.51E-05	9.05E-05	1.21E-04
81	81	81	81	76	81	76	76	76
9	8	3012	190	38	86	295	303	102
16241	16241	16241	16241	15845	16241	15845	15845	15845
89.1	100.3	2.1	8.4	27.4	14.0	6.4	6.2	12.3
1.64E-03	1.83E-03	2.65E-03	5.13E-03	5.91E-03	7.04E-03	1.17E-02	1.18E-02	1.35E-02
1.47E-02	9.82E-03	2.85E-02	6.43E-02	4.63E-02	1.01E-01	1.15E-01	1.38E-01	1.85E-01
Col4a1	*Col4a1*	*Cd59a*	*Dcn*	*Col4a1*	*Pcolce*	*Dcn*	*Dcn*	*Col4a1*
Col6a1	*col6a1*	*tinagl*	*col4a1*	*col1a1*	*col4a1*	*col4a1*	*col4a1*	*col1a1*
Col4a2	*col4a2*	*col4a2*	*col1a1*	*col6a1*	*col1a1*	*col1a1*	*col1a1*	*col6a1*
Col6a2	*col6a2*	*f2rl1*	*col6a1*	*col4a2*	*col6a1*	*col6a1*	*col6a1*	*agrn*
		Figf	*Col4a2*	*Col6a2*	*Col4a2*	*Agrn*	*Agrn*	*Col4a2*
		Col6a2	*Col6a2*		*Col6a2*	*Col4a2*	*Col4a2*	*Col6a2*
		Nup210	*Tgfbi*			*Col6a2*	*Col6a2*	
		Cxadr	*Mmp2*			*Tgfbi*	*Tgfbi*	
		Npr3				*Mmp2*	*Mmp2*	
		Gdf15						
		Pcolce						
		Dcn						
		Aqp5						
		Col1a1						
		Bst2						
		Npr2						
		Pdgfb						
		Enpp5						
		Pvrl2						
		Col6a1						
		Stc2						
		Kcnk1						
		Ednra						
		Nrn1						
		Cdh16						
		Col4a1						
		Ly6a						
		Ctsh						
		Mmp2						
		Neu1						
		Klra4						
		9130213b05RIK						

Table 6. Functional Annotation for Significant Pairwise Gene Analysis

245:450	O	P	Q	R	S	T	U
SPPIR	SPPIR	SPPIR	MF_ALL	SPPIR	BP_ALL	BP_ALL	BP_ALL
-	-	-	5201	-	6694	7275	48513
heterotrimer	structural protein	hydroxylation	extracellular matrix structural constituent	cell adhesion	steroid biosynthetic process	multicellular organismal development	organ development
4	6	5	5	7	71	39	29
1.66E-04	1.92E-04	1.65E-04	2.29E-04	7.08E-03	7.30E+01	1.71E-05	1.64E-05
81	81	81	79	81	14977	128	128
22	108	56	63	345	14.44819603	2270	1416
16241	16241	16241	16377	16241	1	14977	14977
36.5	11.1	17.9	16.5	4.1	0.5	2.0	2.4
1.43E-02	1.50E-02	1.58E-02	2.67E-01	4.01E-01	CYP7B1	5.54E-03	5.68E-03
2.57E-01	2.97E-01	2.56E-01	4.07E-01	1.04E+01	CYP51	3.27E-02	3.14E-02
Col1a1	*Sprr2g*	*Col4a1*	*Col4a1*	*Cdh16*	*mvd*	*Hoxb7*	*Hoxb7*
Col6a1	*col4a1*	*col1a1*	*col1a1*	*pvrl2*	*fdps*	*sprr2g*	*sprr2g*
Col4a2	*col1a1*	*col6a1*	*col6a1*	*col6a1*	*lss*	*fst*	*fst*
Col6a2	*col6a1*	*col4a2*	*col4a2*	*col6a2*	*ndn*	*myo7a*	*myo7a*
	Col4a2	*Col6a2*	*Col6a2*	*Tgfbi*	*Foxd1*	*Foxd1*	*Foxd1*
	Col6a2			*Cxadr*	*Foxm1*	*Nppb*	*Foxm1*
				Klra4	*Tagln*	*Ndn*	*Tagln*
					Hes6	*Tagln*	*Otx1*
					Otx1	*Foxm1*	*Dyrk3*
					Dyrk3	*Otx1*	*Crip2*
					Crip2	*Dab2*	*Htatip2*
					Etv4	*Hes6*	*Etv4*
					Htatip2	*Birc5*	*Grem1*
					Itga3	*Dyrk3*	*Il11*
					Grem1	*H2afz*	*Wnt7a*
					Il11	*Crip2*	*Igf1*
					Wnt7a	*Etv4*	*Col18a1*
					Igf1	*Htatip2*	*Aldh1a1*
					Col18a1	*Itga3*	*Nid1*
					Aldh1a1	*Grem1*	*Irs2*
					Nid1	*Il11*	*Bmp3*
					Irs2	*Wnt7a*	*Psen2*
					Stmn1	*Igf1*	*Ednra*
					Bmp3	*Col18a1*	*Atp6v0a1*
					Psen2	*Aldh1a1*	*Arhgap24*
					Ednra	*Nid1*	*Aplp1*
					Atp6v0a1	*Irs2*	*Scin*
					Arhgap24	*Stmn1*	*Tob1*
					Aplp1	*Bmp3*	*Vgf*
					Scin	*Psen2*	
					Tob1	*Ednra*	
					Vgf	*Atp6v0a1*	
						Arhgap24	
						Aplp1	
						Ercc5	
						Scin	
						Vgf	
						Nppa	

Table 6. Functional annotation for significant pair-wise gene analysis

450:528	A	B	C	D	E	F	G
Category	SPPIR	SPPIR	SPPIR	CC_ALL	CC_ALL	CC_ALL	SPPIR
GO:Number	-	-	-	44421	5615	5576	-
Term	glycoprotein	Secreted	signal	extracellular region part	extracellular space	extracellular region	extracellular matrix
Count	16	10	14	13	13	13	5
%	0.59	0.37	0.52	0.48	0.48	0.48	0.19
p-value	9.89E-06	6.29E-05	4.72E-05	7.07E-05	3.80E-05	1.54E-04	2.22E-04
List Total	27	27	27	27	27	27	27
Pop Hits	3012	1200	2551	2195	2064	2375	190
Pop Total	16241	16241	16241	15845	15845	15845	16241
Fold Enrichment	3.2	5.0	3.3	3.5	3.7	3.2	15.8
Benjamini	8.53E-03	1.80E-02	2.03E-02	2.73E-02	2.94E-02	3.96E-02	4.69E-02
FDR	1.53E-02	9.75E-02	7.32E-02	1.08E-01	5.82E-02	2.36E-01	3.43E-01
Genes	*Wisp2*	*Wisp2*	*Wisp2*	*Wisp2*	*Wisp2*	*Wisp2*	*Dcn*
	Sdc3	*Gdf15*	*Sdc3*	*Enpp2*	*Enpp2*	*Enpp2*	*Col4a1*
	Enpp2	*Dcn*	*Enpp2*	*Col4a2*	*Col4a2*	*Col4a2*	*Prelp*
	Col4a2	*Gpx3*	*Col4a2*	*C3*	*C3*	*C3*	*Col4a2*
	C3	*Col4a1*	*C3*	*Osmr*	*Osmr*	*Osmr*	*Nid2*
	Osmr	*Prelp*	*Osmr*	*Gdf15*	*Gdf15*	*Gdf15*	
	Kcnk1	*Enpp2*	*Gdf15*	*Dcn*	*Dcn*	*Dcn*	
	Scara5	*Col4a2*	*Dcn*	*Gpx3*	*Gpx3*	*Gpx3*	
	Gdf15	*C3*	*Cdh16*	*Col4a1*	*Col4a1*	*Col4a1*	
	Dcn	*Nid2*	*Gpx3*	*Prelp*	*Prelp*	*Prelp*	
	Cdh16		*Col4a1*	*Nid2*	*Nid2*	*Nid2*	
	Col4a1		*Prelp*	*Ctsh*	*Ctsh*	*Ctsh*	
	Prelp		*Nid2*	*Aqp1*	*Aqp1*	*Aqp1*	
	Nid2		*Ctsh*				
	Ctsh						
	Aqp1						

Table 6. Functional enrichment analyses and biological significance of differentially expressed genes from paired comparisons of consecutive time points in spontaneous transformation. Differentially expressed genes were determined using paired comparisons analysis at a significance level p <0.05. All available categories offered through NIH DAVID were *tested for significance based on the observed versus expected number of genes in each category. Uniprot keywords (SP_PIR),* Kyoto Encyclopedia of Genes and Genomes pathways (KEGG), Gene Ontologies b*iological processes (BP), molecular function (MF) and cellular componentry (CC) were determined at a depth of three annotation levels and requiring a minimum of five genes per annotation where p <0.05 after correction for multiple hypothesis testing by randomization.*

3.7.2 Identifying putative oncogenes or tumor suppressor genes

As oncogene up regulation or tumor suppressor down regulation has been suggested as drivers for tumorigenesis, we cross referenced the early, mid, and late lists of genes to determine which are differentially expressed across at least three of the four key transformative stages. Area proportional Venn diagrams show the intersection of these gene lists, and their relation to the findings from the longitudinal time course analysis (Figure 10A). Considering the possible 18,000 genes referenced on the array, relatively few genes belonged to each category. Sixteen genes were found in the intersection of early: middle (170:245:450=16), and two were found in the intersection of intermediate: late (245:450:528=2). A single gene, Necdin (NDN), was the only gene shared in the intersection of all three sets (170:245:450:528=1). STEM analysis places NDN, a gene expressed in terminally differentiated neurons in cluster 6. NDN fits a profile of continuous down regulation (1, -1.92, -3.56, -6.32). Together, the intersection of the combined significant gene lists from the SAM pair wise analysis included 135 SAM genes U 599 STEM genes (Figure 10A). The first in the series of numbers reflects the fold change from the upper term in the column header, the second number for the lower term (Figure 10).

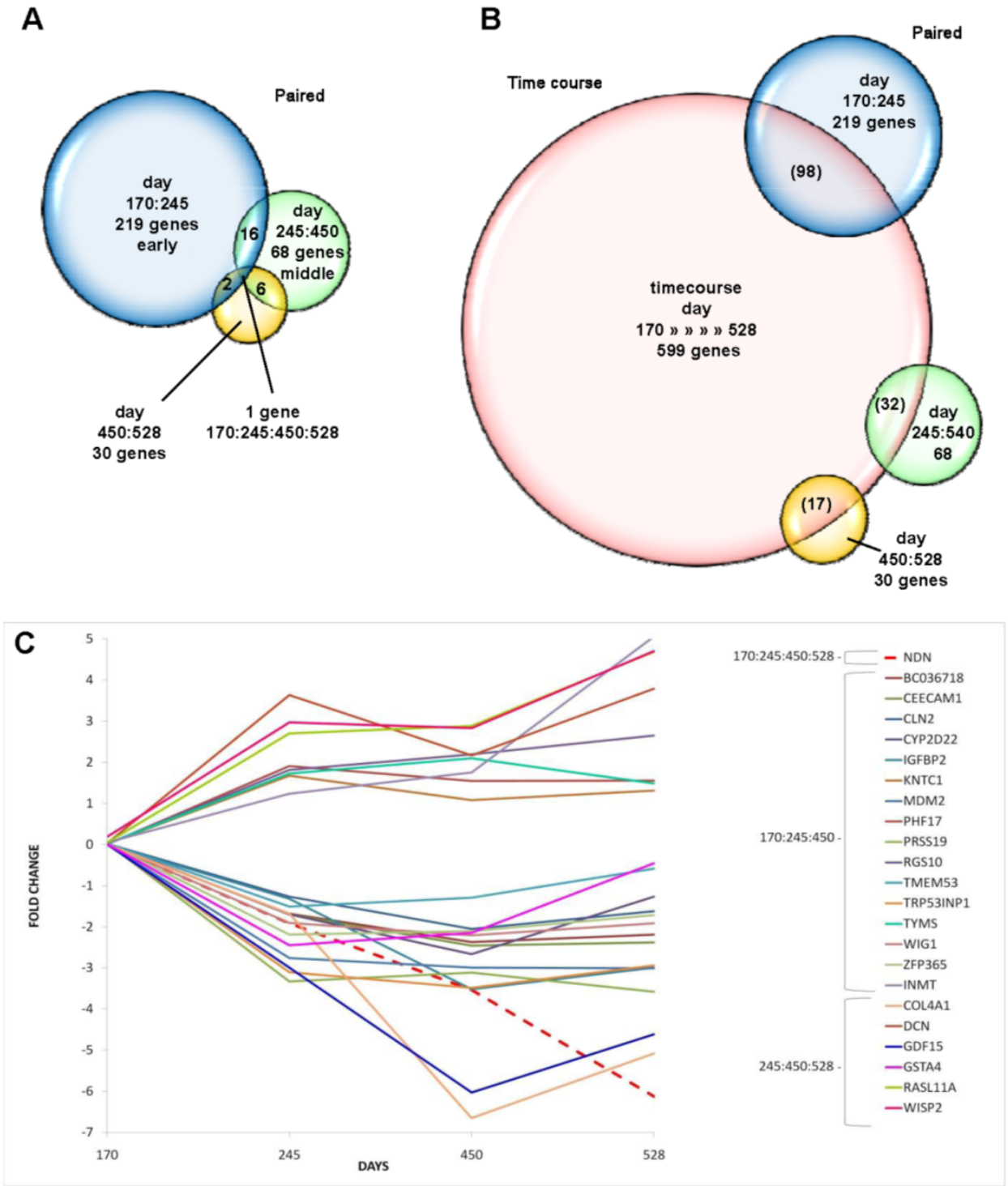


Figure 10. Analysis of differentially expressed genes from each pair of consecutive key transformative stages, and between timecourse and paired analyses. Traditional gene analysis by paired comparisons of consecutive time points yielded 317 genes with significant differential expression where p<0.05 and fold change >2.5 (A). The relationship among 599/ 18,000 differentially expressed mRNA transcripts and the genes discovered by timecourse analysis; 98, 170:245; 32, 245:450; and 17, 450:528 genes from the paired comparison totaled 147/ 599 genes in intersection of these sets and the single gene, NDN that was common among all 3 pair wise analyses (B). Fold change for individual candidate "driving" genes with differential expression of mRNA transcript abundance in more than one pair wise SAM analysis (C).

3.8 Whole chromosome copy number change versus total mRNA transcript abundance by chromosome.

By design, this study afforded the unique opportunity to compare the relationship between changes in chromosome copy number, and changed gene expression level for that chromosome. Additionally, we evaluated if this relationship was different in early transformation compared to late stages. Some chromosomes underwent large changes in copy number while others remained relatively static (chromosomes 12, 1), we sought to define the relationship (if any) between chromosome copy numbers and average mRNA transcript from that chromosome. The relationship between chromosome gains and transcriptome gains and how these values change from one stage to the next (Figure 11 A-D). When expressed as fold change from one transitional stage to the next, correlation and significance decreased with time across key transformative stages. The correlation from day 170 cells to day 245 cells was the highest, but still only accounted for one-third of the variability between whole chromosome and transcriptome (170:245 R2=0.12, correlation=0.35, b=0.27). Comparing fold change for days 245:450 and 450:529, the correlation became progressively lower (day 245:450, R2=0.06, correlation=0.24, slope=0.23; day 450:528 R2=0.02, correlation=0.14, and slope=0.08). These findings suggest that the shift away from high-fidelity genome transfer and system homogeneity is related to the reduced correlation between chromosome: transcriptome.

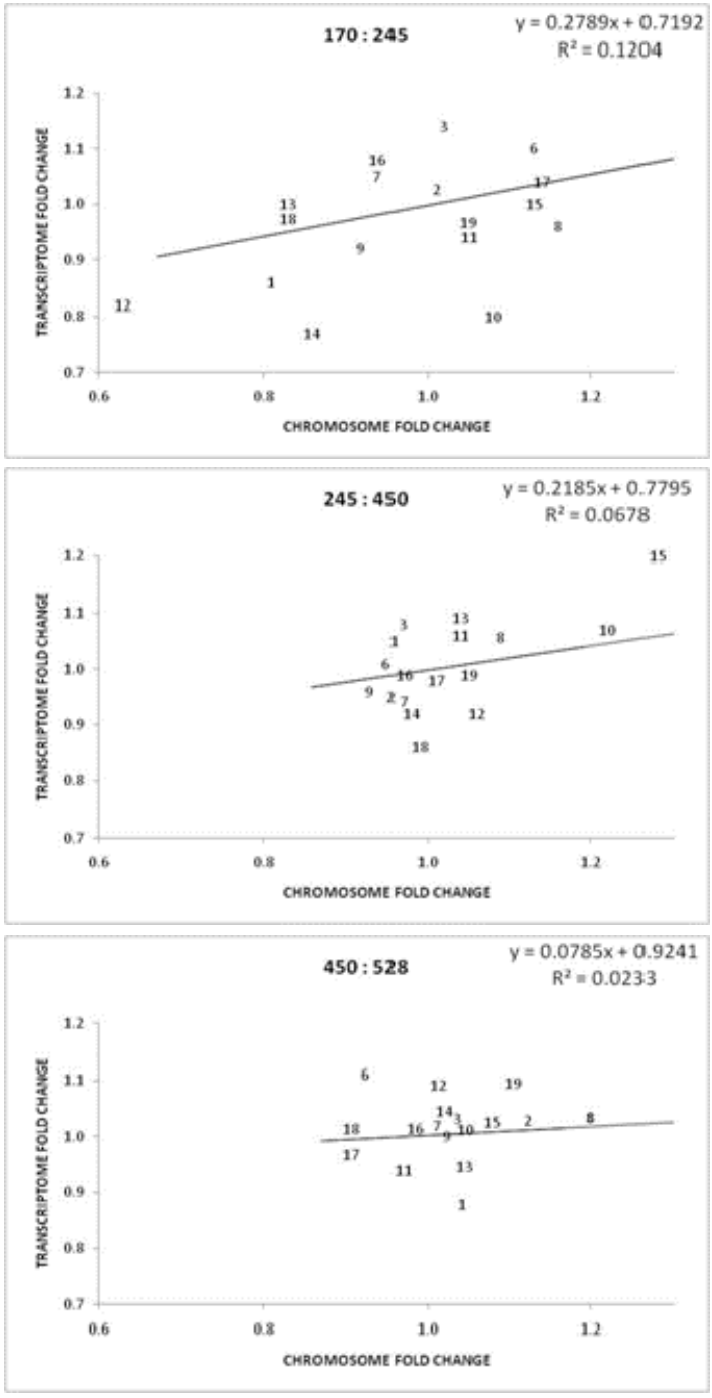


Figure 11. Whole chromosome copy number change for each chromosome (x-axis) and fold change for the total mRNA transcript from that chromosome (y-axis, A-C). For fold change between consecutive transitional stages, correlation and significance increased with transformation from day 170:245 R^2=0.12, correlation=0.35, b=0.27; day 245:450 chromosome R^2=0.06, correlation=0.24, slope=0.23; day 450:528 R^2=0.02, correlation=0.14, and slope=0.08.

3.9 Tumorigenicity of select key transformative stages in vivo

Cells from two key transitional stages (day 245 *versus* 450) were selected based on karyotype and *in vitro* phenotypic differences for *in vivo* tumorigenicity evaluation in both syngeneic C57BL6 and nod-SCID-gamma immune compromised mice (Table 1, 7). Data for the 16 syngeneic mice and 4 nod-SCID-gamma mice are found in Table 7. Each C57BL6 mouse was injected in 3 sites. As intraperitoneal tumor formation was usually apparent only at the latest stages, injections on the left and right back served as sentinel markers to detect tumor progression. Day 245 cells were very weakly tumorigenic, resulting in a single 1.5 mm low grade subcutaneous nodule at day 70 in 1/7 C57BL6 mice (Table 7, Figure 12A). Day 528 cells were significantly more efficient at tumor formation and formed subcutaneous tumors in all (7/7) C57BL6 mice, totaling 10/14 positive subcutaneous sites. These tumors were of higher grade, larger size, and were more invasive than the single tumor found to arise from day 245 cells (Table 7, Figure 12B). Day 528 cells in syngeneic BL6 mice resulted in one of two phenotypes. Either multiple diffusively invasive sites were established on the peritoneal fat with atypical nuclei (arrows, C), or immunoreactive B-cell clusters were found approximating the intraperitoneal fat (not shown). Isolated intraperitoneal injections in nod-SCID-gamma mice gave rise to diffuse nodular intraperitoneal tumors in 2/2 mice. ID8 cells served as positive controls in both types of mice and revealed a similar phenotype to day 528 cells in nearly all measures, with exception of slightly increased subcutaneous tumor size in ID8 cells (Table 7, image not shown). Negative control mice injected with matrigel vehicle occasionally showed a ~1mm acellular collagen deposit at the site of the subcutaneous injection resembling scar tissue (image not shown).

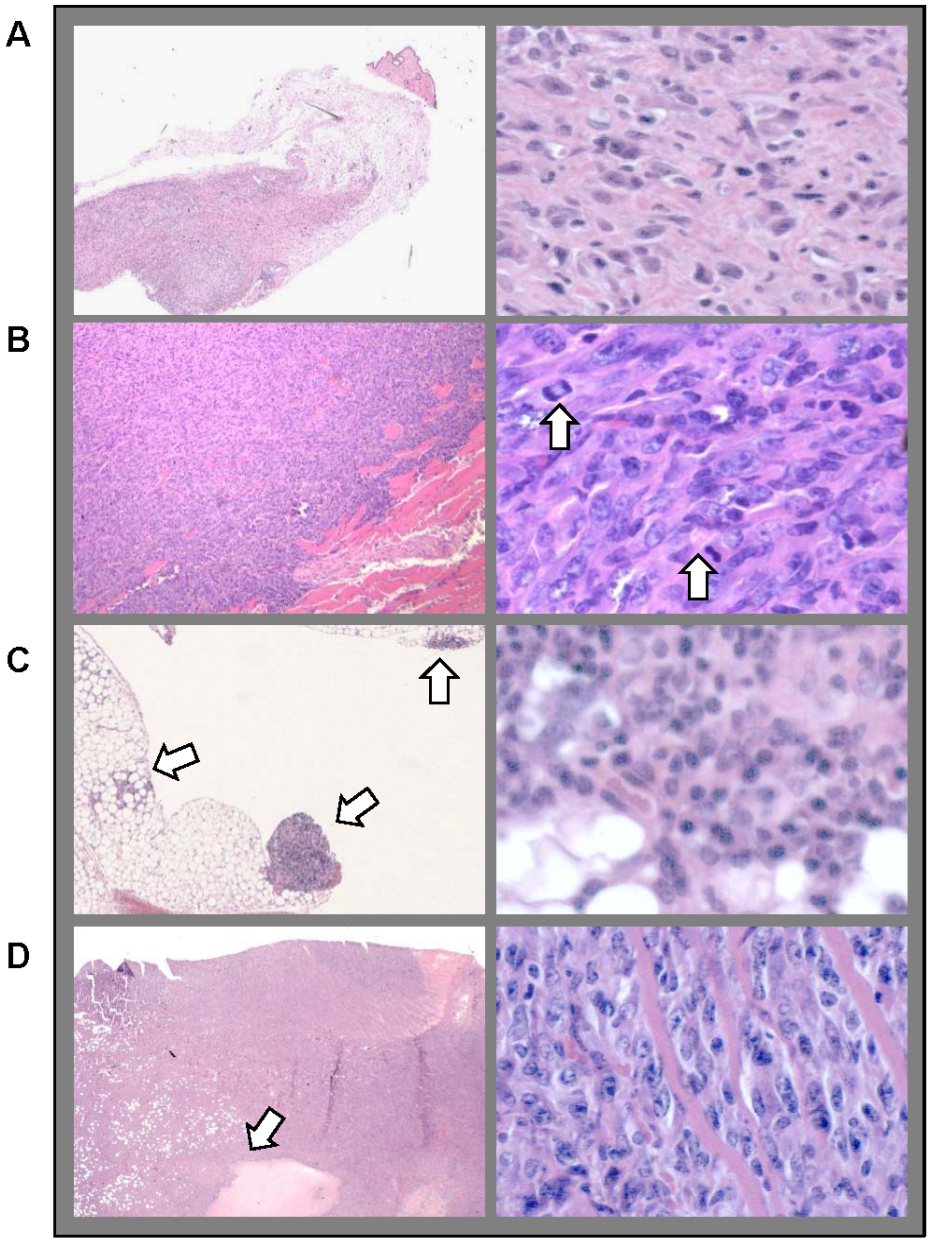


Figure 12. Histopathology of subcutaneous and intraperitoneal tumors in vivo for day 245 vs. day 528 MOSEC allografts. Day 245 cells formed a low-grade 1.5 mm subcutaneous nodule in 1/7 C57BL6 mouse that was contained within the dermis (A). Day 528 cells typically were more infiltrative with diffuse infiltration beyond the dermis and into the muscle layer and numerous atypical nuclei and metaphases were seen per high power field (arrows, B). Intraperitoneal injections of day 528 cells formed either multiple diffusively invasive sites on the peritoneal fat (arrows, C). Nod-SCID gamma mice injected with day 528 cells had more invasive and larger intraperitoneal tumors with invasion into muscle and fat and regions of necrosis (arrow, D). Images were captured following 70 days in vivo growth if allografted cells, harvesting, fixation and staining, using a light microscope at 20x (left column), 600x (right column).

3.10 Characterizing the extent of karyotype heterogeneity in late stage, Brca1$^{\Delta 5-13}$ and harvested tumor cells.

MOSEC with Brca1$^{\Delta 5-13}$ had very high H' and H values despite only 50 days in primary culture and 25 days in culture post AdCre recombination. When comparing the karyotype profile of the injected day 528 cells to each of the harvested tumors, many features of the karyotype were altered between the *in vitro* and *in vivo* karyotype profiles (Figure 13). The Shannon index increases significantly in both Tumor 1 and Tumor 2 over the level of the injected cells. The 4;3 clonal translocation seen in the day 528 cells is present in both tumors at approximately the same frequency (day 528: 32%, Tumor 1: 32%, Tumor 2: 42%). Tumors 1 and 2 additionally had a very high frequency of translocations that were not seen previously in this cell line at 4;5 ~95% of tumor cells, and 5;4 in ~90% of tumor cells. Additionally, several shared minute chromosomes were seen in these tumor cell lines that contributed to their increased variability. NCCA frequency was much greater (essentially 100% of cells contained more than one non-clonal aberration). Chromosome rearrangements were significantly more complex in tumor 2, involving several chromosomes in most single aberrations. Tumor 2 had an increased whole chromosome karyotype heterogeneity by Shannon Index H' compared to tumor 1, (Figure 13, Table 2).

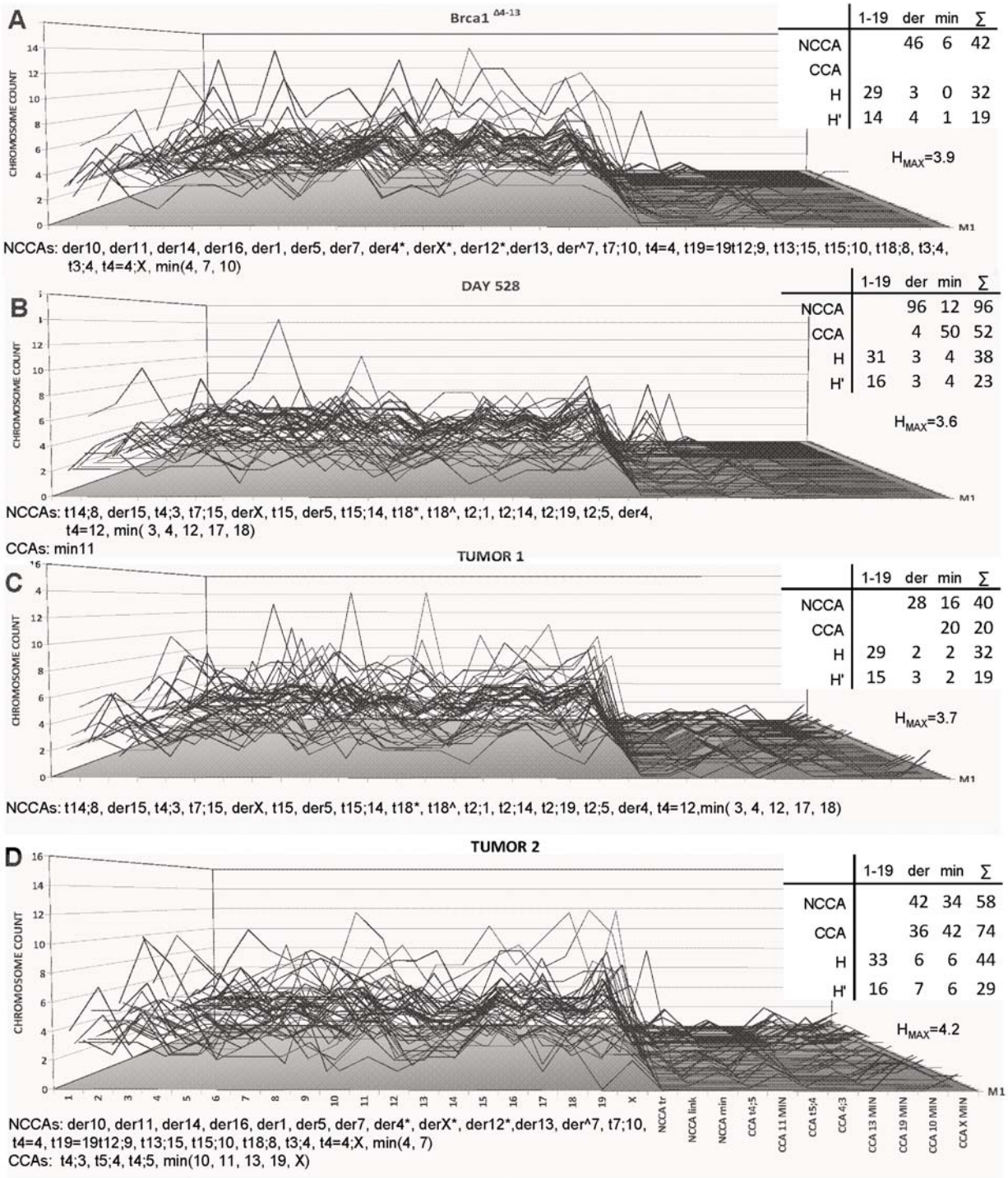
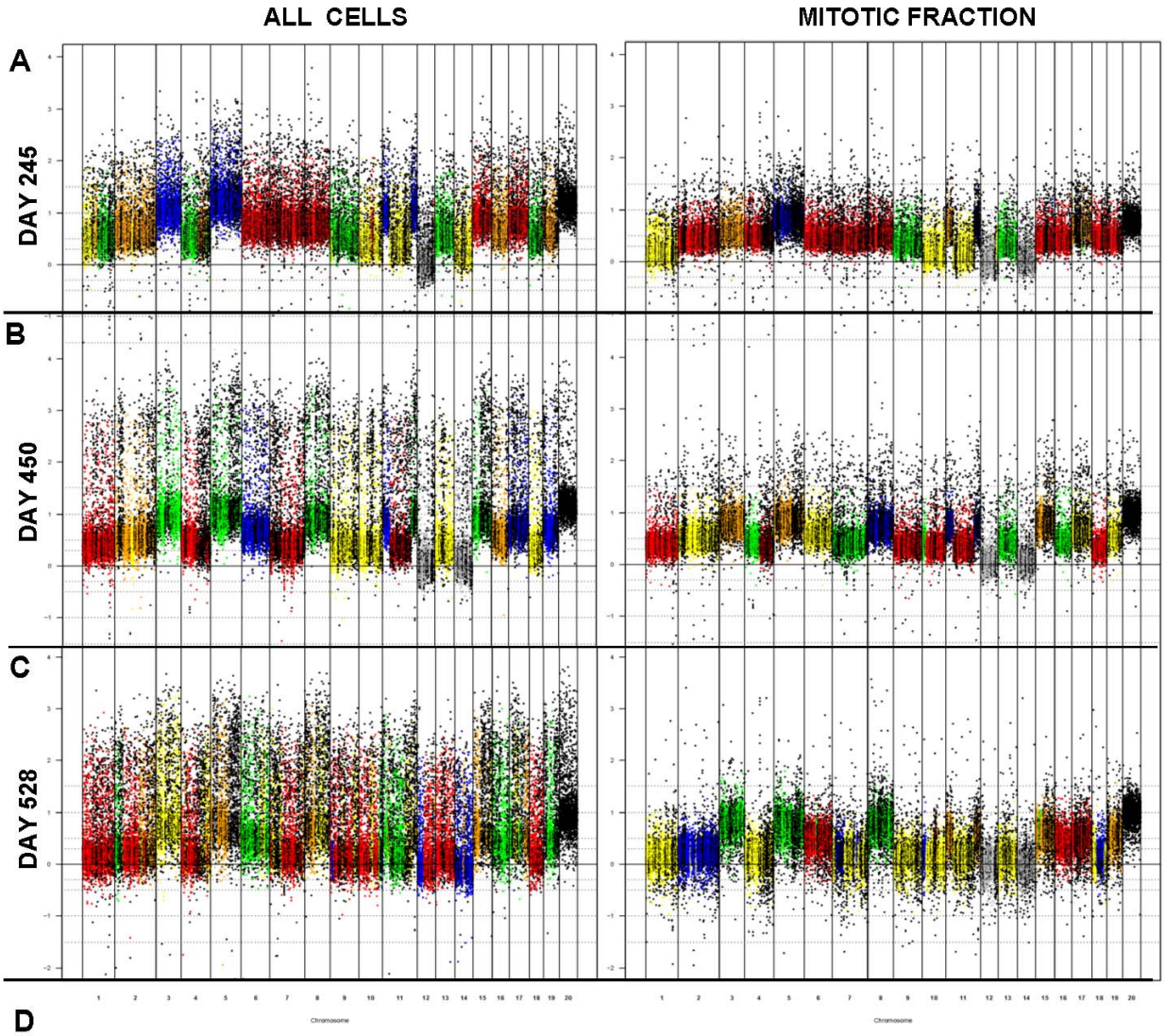


Figure 13. Karyographs showing increasing genomic mosaicism in spontaneous versus gene-based transformation and in two harvested tumors. MOSEC with Brca1$^{\Delta 5-13}$ had the highest diversity of all cell lines despite only 35 days in primary culture and 25 days culture post AdCre recombination. Day 528 cells were injected into C57BL6 mice and were characterized by a clonal 4;3 translocation. 70 days post injection, tumor 1 and tumor 2 were harvested and showed different profiles from the injected tumor and variable frequencies of non-clonal chromosome aberrations. Shannon's Index for population diversity was calculated for each by chromosome and for the entire time point and was significantly elevated in Brca1$^{\Delta 5-13}$ and tumor cells over day 528 injected cells.. Ploidy (y-axis) of clonal and non-clonal whole and derivative chromosomes (x-axis) is shown for 50 metaphases per time point (M1-M50, z-axis). Add p-values for Shannon Index

3.11 DNA copy number analyses for variability, large aberrations, and whole chromosome counts using array competitive genomic hybridization.

To determine the reliability of SKY karyotype analyses, whole genome DNA was extracted from frozen cell stocks grown for 48-72 hours in culture. Cells at transformative stages 245, 450, 528 were harvested as entire cell populations (control, left (Figure 14 A-C)), or mitotic enriched following 12 hour colcemid treatment 0.1 µg/ µl and mitotic shake off (mitotic, Figure 14, right (A-C)). Using PopLowess normalization procedures, data were clustered by k-means and normalized based on the largest single copy number population that could be tracked. This is more appropriate for cancer genomes where chromosome copy number is known to vary from one chromosome to the next. Data were centered by the median of the largest cluster to account for count variability among chromosomes and to allow for universal zeroing. These regions were checked against SKY data for chromosome aberrations and values within one standard deviation of the median for each cluster (colored points) and those that were outside one standard deviation (black points). Colored plots were used to compare counts between control and mitotic enriched fractions (>0.9). The number of probes on each chromosome with copy number >1 SD ouside the cluster median increased with days in culture for both control and mitotic fractions p=0.032 as was less in mitotic fractions compared to controls. This is thought to relate to synchronization of ploidy of these cells as colcemid has been shown to induce DNA synthesis as well as metaphase arrest. As the SKY data was gathered using the same colcemid treatment, comparisons were made between the cytogenetic data and the mitotic aCGH populations. Derivative features that were clonal in spectral karyotype data (CCAs) and also found in the colcemid treated mitotic fraction where a second cluster spans more than 20% of a single chromosome (aCGH (CCA)). The aCGH plots normalized by PopLowess show copy number expansion on within a chromosome where more than one color is assigned within a single chromosome. Overlap between Poplowess and SKY data show agreement in copy number change for t 4;3 and for 11min (Figure 14D). f



D

DAY	CONTROL	MITOTIC
245	1, 4, 11	11 (11min) *
450	11	4 (t4;3) *, 11 (11min) *
528	2, 4, 6, 7, 8, 11, 12	7, 11 (11 min)*, 18

Figure 14. DNA copy number analyses for variability, large aberrations, and whole chromosome counts using array competitive genomic hybridization. DNA extracted from cells at transformative stages 245, 450, 528. Cells were harvested as entire cell populations (control, left (A-C)), or mitotic enriched following 12 hour colcemid treatment 0.1µg/µl and mitotic shake off (mitotic, right (A-C)). Clustering with default of 5 populations, merge cluster criteria of cluster medians different by M<0.2, was performed using PopLowess. Data were centered by the median of the largest cluster and zero and universally corrected by constant so that the cluster with the lowest median matched reference DNA (2n, gray). Values within one standard deviation of the median for each cluster (colored points) and those outside one standard deviation (black points). Colored points were used to determine correlation between control and mitotic enriched fraction (>0.9) and variability as number of probes on each chromosome >1 SD ouside the cluster median increased significantly with with days in culture for both control and mitotic fractions p=0.032 as was significantly less in mitotic fractions compared to controls. Derivative features that were clonal in spectral karyotype data (CCAs) and also found in the colcemid treated mitotic fraction where a second cluster spans more than 20% of a single chromosome (aCGH (CCA)), (D).

Table 7. In vivo tumorigenicity analyses of mouse ovarian surface epithelial cells

Cell Type (days)	245	528	ID8
Subcutaneous Tumors			
C57BL6 Positive Mice *	1/ 7	7/ 7	2/ 2
C57BL6 Positive Sites	1/ 14	10/ 14	2/ 4
Tumor Size (mm)	1.5	11.1 (10.2-11.9)	9.7 (7, 13)
Invasiveness	2/ 5	4.3 (3.7-4.9)/ 5	4.0 (3, 5)/ 5
IP Tumors			
C57BL6 Positive Mice	0/ 7	2/ 7	1/2
Organ involvement	-	3.5	2
nod-SCID-gamma IP Tumors	0/ 2	2/ 2	1/ 1
Organ involvement	-	4/ 10	4/ 10
Total IP Positive Mice	0/9	4/9 *	2/ 3
Average organ Involvement	-	3.8 (3.3 - 4.3)/ 10	3/ 10

Day 245 vs. day 528 vs. ID8 cells (100,000/ site) were injected into sixteen syngeneic C57BL6 mice and five nod-SCID-gamma mice at three sites (left and right subcutaneous flank and intraperitoneal).

Summary data for are presented from gross and microscopic histopthology analyses of tumors harvested 70 days post allograft injection of day 245, day 528, ID8, or vehicle alone.

Invasiveness was scored by assigning 1 point for each of the following reported as mean (range):

Subcutaneous scoring: dermis, subcutaneous fat, muscle, diffuse, necrotic (max= 5 each side).

Intraperitoneal scoring: pancreas, liver, peritoneal fat, reproductive organs, lymph node, spleen, small intestine, large intestine, diaphragm, hemorrhagic ascites (max=10).

**Significant between day 245 and day 528 cells by two-tailed Fisher's exact test $p < 0.05$*

4. DISCUSSION

4.1. Overview

The results of the current study emphasize the significance of karyotype heterogeneity in transformation and tumorigenesis. Specifically, the influence of karyotype heterogeneity on generating genomic diversity that facilitates selection of cancer phenotypes and the differential effects of whole chromosome copy number change on gene expression during the key transformative stages of spontaneous tumorigenesis. Considering cancer cell populations to be complex-adaptive systems that are characterized integrative regulatory systems and by linear and non-linear genotype-phenotype relationships. Therefore, multi-level analyses were performed longitudinally on data from whole genome microarrays, population karyotype profiles, and cell phenotype. Variability of a sufficient magnitude to shift cell populations from homogeneous diploid cells to a mosaic of structural and numerical chromosome alterations reflects low-fidelity genome transfer that began well before the morphologic and behavioral change to the transformed phenotype. Karyotype heterogeneity was quantified by an adaptation of the Shannon Index, and reached a maximum at day 528. After only 50 days in culture, the parallel MOSEC line with conditional *Brca1* inactivation had significantly greater population diversity than the spontaneously transforming MOSEC at day 528. This demonstrates how cancer-associated genes can link micro-evolutionary gene-level change to macro-evolutionary change in karyotype that rapidly facilitates the generation of unique genomes within a cell population. Day 528 cells with heterogeneous karyotypes were tumorigenic and rapidly shifted from the karyotype of the injected line to population karyotypes of increased variability in response to the shift from *in vitro* to *in vivo* environments. Multiple unique fold change profiles characterized the differential expression of ~600 differentially expressed genes throughout spontaneous transformation. Functional enrichment for these genes was significant for DNA/ nucleosome / chromosome related, and cholesterol synthesis/ microsome related categories. The multi-level longitudinal data presented support genomic

instability as a means for increasing population diversity by increasing cell-to-cell variations in karyotype. Thus, genomic instability would also be permissive for rapid adaptation of populaitons of late stage transformed cells. Together, these findings support a genome-centered evolutionary framework for cancer progression that emphasizes cell-to-cell genomic variability as the basis for macro-evolutionary selection and rapid phenotypic switching by whole chromosome copy number change. These data demonstrate the significance, methodologies and rationale for quantifying karyotype heterogeneity in transformation and tumorigenesis. A foundation is provided for incorporating these concepts and techniques into clinical and research based applications for improved cancer detection and treatment strategies.

4.2. Adapting the Mouse ovarian surface epithelial cell model

Similar to previous studies using the mouse ovarian surface epithelial cell (MOSEC) model, we have identified four key transformative stages based on the appearance and disappearance of key phenotypic and behavioral characteristics during spontaneous transformation [54-55, 63]. Several groups have utilized the MOSEC model with attention to a specific hypothesis or with emphasis on a certain feature of the transformation process. These include but are not limited to studies of specific processes (cytoskeletal remodeling [55]), selected gene modifiers of transformation and tumorigenicity (*Brca1*, *p53*, and *Rb* [48]), and pathway modification by specific compounds (alpha-lipoic acid and NF-KB pathway) [100]. Others have taken a genome-wide approach including array based studies that compare genome-wide aCGH analysis of DNA copy number with microarray data for gene expression to compare an early stage MOSEC line and with harvested MOSEC tumor lines [57, 61]. Rather than focus on a specific pathway or stage, the current study provides a global assessment of genome content and gene regulation in relation to cell phenotype throughout the transformation process. The levels studied range from cell behavior to measurements of karyotype heterogeneity for cell populations at key transformative stages and include cell phenotype assays, mRNA transcript abundance, DNA copy number change, the characterization of structural and numerical

chromosome change, and the in assessment of tumorigenic capabilities and tumor histopathology. The design allows for the comparison of a single variable over time as well as an integrative analysis of the interactions between variables. The current study design is unique in this regard and additionally in its characterization of cell populations while preserving resolution at the level of the individual cell.

Several specific changes to the experimental design and methodologies were made to previous publications using the MOSEC model with the following rationale: previous studies using batched surface epithelial cells from more than ten mice to initiate the primary culture demonstrated significant karyotype heterogeneity ranging from 2n – 6n at the earliest time point in this study [55]. The culture was initiated only with cells from a single mouse ovary to minimize the possibility that early stage heterogeneity could be attributed to the batching of cells from many different mice. The key transitional stages of MOSEC transformation have been well defined in previous studies [55] and emerged with predictable sequence in the current study, albeit with a longer timeline. MOSEC displayed slow contact-inhibited growth with low invasion and proliferation capabilities until day 170 and gradually progressed to acquire the capacity for rapid growth, increased migration, invasion with upward stratification, and three- dimensional branching from day 245 onwards. These acquired traits are well established hallmarks of the malignant phenotype and, as previously reported, the current MOSEC transformation shows that a strong correlation exists between these behaviors *in vitro* and tumorigenic potential *in vivo*. Therefore, the genome-wide multi-variate assessment performed is useful not only for time course analysis, but also to evaluate changes in a pair-wise fashion associated with the acquisition of key characteristics leading to the tumorigenic phenotype.

Additional noteworthy differences should additionally be highlighted between our SKY data collection and reporting that extend beyond current standards for cancer cytogenetics. These differences are important because, in addition to reporting the aneusomies found in high frequency in a cell population, this study emphasizes the quantification of low frequency events

and variability from cell to cell. Therefore, high quality metaphases for SKY analysis were selected with minimal overlap and clearly defined borders for each metaphase. This minimizes the possibility that artifactual numerical change or improper chromosome assignment would occur. Additionally, to thoroughly quantify karyotypic variability of cell populations for comparison across different ploidy levels, a high number of complete metaphases (n=50) are counted and reported.

4.3. Stress Induced Senescence as an Incomplete Mechanism of Tumor Suppression

Even when the primary culture was initiated from a single mouse ovary, day <40 cells displayed chromosome copy number change as well as nuclear and cytogenetic morphological abnormalities. These were visualized in images of DAPI stained and SKY painted interphase nuclei and in metaphase spreads. In this earliest evaluation of karyotype change, aneuploidy ranged from loss of a single chromosome, doubling of the entire genome with multiple chromosome fragments and functional defects in extreme cases. Based on the senescence and karyotype data, we can speculate as to possible causes of this early aneuploidy as it relates to ovarian cancer in humans. The incessant ovulation hypothesis is the dominant theory for ovarian cancer causation in humans and suggests that repeated rupture-repair cycles eventually lead to transformation and increased ovarian cancer incidence. According to this hypothesis, successive bouts of apoptosis and regenerative repair of OSE cells at the ovulation site induces genetic instability via oxidative DNA damage, over expression of *p53* at the rupture site and anti-apoptotic *Bcl-2* at the rupture margin, and replication of cells at the margin and their migration to the wound [101]. Survival and subsequent expansion of OSE cells with unrepaired genomic damage and heightened survival potential is theorized to predispose the epithelial cell layer to tumorigenesis [28]. The early genomic instability seen in the current study cannot be due to the inherent MOSEC instability caused by repeated rupture-repair cycles. This is because the first estrus cycle of the mouse occurs at 4-6 weeks of age [102], so harvested

cells from six-week-old mice possibly could have been exposed only to a single cycle of follicle rupture and repair. Therefore, the genomic instability seen in the MOSEC model reflects growth and regeneration occurring as a function of days in culture. As the cells are passaged before reaching confluency and therefore continually divide.

4.2 Cell Senescence

Cellular stress likely accounts for early aneuploidy and is evidenced by the majority of the cell population (>95%) undergoing senescence at day <40. Cellular senescence has been correlated with aneuploidy [103], and is thought to function as a tumor suppressive mechanism in genomically unstable cells [104-106]. The findings that the small minority of dividing cells display aneuploidy with nuclear and cytogenetic abnormalities suggests that, as a potential tumor suppressive mechanism, the halting of genomically unstable cells through cellular senescence is an incomplete mechanism of tumor suppression. Another possibility recently described is the escape of senescence by neosis. Briefly, this process involves the formation of several daughter cells with viable genomes from giant polyploid cells via nuclear budding and asymmetric cytokenesis [107]. The finding that the day 170 culture contained almost no senescent cells suggests incompatibility of the senescent cell phenotype with MOSEC tumorigenesis. Possible mechanisms for this shift include senescence escape mechanisms such as neosis, or progressive replacement due to differential attachment, replication, or death of the senescent versus mitotic cell fractions. Aneuploidy was present in ~32% of cells by small alterations in whole chromosome copy number change and by the prevalence of linked chromosomes. Interestingly, among 50 metaphases, tetra-ploid genomes such as that detected in the day <40 cells were not seen. This suggests that doubling of the genome in the early stages of transformation is linked to reduced viability of cells with 4n genomes.

The increased frequency of murine cell immortalization is attributed to the substantial increase in telomere length in the mouse compared to the human (40-60 Kb, 10 Kb respectively). Additionally, mice have constitutively active telomerase activity in contrast with

the re-activated telomerase activity responsible for transformation of human cells [108-109]. The short time in culture before the majority of MOSEC became senescent combined with the properties of constitutively active telomerase and longer primary telomeres reduces significantly the possibility of telomere associated senescence. The increased stability and maintenance of rodent telomeres could account for spontaneous transformation and tumorigenesis in these lines. However, the ability of the MOSEC to overcome high frequency senescence based arrest suggests that differential entry of human versus mouse cells into senescence is not responsible for their contrasting abilities to immortalize. By comparison, this suggests that telomere induced senescence may be sufficient to prevent tumorigenesis; whereas stress induced senescence may be a less tightly controlled process. Some evidence suggests that different patterns of telomerase activity may be linked to the differing cytogenetic profiles of mouse versus human epithelial tumor cells. For example, the frequency of clonal non-reciprocal translocations is greatly increased in telomerase deficient compared to control Tr*p53* mutant mice, reaching at least one translocation in each chromosome after only a short time in culture [110]. The acrocentric mouse chromosome morphology may therefore be differentially subject to specific aberration patterns compared to human chromosomes. The mouse chromosome centromere is almost directly adjacent to one telomere reducing the physical space in which translocation events might occur by one half [111]. Despite these differences in telomere structure and function, the fact that high-degrees of genomic instability are reported in ovarian cancer cell lines and tumors from both species [87] supports the validity of this model.

4.3 Quantifying Genomic Instability in Cellular Transformation

The application of SKY analysis to early stage cells provides insight into a number of questions regarding early transformative events. Several types of nuclear and cytogenetic abnormalities were detected in both interphase and mitotic cells, which implicates structural and functional pathologies as early contributors to the transformation process. Evidence for both

gene and chromosome based causes of cellular transformation has been provided by experiments in which genes or chromosomes and the rate and frequency of cellular transformation that follows are tracked [112-115]. Unlike other published data, the current methodologies did not alter the cellular chromosomal content or any specific gene(s). Rather, to more closely mimic the majority of human cancers, we observed the process of spontaneous transformation from the normal diploid phenotype over time. Due to the large body of evidence associating multiple checkpoint [116-118] and mitotic associated proteins [119] with chromosomal instability and aneuploidy, all of which may contribute to chromosomal instability in various systems, we sought to determine the effect and extent of early instability rather than determine its specific cause. Evidence for mitotic instability are provided by the images of abnormal nuclear morphologies [120] such as the blebs and nuclear bridges seen in day <40 cells. Cells with the ability to pass through the colcemid induced pro-metaphase block and proceed through mitosis in the presence of colcemid without spindle fibers [99] are identified by two separated centromeres per chromosome. The chromosomes then decondense coalesce and form micronucleated restitution nucleus. Cells in colcemid-telophase or forming a restitution nucleus have been described as those shown here with two separated centromeres per chromosome. This suggests that even in this early stage, the cells have lost the ability to maintain arrest at the spindle checkpoint. Other abnormalities include rare cells with highly aneuploid karyotypes, multiple breaks and structural abnormalities. The fate of these cells is not known, but it has been suggested that for the most part, their viability is reduced compared to diploid cells [121]. This could possibly account for the reduced frequency of day 170 polyploid cells.

Early in the study of cancer, the visualization and experimental manipulation of chromosomes led scientists to postulate that misdistribution of chromosomes might cause tumor development [112]. Since this time, it has become clear that many tumors are not only aneuploid, but also have acquired mutations in oncogenes and tumor suppressor genes, fuelling

the debate over whether genomic instability is the cause or consequence of cellular transformation [122-125]. Independent of the cause for cellular transformation, the implications of genetic heterogeneity as the basis for increased variability, selection, and genetic adaptation are generally underestimated by most in the field [126]. Therefore, to track the effects of early genomic instability during spontaneous transformation of near-normal diploid cells, population based karyotype analysis was performed at each key transformative stage. The level of cell-to-cell genomic heterogeneity is believed to provide a snapshot of chromosomal instability and allows the assessment of cancer progression from an evolutionary perspective by tracking patterns of change at the level of the entire karyotype [127-130]. These data are the first to evaluate karyotype heterogeneity in a longitudinal model that begins from a near-normal relatively homogeneous karyotype where chromosomal instability has not been specifically induced by gene or chromosome manipulation.

The mosaic of cellular karyotypes common to solid tumors and blast-phase hematologic neoplasms are often ignored or filtered to emphasize clonality within a population of cells. Despite the emphasis on pattern detection inherent in most data collection and interpretation, the true biological variability resulting from low-fidelity genome transfer remains apparent in many data sets [131]. As high-throughput molecular technologies allow us to explore the genome at finer molecular resolution, emphasis is placed on filtering and normalizing data so that signal detection is maximized at the expense of variability. In this way, most investigators hope to find common, clinically-relevant mutations in cancer causing genes, over the variable background of intra-tumor genomic noise. Rather than filter or minimize the effect of such genomic variability, our approach for collecting and reporting genomic data adopts concepts from information theory and calculates variability on the basis of whole chromosome copy number data, translocations, minute chromosomes, deletions and structural defects. Using these inclusive criteria, a striking difference is seen in the karyotype variability between day 170 cells and later stage cells, between allografted cell lines and harvested tumors, and between

spontaneously transformed versus *Brca1* conditional knockout MOSEC. The karyograph format provides a clear picture of cell-to-cell variability for comparison with the variability induced by the conditional inactivation of *Brca1* in this model.

Quantification of genomic heterogeneity over time provides significant challenges. Nevertheless, we sought to characterize and quantify genomic heterogeneity throughout spontaneous transformation, with the concept that heritable variability at the level of the entire genome system is the basis for evolutionary selection. Though not intuitive in the context of linear outcome prediction, we have chosen to incorporate the reporting of cell-to-cell genomic variation at multiple levels into our experimental design, as we believe it to be inextricably linked to the emergent properties of cancer cell populations. We encourage the development of high throughput technologies that comparatively will provide cellular resolution rather than molecular resolution. Population profiling of transformation, tumorigenesis, drug resistance, and variable responses to therapy will be more easily employed. We have provided a basis for quantifying karyotypic heterogeneity in cancer cell populations which is intuitive and describes the extent of genomic variability by whole chromosome, derivative chromosome, and as a total population index. Taken together, these data demonstrate the significance, methodologies and rationale for quantifying karyotype heterogeneity in transformation, tumorigenesis, and clinical cancers, with hope towards incorporation of these features into prognostic and therapeutic applications.

The mechanism leading to such genomic mosaicism is, for most clinical cancers, not as significant as the magnitude and extent of this phenomenon. Cell-to-cell variability within a population of cells in culture, or from clinical tumors has been estimated in a number of ways by different groups [72]. Some algorithms for quantifying chromosomal or karyotype instability are based on the mechanisms by which they are thought to arise. Most support two sets of calculations for the contributions of whole versus translocations or marker chromosomes. Recently, in an set of papers by Castro et. al., the concept for quantifying inter-tumor karyotype heterogeneity among data from the Mittleman database was put forth using concepts based on

Shannon Entropy [80]. By this measure, tumor heterogeneity among samples was paired with data from SEER and showed that lower diversity of total chromosome number across samples correlated with better survival statistics. Additionally, the stochastic nature of chromosome change was shown by the relationship between global spread of numerical and structural chromosome abnormalities with sample number. If chromosomal instability is a necessary and ongoing characteristic of solid tumors, continuing to increase sample size in search of recurrent patterns of derivative chromosomes linked with cancer phenotype tumors [132-138] seems illogical. Independent of the mechanism, if random aberrations are produced at an increased rate and population size increases, some aberrations may appear more frequently if they confer a relative increase in fitness in that environment [125, 139].

The clonal selection theory describes the dynamics of a population of cells that, after undergoing random mutations, environmental pressures eliminate the least viable genotypes. If we assume that viable mutant phenotypes are rare, the degree of karyotype heterogeneity would be expected to decrease over time [140]. Our data, most clearly shown by the 50-cell karyographs characterizing each transformative stage or condition, suggest that unique and viable mutant genotypes are much larger in number than previously expected. By our measure, each of the fifty cells is unique by SKY analysis once it has diverged from the diploid genome, providing evidence for at least 50 different viable phenotypes per flask. As the literature suggests, a significant amount of additional variability likely exists in each cell that is too small to be detected with SKY analysis [141]. Our SKY data show that more than half (25/50) of the karyotypes are unique even in the earliest stages of transformation. When combining this finding with the concept that clinical tumors are detected when comprised of approximately 12 billion cells, we suggest that change at many levels including the whole chromosome, derivative chromosome, indels, loss of heterozygosity, and base mutation contribute to clinical tumors in which each cell, by its genome, is most likely unique. Although some may consider the measurement of genomic heterogeneity to be without clinical significance, we suggest that the

opposite may be true. This is because the resulting mosaic of cellular genomes provides the large-scale variability by which different phenotypes can be selected by drug treatment, or can adapt to a new environment during metastasis. By acknowledging the rate and magnitude of the heterogeneity formed in spontaneous versus gene-based transformation models (such as the conditional *Brca1* conditional inactivation model), we demonstrate the necessity for measuring karyotype heterogeneity and stability of cancer cell lines as outcomes variables in a wide variety of cancer research projects. These data suggest at least in the case of *Brca1* inactivation, this genetically induced model of cellular transformation is significantly different in terms of its ability to generate heterogeneity. As the variability seen in our allografted cell line and in the work of others [142] affords cell populations the ability to readily undergo rapid phenotypic shift, it seems that karyotype data should be included in experimental design and data interpretation using unstable cell lines.

The findings of a small Shannon Index value for chromosomes within the day 170 cells was not surprising considering the small variation in count for each chromosome. From day 245 onwards, the increased cell-to cell variability as indicated by the rise in the Shannon Index for each chromosome and for summed whole and aberrant chromosomes within that transitional stage, suggests an extremely low fidelity process for genome transfer. The small but significant reduction in Shannon Index values with the onset of several clonal chromosome translocations suggests decreased variability from the previous passages. As a unique event occurring in the late stage cells, we conclude that cells with this t4;3 translocation most likely have descended from a common cell, and now account for ~30% of the cell population. Interestingly, when these cells were isolated into subpopulations the variability of the remaining chromosomes was greater in cells with the 4;3 translocation than in those without, implying that selection may have occurred for this translocation, but low fidelity transfer of the remainder of the genome persisted.

Clonal and sub-clonal evolutionary processes have been previously described in induced cellular transformation [143-145]. In the current model, karyotype change is tracked

beginning from a normal diploid cell as might occur in a typical cancer patient without an inherited aneusomy or known heritable gene linked to increased cancer risk. In our spontaneous transformation model, cells demonstrated a weakly positive tumor phenotype after 245 days, and displayed high-grade tumorigenesis in 100% of injected animals at day 528. Our data support that by whole-genome analysis, clonal and sub-clonal evolutionary processes as described in the literature do not occur. When chromosome copy number changes, data for derivative chromosomes, and phenotypic data are taken into account, the evidence for clonal and/or sub-clonal populations of cells is largely nonexistent. However, when measuring clonality by emphasis on a single translocation and ignoring other variability, it appears ~30% of the day 450 cells may have originated from a common predecessor cell as marked by the 4;3 translocation. Background heterogeneity of the cells with the 4;3 translocation was the same as those cells without the translocation, suggesting that this marker of clonality may be fairly insignificant on the background of genomic variability.

The reason why clonal expansion is more clearly defined in other models probably extends beyond simple differences in reporting methodologies. Typically, induced transformation requires colonies to be picked that are all directly descended from a single cell. The fastest growing are the most likely to be picked and, as the cells in that colony recently descended from one cell with the same genomic disruption leading to its increase growth rate is more likely to be clonal. Other studies measure lower numbers of cells, typically only ~10 metaphase cells which are then subjected to pre-analysis filters to remove events with frequencies less than 20%. The concept that heterogeneity in the current study is created by artifact is unlikely based on the uniformity of the day 170 cells and the consistency in median gains, losses, and distributions of specific chromosomes over time. Few have studied karyotype change over time, but it has been suggested that following induced transformation and over 60 passages in culture, newly derived clones with “quazi-stable” karyotypes could be maintained in culture for 10 subsequent generations [146].

4.5 Genomic variability in spontaneous versus genetically induced models for transformation

By comparing the genomic mosaicism of MOSEC with conditional inactivation of *Brca1* (*Brca1*$^{\Delta 5-13}$) at day 79 with spontaneously transformed MOSEC, we provide insight into some of the unexplained features of clinical and laboratory findings relating to ovarian cancers in patients with inherited *BRCA1* mutations. Previous attempts to transform human ovarian epithelial cells by similar means have failed to spontaneously generate tumorigenic cells. Therefore, we find the MOSEC model well suited to investigate the role of genomic instability in spontaneous transformation and tumorigenesis.

Additionally we can shed some light on the confusion between genes as causative agents for cancer, and the difference between micro and karyotypic macro-scale evolutionary change. Among carriers of the *BRCA1* mutation with family histories of breast and or ovarian cancer, the lifetime risk of developing ovarian cancer has been estimated to range from 11-66% [147-148]. This six-fold spread in the range of ovarian cancer risk has been attributed to a number of environmental factors and detail regarding specific gene mutation. The current findings of rapid and widespread genomic mosaicism shown by the *Brca1*$^{\Delta 5-13}$ MOSEC suggest that environmental and other factors likely all contribute to promotion of the malignant phenotype by selection from a genomically heterogeneous cell population. The data also may explain why, in contrast to sporadic disease, *BRCA1*-related familial ovarian cancers are more frequently multifocal, with genetically distinct clones involving multiple sites [149]. The inherent heterogeneity of these *Brca1*$^{\Delta 5-13}$ cells gives rise to the possibility that certain karyotypes may be well matched for specific microenvironments or niches within the peritoneum in which the local environment would provide the basis for selection. The findings that the Shannon Index for each chromosome in the *Brca1*$^{\Delta 5-13}$ MOSEC was significantly elevated over the most variable of the MOSEC stages suggests that genetically distinct subtypes may arise more frequently and to more rapidly adapt to new environments (multifocal nature, drug resistance,

progression). As a functional Brca1 protein is required for reliable DNA double strand break repair by homologous recombination, we expected increased non-homologous end joining events [150-151] to be marked by increased translocation frequencies. Increased translocation frequencies were seen to some extent, but low numbers of translocations prevented thorough analysis. This phenomenon may become more prominent with increased time in culture following conditional *Brca1* inactivation. The data suggest that a large increase in whole chromosome counts due to improper segregation events precedes the emergence of instability caused by translocation events. It has been suggested that the gene expression profiles in sporadic ovarian cancers have been reported to fall into two categories that resemble the ovarian cancer gene expression from patients carrying either *BRCA1* or *BRCA2* mutations [152]. This seems to conflict with the early widespread heterogeneity documented in the $Brca1^{\Delta 5-13}$ line, but may be resolved by the different selective pressures created by differing environments and clinical phenotypes of *BRCA1* versus *BRCA2* tumors compared to early stage cells *in vitro* that have a relatively homogeneous and constant cell culture environment. These data can be used as an example of how microevolution can facilitate and predict of the onset of macroevolution in certain types of cancer. *BRCA1* and likely other widely accepted oncogenes or tumor suppressor genes have profound and swift impact on genomic instability compared to spontaneously occurring transformation. This demonstrates a link between micro-evolutionary alteration in a single gene and the jump to macro-evolutionary large scale rearrangement of the entire cellular genome. The current data support a theory for cancer causation by genetic alteration that occurs through widespread genomic instability. This instability begins as early as day 40, far ahead of the emergence of specific traits of the cancer phenotype supposedly linked to growth or invasiveness. Conditional inactivation of *Brca1* in MOSEC in situ by AdCre recombination within the bursa results in the development of preneoplastic changes, such as hyperplasia, epithelial invaginations, and inclusion cysts but failed to yield epithelial ovarian cancer even after a one year follow up. The lack of expression

of known human ovarian cancer genes suggests that *BRCA1* mediated transformation may occur differently in mouse versus man. Additionally, it is possible that the time was not sufficient for the pre-neoplastic changes to develop into tumors [48].

The findings that certain chromosome counts remained relatively stable (e.g. chromosome 1, 12), while others were highly unstable (e.g. chromosome 2, X) is noteworthy, but the mechanisms and significance of these findings is not well understood. The differential variability of certain chromosomes could be related to an inherent stability of each chromosome in the segregation or selection process. In this case, if the experiment were repeated, as in the *Brca1* conditional knockout model, these may again emerge as most and least stable. The second and more likely possibility is that an early random expansion or imbalance in copy number increases the likelihood of persistent or worsening change. It is possible that the variability seen in chromosome count reflects the ability to regulate gene expression from genes on that chromosome. This is supported by data from the X chromosome, which consistently displayed the greatest amount variability in whole chromosome copy number. This hyper-variability may relate to its ability to its ability to repress transcription from the entire chromosome as a Barr body, a phenomenon that is not appreciated when chromosomes are condensed at metaphase. However, several studies have found that ovarian cancer cells lack a normal inactivated X chromosome (X_i) and manifest at least two active X chromosomes (X_a), suggesting a selection for multiple X_{as} that may be linked to over expression of X-linked genes [153-155]. These findings regarding dysregulation of heterochromatin at the X chromosome may mark a broad dysregulation of the nuclear hetero-chromatic compartment responsible for global perturbations of gene-expression promoting tumorigenesis.

4.5 Array analyses for gene expression

The majority of short time series expression data analyses do not utilize the scale and sequence information inherent in time-series data sets. Some of the typically used methods for analysis are hierarchical clustering [156], *k*-means clustering [157], self-organizing maps [158],

and sets of paired comparisons [159]. It is important to understand that these commonly used methods do not test for differential expression based on event sequence, and so random permutation of the order of time points would not change the results of these analyses. As phenotype data from the current study clearly suggest a progressive transformation over time, a more appropriate and robust method for time course data analysis was performed using extraction and analysis of differential gene expression (EDGE) [92] coupled with time-course based cluster analysis with gene Ontologies using the short time series expression miner (STEM) [95] software. Nearly 600 genes were identified and clustered by profile according to each gene's fold-change pattern of significant differential expression over time. This analysis was ideally suited for the sequential analysis of key transitional stages, as the time between key transitional stages was relevant to the accumulation of changes in cell phenotype, but time points were not equally spaced.

4.5.1 Longitudinal Gene Expression Analysis Suggests that Six Hundred Genes and a Variety of Fold Change Profiles Play a Role in Tumorigenesis

Tests for functional enrichment of the genes linked to each profile were significant for Gene Ontologies annotation categories in two of the eight profiles. Differential expression in profile 2 was characterized by increasing transcript abundance of ~3-4 fold (170:0; 245:3; 450:4; 528:3) and this profile was enriched in functional categories concerning chromatin, DNA packaging, conformational change, chromosome and nucleosome organization, RNA processing, macromolecular complex assembly and subunit and spliceosomal complexes. This profile summary fits directly with the dramatic changes seen in chromosome copy number and transcription profiles throughout transformation and may be particularly relevant to trends in chromosome number and distribution as well as in karyotype heterogeneity and altered differential gene expression patterns over time. The second significant profile (5) shows a trend of increasing transcript abundance across all transformative stages (fold change: 170:0; 245:1; 450:2; 528:3). This profile was significantly enriched for cholesterol biosynthetic and metabolic

processes and genes associated with microsomal and vesicular fractions. A wide range of integral and lipid membrane associated protein types are contained within the microsomal and vesicular fractions including pores, channels, pumps, transporters, trans membrane receptors, and cell-adhesion proteins. The functional enrichment of genes linked to cholesterol biosynthesis and metabolism support emerging evidence of off-target anticancer effects of commonly prescribed Statin drugs (HMG-CoA reductase inhibitors) [160]. By reducing available mevalonic acid, this class of drugs exerts pleiotropic effects including blockade of the G1-S transition effecting cell proliferation and on many essential cellular functions including differentiation, survival, and the regulation of cell shape and motility [161-162].

Pair wise analysis using significance analysis of microarray (SAM) identified specific lists of differentially expressed genes for comparison with morphological and behavioral change. The largest changes in mRNA transcript abundance occurred in the early transition (170:245), which parallels the largest shift in karyotype diversity and chromosome count. The day 245:450 mid transition had relatively small changes in expression despite over 200 days of continuous culture. Phenotypic change included the acquisition of a diffusively invasive phenotype and a large increase in cell branching morphology in 3 dimensional culture. Day 245 cells produced low-grade locally confined tumors at low frequency (1/14 sites). Significant functional enrichment for GO categories, included categories relating to the extracellular space/ matrix, glycoprotein related cell signaling, and changes in cell binding and adhesion. These gene expression array results are concordant with previous work by Roberts et. al. [55], who used immunohistochemistry on various MOSEC transformative stages and demonstrated alterations in focal adhesion complexes and cytoskeletal elements. These categories for functional enrichment generally are well matched with changing MOSEC phenotype including increased growth capabilities, acquired focal invasiveness, and the loss of contact inhibited growth. Approximately half of the differentially expressed genes determined in each SAM pair wise analysis overlapped with genes detected using STEM time series analysis. Genes in the null

set (Φ) are significant for that specific transitional stage, but are outside the time course analysis. As fold change cut offs were the same across both types of expression, Φ genes likely reflect the increased power to detect change following normalization only for that specific pair wise analysis.

4.5.2 Driving genes in transformation as anti-cancer targets

Due to the large numbers of studies attempting to find cancer causing genes through the high-throughput screening of large numbers of patient samples, we sought to determine the rate at which early up-regulated genes were detectable in the transformed late stage cells. To identify these putative "driving genes" in transformation, the intersection of gene lists generated from SAM pair wise analyses were evaluated. Three lists were generated that consisted of genes from the intersection of each paired comparison (170:245, 245:450, 450:528). The number of differentially expressed genes on each pair-wise analysis list became smaller as transformation progressed. This suggests that later stage samples are more closely related to each other than earlier stage samples. Accordingly, genes found in the intersections of these were few in number (170:245=16; 245:4501=6; 450:528=2). In fact, only a single gene, Necdin (NDN), met criteria for significant differential expression across the three paired comparisons. NDN shows current major associations with neuronal migration and the human disease Prader-Wili Syndrome. A lone entity among 600 genes, the role of Necdin in cellular transformation is developing and others report NDN as a target of P53 with a role in hematopoetic stem cell senescence. Combined with evidence for NDN-mediated down regulation of *p53* and other diverse roles for NDN in tumorigenesis, data from the current study provide further rationalle for continued investigation of NDN as a potential tumor suppressor gene [163].

4.6 Chromosome Copy Number Change Influences Gene Expression in Late Stage Tumorigenic Cells

Karyograph and Shannon Index data provide qualitative and quantitative evidence for differential chromosome instability and karyotype heterogeneity across the key transitional stages. The concept of, “gene-dosage effect” is familiar to most geneticists, as is the notion in cancer genomics that that gene-dosage might contribute to the malignant phenotype. As of yet, gene regulation by this mechanism has not been evaluated across various transformative stages. In the early transformative stages (170:245) chromosome count (5% per chromosome) is not related to mRNA transcript abundance (range: 3-8% per chromosome). However, when this phenomenon is tracked over time, the strength of this relationship becomes less and less clear. Fold change in chromosome copy number and mRNA transcript abundance are largely unrelated, reaching a minimum between 450:528. Unlike our late transition findings between day 450:528, a close relationship between chromosome copy number and transcriptome has been reported in sub-clone derivatives of previously established cancer cell lines in culture [164]. Oppositely, in the case of chromosome gains on the background of a normal diploid karyotypes, such as human autosomal trisomies, dosage compensation for individual genes can be stage- and tissue-specific [84, 165-166]. The susceptibility of cells to gene dosage effect or their ability to compensate for altered gene dosage is mediated presumably by feedback at the level of transcription or mRNA stability; however, similar translational and post-translational effects might exist that would not be detected. Gene dosage effect by whole chromosome count is not a feature of the late transitional stages of spontaneously transforming MOSEC (245:450:528), as chromosome copy number and gene expression are unrelated. Subcloned MOSEC lines recently derived from a single transformed progenitor may exhibit such features, however this did not occur under the spontaneous transformation model used in the current study. The current data are unique in that they resolve the discrepancy between gene-dosage effects in normal diploid versus transformed cells by taking a whole genome multi-level approach to understanding cancer genomics. The data suggest that gene-dosage effect may be a characteristic of early stage cancer cells and that the relationship may exist when

chromosome copy first depart from diploid status and vary over a large range from one transformative stage to the next. This provides further support for the theory that karyotype based evolution is significant to cancer cell populations in that low fidelity genome transfer and high variability from cell to cell is linked to dysregulation of mRNA transcript abundance and uncoupling of the gene-doseage effect [164]. Karytype evolutionary patterns that have previously been linked to a tight genome: transcriptome relationship have significantly less variability from cell to cell and less change in chromosome copy numbers over all.

4.7 Linking Changes in Tumor Karyotype Profiles to Environmental Shift

By continuing to evaluate the key transformative stages during *in vivo* tumorigenicity assays, several interesting findings emerged. The weak tumorigenicity of the day 245 cells that were allografted into C57BL6 mice was expected based on *in vitro* data and link relatively low Shannon indices to low tumorigenicity. The findings that day 528 cells formed subcutaneous and intraperitoneal tumors of larger, and higher grade than day 245 cells concur with previously published characteristics of allografted cells using this model system [54]. These tumorigenicity rates for late stage MOSEC are somewhat lower than previously reported rates [55], but match the tumorigenic potential of these cells as assessed *in vivo.* The invasion and proliferation assays from the current study are slightly less aggressive at all transformative stages despite longer transformation times [55], which could be attributed to the small number of seeded and early stage mitotic cells. Considering the genomic instability and karyotype variation in this longitudinal transformation model, previous “runs” beginning from primary cells could likely have different end phenotypes. The variable karyotypes of the late stage injected MOSEC and the significant phenotypic shift that occurred between cell injection and harvest, could account for the variable tumor formation. Based on the findings that both tumors contain the same four to five high-frequency translocations these cells have likely descended from a common predecessor. It is unlikely that this translocation itself increases tumorigenicity as in the

BCR:ABL fusion protein of CML, but rather marks a recognizable feature of a genome that at one point had a selective advantage over the other cells. The significance of these markers may be interpreted in several ways. Perhaps this set of markers represent a genome that facilitated survival in the initial transition from *in vitro* to *in vivo* environments, or if these markers were accumulated in a stepwise fashion and reflect cellular adaptations required to survive as a larger and progressively more invasive tumor. Additionally, the contributions of gene alterations associated with these visible markers versus gene alterations invisible to SKY analysis is beyond the scope of the current studies. This could be interpreted to mean that this clonal set of markers are part of a very large set of genomic, epigenetic, environmental, and population based variables that facilitate positive selection for the tumorigenic phenotype. Alternatively, it is possible that this set of karyotype markers are unlinked to phenotype and that they independently arose in both tumors, and as multiple independent events in the day 450 cells. If this were true, we suspect that fragile sites or other inherent features of the genome could account for the increased frequency and repeated appearance of this set of markers.

A role for the immune system in suppressing, selecting for, or eliminating injected tumor cells from the peritoneal cavity is of particular interest as ovarian cancer cells evade the immune system by creating a highly suppressive environment in the peritoneal cavity. A recent clinical cohort study in of 500 ovarian cancer patients correlated intraepithelial CD8+ T-cells with improved clinical outcomes for all stages of ovarian cancer [167] brought importance to the study of immune system based therapies for ovarian cancer. The immune function of the C57BL6 mice was evidenced by the active aggregates of mesothelial cells with inflammatory cells seen in association with peritoneal fat. Although subject numbers in the current study are small, the larger tumor burden in the nod-SCID-gamma mice and the occurrence of invasive tumors in 2/2 animals supports our current knowledge of immune regulation in tumor establishment and progression. Importantly, these data suggest that the MOSEC model is well suited for ovarian cancer studies focused on the dynamics between the immune cells,

cytokines, and other regulatory molecules in cancer progression in the context of vaccine or other immunologic therapies.

The population karyotype data from injected late-stage versus harvested tumor cells provides valuable insight into the possibility of cell selection and the significance of variability and low fidelity genome transfer in tumorigenesis. The two tumors evaluated are from different mice injected with the same late-stage cells. However, both share the 4;5 translocation at the same frequency as the day 450 cells, but both display two additional translocations at frequencies greater than 90%. Based on the rates and patterns of these aberrations, it seems more likely that the environmental shift selected for a cell subtype that was present at low levels, rather than the set of translocations arising independently in both mice. The former theory seems more likely, although the latter cannot be completely ruled out. Finally, the high degree of complex non-clonal translocations in tumor two (T2) compared to tumor one (T1) may be linked to tumor histopathology. The size and growth rate of T2 was significantly elevated over T1, is likely responsible for the large necrotic zone seen on histopathologic analysis. Necrosis is a well known response to oxygen debt and is related to tumor size and perfusion, with cells adjacent to the necrotic zone subjected to low oxygen tensions [168-169]. Hypoxic cells continue to divide and, as they slowly proceed through the cell cycle, increase genomic instability and promote tumorigenesis [170]. These data warrant further investigation as they link genomic instability with hypoxia, a phenomenon involved in radio-resistant tumors. The concept that random chromosomal change in response to hypoxia could increase variability and provide the basis for rapid evolution and phenotypic shift suggests a low fidelity of genome transfer in the larger, hypoxic tumor, suggestive of a genome based, rather than gene based mechanism for hypoxia in resistance to chemotherapy and radiation treatments.

4.8 Future Directions and Clinical Considerations

The findings of the current study raise several basic and clinical scientific questions regarding karyotype heterogeneity and many complications known to clinical oncology including metastasis, drug resistance, and immune system evasion. Both the prognostic significance of chromosomal instability and its facilitative role in drug resistance and metastasis suggest that genomic instability may likely be a modifiable factor in clinical cancer outcomes. We anticipate the key future challenges will lie in the determination of how genomic instability based therapies can be incorporated into the current clinical standards for cancer treatment. The elucidation of the mechanisms underlying aneuploidy and genomic instability is only one piece of this puzzle. More important is to determine how our interventions on genomic instability change cancer cell evolution and clinical outcomes. Our analyses suggest that the observed karyotype heterogeneity differs with key transitional stages of transformation, but leaves several unanswered questions. These include our ability to alter genomic instability of cancer cells *in vitro and in vivo* with pharmacologic intervention and the outcomes of these alterations on cell phenotype and clinical outcome. One mechanism by which genomic instability could be exploited for treatment would be by inhibiting mitotic-checkpoint signaling. This has been shown to be lethal in unstable cells as the consequence of massive chromosome loss [171]. Genomic instability additionally holds promise in the design of adjuvant chemotherapies that minimize new mutations that allow them to develop resistance to other cyto-toxic agents. The ability to identify and interpret information contained in measurements of cellular heterogeneity will, within an evolutionary framework for cancer progression, provide insight that is not attainable by traditional analyses of genomic material. This is because traditional analyses are performed on averaged over a heterogeneous population.

In conclusion, early genomic instability generates variation in tumor cell populations well before the onset of the transformed phenotype. By reporting complete karyotype data for a large number of metaphase cells at key transformative stages, the contributions of population karyotype dynamics in tumorigenesis can be evaluated. The fact that gene-dosage effect by

changes in whole chromosome copy number are most significant in tumorigenic late-stage cells suggests a mechanism by which genotype is differentially linked to global gene expression profiles between normal and cancer cells. The magnitude and early onset of genomic heterogeneity in the conditionally inactivated *Brca1* MOSEC demonstrate how a gene linked to increased cancer incidence can predictably induce rapid and widespread genomic instability at a much faster rate than the spontaneously transformed line. The widespread instability of this population suggests that targeted therapy will likely result in selection, then expansion and progressive increase in the heterogeneity of resistant cells. Finally, the late stage cells with heterogeneous karyotypes are shown to undergo rapid karyotypic shift from the profile of the injected line to a profile of increased variability following environmental shift. Together these multi-level longitudinal data support genomic instability and increasing cell-to-cell variations in karyotype as a means to increase population diversity and permissive for rapid adaptation by gene-dosage effect in late stage transformed cells. By understanding how cancer cells continue to generate heterogeneous karyotypes versus commit to senescence or cell death pathways support will be provided for the treatment designs affecting this common feature of all clinical cancers. In the science of cancer complexity, a shift away from observational techniques to the application of quantitative measures for genomic heterogeneity is established. Like other genome-wide techniques in molecular biology, the continued development of high throughput machinery will benefit clinicians and investigators in the field by permit the quantification of genomic heterogeneity by streamlining the data acquisition process while maintaining cellular resolution. The universal features of within tumor genomic variation and cancer cell evolution can be evaluated in a wide variety of clinical and research settings and the outcomes applied to improve cancer detection and treatment strategies.

5. THEORETICAL CONSIDERATIONS

Since the War on Cancer was officially declared nearly four decades ago, the United States has spent over $1.5 trillion dollars on cancer research and treatment. A large amount of data has been generated through the development of high throughput genomic strategies, multi-center collaborative efforts, and the exploration of new therapies including immunologic [172] and gene based [173] therapies. Despite these advancements, overall cancer mortality rates have only fallen by approximately 1% since 1975 [174]. With some rare successes in hematological cancers [175] and in cancers where inheritance or environmental exposure increases surveillance and early prophylactic removal of at-risk tissues [47, 176-177], the outlook for most people diagnosed with cancer today is not much different than it was a generation ago. Therefore, a critical review of the assumptions inherent in the established paradigm currently dominant in cancer research is provided.

Understanding the evidence for the current paradigm and the accumulation of anomalies against it is of utmost importance, as scientific paradigms imbue the minds of researchers in the field and influence numerous aspects of how science is conducted, how experiments are designed, and how data are interpreted. The established paradigm defines genes as primary agents in cancer causation and describes a stepwise clonal evolutionary process where gene mutations account for the progressive acquisition of the hallmarks of the clinical cancer phenotype. Oppositely, the proposed paradigm emphasizes wide-spread heterogeneity of cancer cell genomes within a given population. Rather than focus on individual genes and their direct linkage to specific acquired characteristics of the cancer phenotype, the new paradigm considers genomically unstable cancer cell populations as complex-adaptive systems that are not governed by linear genotype-phenotype relationships. The proposed paradigm places greater significance on the clinical utility and predictive power of gene-level change on a normal diploid genomic background, and acknowledges it as “micro-evolution.” Cancer cell populations evolve by a novel mechanism that is central to the proposed paradigm. This rapid

and widespread genomic reorganization can be monitored within a cell population and is termed “cellular macro-evolution.” Central to this paradigm shift are several pertinent conceptualizations surrounding the magnitude and extent of the karyotype heterogeneity discovered within random samplings of cell populations at key stages of cellular transformation and tumorigenesis. These include current working definitions of clonality in cancer, cancer cell populations as complex-adaptive and non-linear systems, microevolution versus macroevolution, cellular versus molecular resolution, genotype-phenotype linearity in stable versus macro-evolutionary environments.

Paradigm Shifts in Cancer Research

Over one hundred years ago, chromosomal aberrations were linked to cancer. At first this link was based on observational data by Von Hanssemann [126]. Soon after, Boveri reported the growth of tumor like structures in sea urchins after manipulating their chromosome numbers and inducing multipolar mitotic events. This work resulted in the formulation of a theory that implicated unequal mitotic events and the resulting chromosome aberrations that ensued in the formation of cancers [178]. The significance placed on gene-level change by the established gene-centered paradigm is rooted in a series of exemplar experiments and has maintained momentum through *in vitro* experiments and a handful of clinical successes. Each of these experiments involves gene alteration of an otherwise normal genome followed by selection for the transfected cells with strong growth phenotype. Each exemplar experiment which initially implicated a specific gene or genes in cancer causation has subsequently been tied to genomic instability and karyotype abnormalities [179][180]. For example, the discovery of the first gene (Src) that could cause cancer in normal animals [181], was soon followed by publications demonstrating a clear role for Src in genomic instability [179]. Similarly, in Knudson’s 2-hit hypothesis for inherited versus sporadic cases of retinoblastoma (Rb) [182], Rb was reported to cause mitotic recombination errors that lead to homozygosity of the inherited mutation [180],

and disruption of the Rb pathway was linked to mitotic spindle checkpoint disruption and aneuploidy [115]. "The clonal evolution of tumor cell populations" describes an evolutionary view of cancer as being caused by gene mutation mediated increases in variability that are subsequently followed by selection or clonal evolution of tumor cell populations tracked by signature karyotypic features [143]. Nowell had recognized the potential for inter-tumor variability by studying karyotypes cell-by-cell with relatively simple techniques. According to the current data, which demonstrates unique karyotypes for each cell in a randomly selected population of 50 metaphases, the magnitude and extent karyotype heterogeneity has been grossly underestimated, as evidenced by the co-existence of numerous highly variable karyotypes from actively dividing cells [143]. Banding techniques for karyotyping developed in the 1970s led to discovery that specific chromosome aberrations could be linked to gene-based causes for cancer. This shift from the microscope to the molecule marked the beginnings of the established gene-centered paradigm that currently dominates the field of cancer research.

The gene-centered paradigm for cancer research regards cancer as a disease that is caused by the stepwise accumulation of gene mutations and of clonal expansion from a single cell. This paradigm focuses on the one-to-one relationship between alterations in cancer causing genes, and the acquisition of key characteristics of the cancer cell phenotype as the mechanism for cancer progression. The predictability of phenotypic change resulting from alterations in cancer causing genes and the linear progression to reach the transformed phenotype highlight a major set of assumptions contributing to the gene-centered nature of the established paradigm. Two theories have provided the foundation for the current paradigm that are not mutually exclusive: the theory of gene-mutation and the theory of somatic evolution for clonal cellular transformation. Gene-mutation theory states that cancer progression is driven by sequential somatic mutations in five to ten specific genes, each of these accounting for a specific facet of the cancer cell phenotype [183-184]. This theory is also gene-centered, in that change at the level of entire chromosomes or karyotypes is considered epiphenomena of the

altered growth and metabolism. The second theory underlying the established gene-centered paradigm is the theory of somatic evolution for clonal cellular transformation, which explains tumor progression as the result of acquired genetic instability and the sequential selection of variant subpopulations, the most advantageous of these leading to clonal expansion and transformation. These theories are synergistic, and when integrated support the gene-centered paradigm by describing cancer progression according to a linear path of Darwinian microevolution, where clonal selection due to specific gene mutations is responsible for the conversion from the normal somatic cell phenotype to the eventual death of the host organism.

Dissidents of the current paradigm suggest that careful examination of the existing data regarding genetically linked oncogenic transformation contains a significant number of anomalies. The first is that transfection with sets of oncogenes results in transformation *in vitro* in a very small minority of cells (1:100,000), [3, 185]. The second anomaly is that early transient gene alteration is sufficient to increase the rate of transformation, but is not required in later stages of progression where genotype-phenotype alterations are purported to continue to contribute to progression towards the malignant phenotype. Following these early transient gene alterations, genomic instability emerges and persists into late stages despite a loss of the corresponding gene products [114, 186-189]. The third anomaly pertains to inherited cancers and transgenic animal models of cancer causation. When cancer arises by these means, the gene mutation supposedly causative for oncogenesis imbues the entire organism. However, such inherited mutations result in cell-type specific cancers in a small number of cells and frequently arise after a long time period despite the equal presence of cancer gene alterations in all cells [186-188, 190]. An example of this is in human cancers is in cases of inherited *BRCA1* mutation. Ovarian cancers arise only in some women carrying the mutation, is epithelial cell type specific, and typically arises at approximately 50-60 years of age [191]. For reasons beyond our current understanding, *BRCA1* mutations additionally increase the risk for breast,

fallopian tube, prostate, and a subset of hematologic cancers, but do not similarly effect other organs [192].

In response to the above questions of the gene-centered approach, the proposed paradigm is centered on genomic heterogeneity of cancer cell populations and places gene-level alterations in a diminutive role and is therefore termed, “genome-centered”. Despite a large body of evidence that supports the genomic instability in cancer causation, the popularity of this theory has been surpassed by gene-centered viewpoints in the last 25 years [83, 122-124, 128-129, 139, 145, 193-202]. Numerous classes of non-genic causes of cancer have been reported and include environmental insult by non-genotoxic environmental carcinogens [193, 203], and failures of replication timing [204], condensation, cytokenesis, kinetechore migration, centriole duplication, and spindle formation [205]. These mechanisms all destabilize the genome and result in a massive increase in genomic variability through the unequal distribution of chromosomes among each generation of daughter cells. The new “genome-centered” paradigm differentially defines the role of gene level changes before and after the major rise in population diversity by the generation of new and variable genomes. The paradigm defines gene-level alterations as micro-evolutionary change which can predictably contribute to cancer causation through increasing the likelihood of genomic instability in stable, near-normal cells. Under the genome-centered paradigm, gene level alterations are much less predictably related to cell phenotypic change once population diversity has risen, but may be linked to phenotype change through karyotypic shift.

The genome-centered paradigm defines the genome context as follows: *The entire genome in its orientation including all heritable modifiers of that genome*. This is significant as the proposed paradigm places theoretical and functional emphasis on the genome context (rather than any of its individual constituents) as the main platform for evolutionary selection [86, 206]. The conditions required for natural selection are well known to evolutionary biologists and are entirely dependent upon genetic variability within the population of cells. These conditions

include 1) the existence of variation in the population, 2) that variation must be heritable, and 3) that variation must affect survival or cell fitness. Based on these tenets, the proposed novel theories for cancer biology similarly regard variability as the universal feature of success for cancer cell populations as such variation facilitates rapid phenotypic shift of cell populations and rapid adaptability to new environments such as those encountered by metastatic or drug-treated cells. Spontaneous and inherited clinical cancers, cancer cell lines, and tumors arising by oncogene transformation universally display aneuploid karyotypes and genomic instability [146, 186, 207-211]. The association of increased population fitness with increased genetic diversity has been described in cancer cell drug resistance [87, 142, 212], and is linked to decreased patient survival times in clinical cancers [213]. On the basis of these findings and the results of the current data set, it seems the gene-centered paradigm and the step-wise clonal expansion theory for cancer causation place the investigator on a certain observational plane. From this viewpoint, the extent and significance of high-level genomic variability as facilitators of rapid genomic and phenotypic switching are greatly underestimated. The genome-centered paradigm for cancer research places emphasis on the differential predictability of genetic cause and effect relationships. These are dependent on the genomic composition of the cell population. Specifically, the proposed paradigm calls for two sets of rationale, one of which can be applied to genomically stable cell populations that are homogeneous by karyotype and constrained by normal organismal function. The second set of guidelines is to be applied to genomically unstable cells comprising the majority of clinical cancers that display high levels of genomic variability facilitatory for large-scale macro-evolutionary change.

Clonal Misconceptions of the Gene-Centered Paradigm

Under the established gene-centered paradigm, a generation of cancer researchers has been indoctrinated with the belief that clonality is central to cancer formation. A systematic and self-renewing bias favoring "clonality" is propagated by the continued use of tools and

techniques that systematically reduce the measurement and reporting of intra-tumor genomic variability. The ramifications of introducing such a bias are reflected in our current working definitions of the clonality of cell populations in cancer research. For example, clonality among cancer cell populations is frequently concluded on the basis of a single point mutation shared among populations of cells. This working definition of clonality is readily accepted under the current paradigm, despite the unexplored sequence diversity and context of the other 5 billion bases comprising the genome [214-218]. The acceptance of cancer cell "clonality" on the basis of a single genetic marker is inextricably tied to concepts of classic Mendelian genetics, where the genotype: phenotype relationship was taught using a single locus and the resulting characteristic phenotype. Thus, our fundamental understanding and working definitions of the genome are rooted in linear single-gene "genotype: phenotype" relationships. The proposed paradigm shift dispels the belief that cancer cell populations are characterized by clonality and particularly emphasizes the inappropriateness of considering them as such in scientific and clinical settings.

The incorporation of, "clonal" into the vernacular of cancer cell researchers has not occurred by misunderstanding of meaning or definition of the term, but by the inappropriate frame of reference from which clonality is measured. By their strictest definition, clones are sets of genetically identical organisms. More commonly, clones are defined as a set of cells that share a common genotype owing to descent from a common ancestor, and genotype is defined as an organism's full hereditary information or genetic constitution including expressed and non-expressed DNA. However, high rates of genomic alteration across many levels of the genome have been reported and this property of cancer cells renders them incapable of high-fidelity genome transfer [141, 219-220]. Therefore, the term "clone" by this definition is not applicable to cancer cell populations as they are neither genetically identical to each other nor do they share the same full hereditary information. By the alternate definition of "clonal," meaning that a cell population has descended from a common ancestor, all cells comprising all normal diploid

and cancer cells within the same organism would be “clones” of one original zygote. It is important to recognize that the misconception of clonality in cancer biology is propagated by the very system used to describe cancer cells. Similar to gene-centered studies, the identification of recurrent or clonal chromosome aberrations in cancer cytogenetics is also systematically prioritized over random change. Specifically, the International System for Human Cytogenetic Nomenclature (ISCN) states that a clone may be marked by a *single* aberration occurring at a frequency of only 20-30% [221], “A clone may not be completely homogenous… It will always mean at least two cells (out of ten) with the same aberration.” The guidelines recommend non-reporting of this detail, allowing the author to provide his or her own operational definition of “clonal”, by which all other features of the population are described. Rather than acknowledge or quantify the fact that these cell lines are in a state of extreme genomic instability, the guidelines recommend, “Effort then should focus on describing the sub-clones so that clonal evolution is made evident… the composite karyotype can be created with all clonally occurring abnormalities combined into one karyotype” [221].

The ramifications of the continued search for clonality are significant to the field for several reasons. Unable to define a recurrent pattern in this multi-tiered unstable system of hyper-variable cancer genomes, many investigators have come to regard random karyotype change as insignificant and with limited diagnostic and prognostic value [222-224]. However, quantification of this universal feature of cancer phenotypes is of the utmost importance as genomic instability itself is known to characterize tumor development [86, 128, 206], immortalization, metastasis, and drug resistance [87, 142, 212]. In line with these concepts, the data presented in this work documents a high-degree of genomic variability by karyotype heterogeneity which occurs far before the acquisition of the tumorigenic phenotype. This variability supports a pattern of stochastic rather than stepwise progressive evolution. The dynamics of whole genome mediated macro-evolution by karyotypic heterogeneity (rather than micro-evolutionary change) is shown by population karyotype reports that trace micro-

evolutionary change in spontaneous transformation by measuring the patterns of karyotypic heterogeneity over time. The new, genome-centered model provides the opportunity to re-write the concept of clonal cellular evolution by documenting the timing and characteristics of macro-evolutionary karyotypic change in relation to the transformation process.

As the implications of clonality captured at a single locus or level are assumed to have strong biological and clinical relevance, the term “clonal” to describe cancer cells is not simply a matter of parlance, but a key term with wide-spread influence on experimental design and data interpretation [183, 225-226]. The rationale for this supposition is that the clonality determined by measurement at a specific marker is conceptually extended to encompass the remainder of the genome. Therefore clonal changes should be linearly predictive of phenotype. The accompanying data set demonstrates by quantification of complete karyotype data that subpopulations of cells with a specific translocation are no more homogeneous than the remainder of the cells comprising that population. Thus, the stochasticity of low fidelity genome transfer invalidates the assumption that any given genomic marker denotes relative genomic homogeneity within the marked sub-population. With this understanding, the significance of these markers is greatly reduced in the acute setting, and particularly when gene products of these markers are challenged with targeted therapies.

Reductionism and Holism in Cancer Genomics

The established gene-centered paradigm has been supported by data generated within the framework of a molecular reductionist approach to cellular biology. Specifically, methodological reductionism is the position that the best strategic approach to problem solving in science centers on reducing explanations to the smallest possible entities. This deductive strategy is inextricably tied to the gene-centered viewpoint in that the linearity of hierarchical genomic organization is assumed to transcend multiple hierarchical levels (phenotype, protein function, protein structure, amino acid residue, RNA transcript, DNA sequence). Multiple exceptions to

the rule of linear, one-to-one relationships assumed by the gene-centered paradigm exist. If cancer cell evolution proceeded by the stepwise acquisition of several gene-level alterations, then correction of a single imbalance should have only a modest inhibitory effect. However, the re-introduction of a wild-type tumor suppressor gene (for example, *Tp53*, *Rb*, or *APC*) into human cancer cells where the respective endogenous gene is inactive usually promotes severe growth inhibition, apoptosis, or near total inhibition of tumorigenesis [227], rather than the predicted equal and opposite modest inhibitory effect. A second example is that the continued over-expression of some genes (e.g. CCND1) seems critical for maintaining the cancer phenotype in several human cancer cell lines, as antisense CCND1 reversed the cancer phenotype towards normal. Unpredictable and non-linear changes were demonstrated in both the pancreatic and esophageal cell line, which additionally developed hypersensitivity to chemotherapeutic agents, and phenotype reversion with persistence of elevated CCND1 levels [228]. The misapplication of reductionist logic becomes increasingly evident when considering that genotype-phenotype relationship also assumes the intermediary cellular network of actively transcribed genes, their regulatory elements, and the function of their gene products to be normally regulated, predetermined, and therefore predictable. Considering that gene-level changes frequently result in variable and unpredictable phenotypes arising from complex combinations of genes and gene regulatory events, the assumption that predictability even of similar magnitude might persist is likely incorrect. Thus, the reductionist approach is more reasonably applied to predict interactions between components in small groups with clear margins, and it offers no mechanism for the determination of how the emergent and adaptive properties of these cancer cell populations come to be. From the genome-centered perspective, reductionist logic is misapplied because it underestimates the complex adaptive nature of cancer cell populations.

The underlying theme of complexity theory is that local interactions between parts of a system can lead to global properties that are neither linear nor proportional to the sum of its

parts. In assuming cancer cells to be complex-adaptive systems, a genome-centered holistic approach is favored over reductionist logic. Linearity and predictability assumed by the established paradigm underestimates the complexity of cancer cell systems because whole dimensions of interaction, regulation, and adaptability are overlooked. As complex-adaptive systems, cancer cell populations are characterized by innumerable multi-faceted and dynamic interactions at the level of molecules, genomes, cells, and microenvironments, each of these having characteristics of a non-linear system. Cancer cell populations also meet the criteria for complex-adaptive systems because of their ability to alter the strength of interactions within the network in a way that maximizes the average fitness of the cell population. Examples of emergent properties of cancer cell systems arising from complex local interactions include: the ability to organize spontaneously into different morphological patterns, the acquisition of drug resistance, and the rapid development phenotype alteration by karyotype shift [164, 229]. The significance of intra-tumor genomic heterogeneity relative to the individual cancer gene in facilitating these emergent properties has been demonstrated by data from our group and others [83, 86, 124, 128-130, 212]. The holistic approach taken by the genome-centered paradigm incorporates the concepts of natural and generalized Darwinian evolution with the regulation that cancer cell populations be considered complex-adaptive rather than linear and constrained by karyotype. The appropriateness of assumed linearity between genotype and phenotype can be assessed by monitoring the genomic variability and stability of the cell population. The complexity of many cancer cell populations allows for similar phenotypes to be associated with a number of alternative genome-systems or sequences of cellular events. This aspect of complexity dictates that the causative events in phenotypic cellular evolution can be probabilistically speculated upon, but not unequivocally determined. Within this framework, the proposed paradigm implicates the interplay between deterministic chaos, complexity, self-organization, and natural selection as driving forces for conversion to the malignant phenotype.

The rationale for shifting from a holistic viewpoint rather than a reductionist approach is largely provided by the unique features of cancer cell populations that cannot be explained by our current understanding of the governance of eukaryotic cells with normal genomes [230]. Widely held viewpoints based in Darwinian evolutionary theory have established a teleological dogma of cause-based-on-function and emphasized a role of natural selection for the fittest individuals during a continuous evolutionary process. Within the framework of Darwinian microevolution, the concepts of descent with modification, and the advantages provided by slight and successive variations, it is not possible to understand the rapidity of genotypic or phenotypic adaptability of cancer cell populations. The holistic viewpoint proposes that cancer cell populations characterized by widespread genomic instability are precisely the irreducibly complex system that Darwin himself determined exceptions to his theory [231]. The widespread genomic heterogeneity demonstrated in the current data set also suggests a great deal or redundancy within the genomes of cancer cells as numerous viable karyotypes are seen. This finding supports an alternative under the proposed paradigm shift and de-emphasizes any specific micro-evolutionary change in the determination of causality and its linkage to specific features of genes or genomes. This traditional objective of Mendelian genetics, to relate cause and effect in a predictable manner, cannot be achieved in the case of a complex-system because the defining features of complex-systems are those of chaos and complexity [232-233], rather than constancy and linearity. The interconnectedness (i.e., existence of multiple links between elements) of the system, dictates that neither the cause nor the driving force for evolution of the entire hierarchical system can be identified unequivocally. The holistic viewpoint proposed by the genome-centered paradigm places cancer causation and progression at the intersection of chaos, complexity, self-organization, and selection by genome-system. In understanding that entirely different conceptual frameworks pertain to normal versus hyper-variable genome systems, the rationale for measuring the genomic

variability and stability of cell populations becomes clear, is as it provides an index for weighting the significance of gene-level change.

Macroevolution versus Microevolution

The genome centric paradigm clearly delineates normal micro-evolutionary change of normal cells from macro-evolutionary mechanisms of cancer cells on the basis of their respective evolutionary genomic constraints. That is, normal diploid cells evolve by micro-evolutionary change and transformed cells are capable of macro-evolutionary change by altering their entire karyotype in a single cell division. By incorporating these differential mechanisms of evolutionary change under the genome-centered paradigm, it becomes clear how the predictability with which specific genes increase the transformation rate of diploid cells in culture could be far different from the predictability of such gene alterations in solid tumors. Even if the oncogenes purported to be the root cause of cellular transformation are discovered by large-scale tumor sample analyses, they will have limited potential as therapeutic targets. This strategy involves specific targeting of a gene or pathway in a complex-adaptive system, which is problematic because cancer cells are not constrained by usual growth and division checkpoints or by the functional roles they once held in their tissue or organ of origin [164]. In this setting, targeted therapy will likely selectively kill only a fraction of the cells. It has been shown by karyotypic and phenotypic change that cancer cells are capable of undergoing rapid macro-evolutionary phenotypic shifts [164] and that the potential for drug resistance by genome level cell-evolutionary mechanisms is large [234].

Thus, gene and genome centric paradigms can be reconciled by understanding that genetic cancer causation can be described as micro-evolutionary to macro-evolutionary transition in cellular transformation and tumorigenesis. The genes that cancer researchers believe are important for cancer treatment are those that function in cancer causation to bridge the gap between micro-evolutionary change and Macro-evolutionary change. Whether by non-heritable

transient changes in protein abundance or activity, heritable gene-level molecular change such as sequence mutation or heritable epigenetic alteration, or large scale physical rearrangement of the chromosome complement by rearrangement of whole or partial chromosomes, the specific initiating events by which cells transform to become tumorigenic are not as significant as the consequence of instability itself. Under the genome-centered paradigm, this shift in focus is particularly important, as genome level change is thought to account for macro-evolutionary adaptation of cancer cell populations. The genome-centered paradigm considers cancer as a disease of odds with an almost infinite number of possible combinations of causative factors, whose behavior is best predicted by the acquisition of macro-evolutionary change and subsequently by environmental selection. Emphasis is placed on the presence of macro-evolutionary genomic change, rather than on the initial destabilizing events, because the switch to macro-evolutionary change makes the gene alteration unpredictable, and more importantly, provides the variability on which selection may occur.

Once widespread genomic variability is evident, genome-centered logic leads to consideration of cancer cell populations as complex-adaptive systems with focus on *intra-tumor* rather than *inter-tumor* variability. Intra-tumor variability or cell-to-cell variability within the same tumor exists at multiple levels. These levels include cell phenotype and ploidy level [235], chromosomal aberration [136-137, 236-244], DNA copy number [245], single nucleotide polymorphisms [246], microsatellite shifts [247-248], base mutations [219-220], and loss of heterozygosity [249]. The current data demonstrate how the largest changes in gene expression coincide with the largest shift in karyotype. In transformed late-stage cells, change in whole chromosome copy numbers appears to regulate gene expression. This signifies a major shift in the regulatory capacity of cells to govern their genome and is marked by altered and heterogeneous karyotypes among cell populations which generate large-scale variability. Change by macro-evolution redefines the complex-adaptive system of the cancer cell

population by changing chromosome copy numbers and the relative orientation of genes on each chromosome.

Measuring Signal versus Noise in Cancer Cells

Despite the variability known to exist in cancer cell populations, the relative importance of pattern identification over characterization of the true biological variability of cancer cell populations should be given careful consideration. This preference is made clear by the majority of investigative techniques commonly utilized to test hypotheses under the gene-centered paradigm. For example, the widespread acceptance of techniques that average and batch genetic material from cell populations with high levels of genomic heterogeneity [250-251] is one way that the average genomic signal from a mixed population of cells is emphasized over the quantification of genetic diversity. These techniques include a wide variety of extremely common techniques such as harvesting of extracted nucleic acids, proteins, and other cellular components from lysed populations of genomically heterogeneous cells. Secondly, the manifestation of gene-centered logic in experimental methodologies is the significance placed on increasing the resolution of molecular techniques while failing to maintain resolution at the level of the individual cell. Population averaged data that are reported without measure of the deviation from cell-to-cell certainly has earned its rightful place in molecular problem solving, but only when taken in the context of the present cell diversity for that population and the stability of the population in its environment. For understanding the inter and intra-tumor genomic heterogeneity of cancer, it becomes clear that the pattern of increasing molecular resolution and increased scope and size of project are likely due to the difficulty in generating signal over the background noise of true biologic variability. Background biological variability particularly presents a formidable challenge for noise reduction in epithelial cancers. For example, in the case of epithelial ovarian cancer heterogeneity has been reported from the level of histological subtype to DNA sequence [32]. Rather than focus on analytical techniques that emphasize

genomic signal over the true “noise” of cancer cell systems, the current data set and work by our group [128-130, 193] and others [115, 252-254] have focused the experimental design and techniques to quantify genomic heterogeneity as it relates to macroevolution and cell phenotype.

Several large-scale projects such as The Cancer Genome Atlas (TCGA) and the Cancer Biomedical Informatics Grid (CABIG) are underway in cataloging the sequences of hundreds of patient tumor samples. The rationale for funding such large-scale projects is that the identification of gene mutations common among clinical tumor specimens will yield biologically-relevant therapeutic targets for clinical cancer treatment [255-257]. In such studies, a wide variety of low frequency genetic alterations have been discovered. For example, a team of forty-two scientists searched for common mutations among 18,000 genes and 11 tumor samples. Most of these genes were found to be mutated in less than 10% of samples. Of the cancer genes selected for the validation set, 15/40 were not found to be mutated in any of the 96 patient cancers. In a similar vein, gene-centered rationale has also provided the impetus for recent sequencing of entire cancer genomes from batched extracted DNA. This and other similar studies have proven to be extremely complicated. In sequencing the melanoma genome, the authors describe the difficulty in mapping the genetic components correctly onto each chromosome [256-257]. Additionally, repeated tweaking of the bioinformatics algorithms and the analysis of 200 billion data points was required to determine the 33,000 mutations characterizing the melanoma cell line [256-257]. Genome-centered logic suggests that the *intra-tumor* (cell-to-cell) heterogeneity of this line has greater significance than the mutations with significant commonality among cells to generate signal above the biological noise. Heterogeneity of the cell population at this level was readily apparent among 10 metaphase cells: 6/10 metaphases have 3 Chr. 11, four of these with an additional unbalanced der(?)t(1p?;18q?), while 4 other metaphases have 2 Chr. 11, and one der(?)t(11;18). Based on this karyotype data, it is clear that several different versions of the melanoma genome system

co-exist within one cell line. This variability brings into question how the results of such sequencing experiments should be interpreted and the composition possibly of sequence of the dominant signal over the noise of the mixed genome system that was analyzed. The proposed paradigm shift additionally questions the significance of these data based on the transient nature of any cancer genome sequence given the rate of mutation (33,000/ division) at the level of the DNA base. Finally, the likelihood that targeted therapies for clinical cancers aimed at gene products of mutated sequence from this melanoma line will result in cancer cure is extremely small considering the combined heterogeneity and adaptability of the genome system.

Basic scientists are beginning to measure behavior of cell populations in lower organisms on a global scale and on multiple levels (genome, transcriptome, metabolome) at the same time. These strategies are advantageous because they provide information as to the regulatory control among multiple levels of cellular dynamics. Using these techniques, a complex-systems based approach to data interpretation can be employed in the evaluation of cancer cell populations. However, results generated from such studies must be considered in light of data that shows extreme cellular variability additionally occurs at the level of transcription and translation when such outcomes are measured at the level of the individual cell [258-259]. For example, when comparing metaphase-CGH to SNP-CGH to array-CGH, molecular resolution is clearly increasing while information quantifying cellular variability is lost. The strength of data recorded on the content or behavior of individual cells within a population such as SKY karyotype analysis, *in situ* hybridization, metaphase-CGH, flow cytometry, immunohistochemistry and others is their ability to measure outcome variable(s) while maintaining cellular resolution, thereby allowing for quantification of variability about signal strength, and position relative to other structures. One such technology with the capacity to characterize heterogeneous cellular responses is histomics. This technique involves the collection of proteins from a population of cells, the generation of monoclonal antibodies from

the batched cellular responses, and the application of these antibodies to the cell population/tissue from which they were collected to determine the origin, sub-cellular location, and cellular variability of the response *in situ*. When considering the effort required to explore the cancer genome that involves gene-seeking or sequencing of heterogeneous populations, these high-throughput technologies with high molecular resolution seem less attractive.

As high-throughput techniques continue to be developed to support cell-by-cell experimental methods, data from batched cellular material can be interpreted within the proposed genome-centered paradigm. These data will provide information as to the variability of the cellular response and the predictability and permanence of the response in different environments. This is particularly important in cancer research, aging, or in fields where internal or external exposure to genotoxic agents jeopardizes the integrity of the genome-system. In these cases, the fidelity of genome transfer is likely low, and the system can transition to macro-evolutionary genomic shift, rather than micro-evolutionary Darwinian evolution. For example, the current data set displays tumors with a set of similar marker chromosomes and two different patterns of cytogenetic diversity. The widespread genomic instability of the injected cell line suggests that this experiment, if repeated, would likely emerge to showcase a different pattern of high-frequency clonal translocations. If cell survival is challenged, low-fidelity genome transfer suggests that the genome context under which that data set was collected is transient because variability and macroevolution can occur within that population of cells. Particularly in the case of unstable genomes, averaged data without the reporting of intra-population variability provides no insight into the long term impact of the discovery.

Clinical Considerations under the Genome-centered Paradigm

Although concepts such as paradigmatic shifts and altererd conceptualizaiton of gneomic variability in cancer cell biology may seem of little clinical value compared to the promise of successful trageter therapies, these theoretical shifts have real and practical clinical impact in a

number of areas. While acknowledging the complexity of cancer cell populations and their propensity for differential gene regulation and signal transduction, the majority of investigators utilize laboratory cancer models and data interpretation within the framework of the gene-centered paradigm. The focus on pattern matching and noise reduction continues to generate short and fluid lists of targets for anti-cancer therapies and are considered suffieicnt rationale for the continued search for the genetic "Achilles heel" of cancer cells {Weinstein, 2006 #5446}. However, the empirical data generated from clinical trials of targeted therapies supports their use in combination with at lest one or two other chemotherapeutic agents [260]. Even when multiple targeted therapies are used in combination, these do not provide a cure. From a clinical standpoint, the diminutive role placed on cancer cell heterogeneity by the gene-centered paradigm seems somewhat unique to therapeutic design for solid tumors. Cancer cells have a 200 fold increase in mutation rate compared to normal cells {Bielas, 2006 #2050;Bielas, 2005 #2051;Loeb, 2008 #2049}. Additionally, the rate of chromosome mis-segregation events (one out of five divisions [121]). Each of these missegregation events offsets the DNA doseage by 61-197 mega-base pairs, therefore further contributing to the variability of the genome within a single population. Considering that clinically detectable cancers have already reached 10×10^9 cells, a modest 10% increase yields 6.3×10^{12} new DNA base mutations coupled with 7×10^7 abnormal chromosome segregation events. By these measures, the genome centric viewpoint reasons that low-fidelity genome transfer and the ensuing multi-level genomic heterogeneity should be recognized as the major mechanism of drug resistance in solid tumors. Disease with high-level variability should then beaddressed with first-line therapy as in the approach for treating many genomically heterogeneous and rapidly mutating diseases caused by prokaryotes and viri. The search for targeted therapy in the face of such genomic diversity hinges upon the rare success of certain targeted therapeutics in clinical cancer treatment. However, reconsideration of genome dynamics linked to responsive versus resistant cases actually highlights the importance of genomic variability over specific genes. For example, the high

remission rates and low toxicity of the targeted BCR-ABL tyrosine kinase inhibitor has extended 5-year survival rates to ~90% in chronic myelogenous leukemia (CML) patients and has ushered in an era of targeted therapy. CML however is unlike most cancers in that its cause is believed to be a single genetic event that is visible by a high frequency 9;22 chromosomal translocation. This translocation creates a fusion protein with abnormally high tyrosine kinase activity [132] and, like other members of the hematologic malignancies, the frequency of random aberrations and the karyotypic variability of CML cells is relatively low when compared to solid tumors [79]. Like other cancer genes, *BCR-ABL* has been linked to genomic instability, mitotic errors, and unfaithful DNA repair, perhaps aiding in the transition from chronic to blast-crisis phase [261-263]. These changes have been noted at diagnosis. Therefore, it is not surprising that progression for patients on chronic Imatinib therapy occurs at a rate of ~15% over 54 months. Leukemic blasts from resistant individuals were first shown by sequence exploration to harbor disabling ligand association mutations [175, 264]. Subsequently, resistance mechanisms by chromosomal aberrations and copy number alterations of BCR-ABL tyrosine kinase [175, 265-267] were also revealed. But in patients with blast-crisis, where genomic instability is significantly elevated, remission was unpredictable and short lived. Patients typically underwent dramatic relapse just weeks after a complete cytogenetic response [268-273]. The relative success of targeted therapy in chronic versus blast phase CML demonstrate that gene-centered approaches have much greater utility in the context of low genomic instability. Specifically, targeted therapy was successful when the phenotype was homogeneous, the karyotype relatively stable, and the gene target most clearly mapped. In epithelial tumors, the complexities of the predominant karyotype and high degree of clonal heterogeneity in the form of cytogenetically unrelated clones makes characterization by established means virtually impossible. Under the gene-centered paradigm, genomic heterogeneity of solid tumors represents a further dimension of complexity in analytical problem

solving that is greatly attenuated in clinical cancer conditions still evolvolving by micro-evolutionary change.

The environment of homogeneity and relative genomic stability seen in successful targeted therapeutic intervention is not characteristic of most clinically relevant cancers [80, 213]. Therefore, continued application of gene-centered approaches to problem solving within a complex adaptive system reflects an over-extension of traditional genetics to a hyper-variable and rapidly evolving non-linear system. Genome-centered reasoning suggests that continued investment in strategies which discount the existence of cell-to-cell variability and rapid phenotypic shift will likely not yield results leading to clinically successful treatment. The distinction between micro versus macro-evolutionary mechanisms in normal versus cancer cells provides rationale for the relative predictability of the genotype: phenotype relationship in each type of evolutionary mechanism. For example, in heritable gene mutations in the context of a normal diploid background there is a fair amount of predictive power between gene alteration and increased cancer risk. Similarly, the altered *BCR-ABL* gene has greater predictive power in chronic phase CML. This greater predictive power occurs when genomic instability is relatively low, and the target is clear and frequent among the cell population [79-80]. The differences between micro and macro-evolutionary mechanisms additionally provide contextual insight into why multiple studies have failed to determine significant lists of validated genes from heterogeneous clinical tumor samples [250, 274].

Clinically, the challenges posed by the application of targeted gene-therapies to a complex system with multi-level heritable and non-heritable variability have been evidenced by the multi-modal acquisition of Imatinib resistance in blast phase cells. The variability of resistance mechanisms and mutability of blast-phase cells are similar to the varibility and drug resistance mechanisms previously documented for solid tumors. These mechanisms include gene amplification of associtated proteins or pathways, such as the dihydrofolate reductactase and androgen receptors [275-277], the acquisition of entire chromosomes [278-280] and the shifting

of entire karyotypes [229]. An example of multi-modal resistance in a single cancer type is that of estrogen-receptor positive (ERα+) breast cancers. In this subtype of breast cancer alone, known resistance mechanisms include, deletions involving estrogen receptor-α [281-282], altered gene expression [283], interference or cross-talk between growth factor signaling pathways EGFR/HER2 [284-285], protein-protein interactions among SERM, ER, and co-regulatory proteins [286], and the background genomic profile of the cancer patient such as metabiolic polymorphisms including *CYP2D6*. Considering the multi-modal mechanisms of genomic diversity documented in solid tumors and late stage hematologic disease, the gene-centered approach to gene-discovery for targeted therapy by analysis of late stage solid tumor samples seems less likely to yield long-term remission or cure. Variability within a single tumor hinders targeted therapeutic approaches in two ways: 1) the biological noise truly existent in the sample makes pattern finding difficult, even gene alteration existing in the majority of cells can be difficult to find, 2) if these signals are detectable, the targeted therapeutic approach will select for cells able to survive, appearing as a drug resistant population that has the capacity for high mutation rates.

Under the guise of that clonal tumor populations dominate most cancers, investigators look toward oncogenes and tumor suppressor genes identified from solid tumor patient samples to provide the basis for designing targeted anti-cancer therapies. However, within the framework provided by the genome-centered paradigm, the significance of measuring, controlling for, or reducing genomic instability must be incorporated in the development first line or adjuvant therapy of a logical approach to prevent further diversification of cancer cell populations is to be implemented. Not surprisingly, heterogeneity exists even in the response of cancer cells to genome-stabilizing treatments [287]. Therefore, multi-pathway or multi-level approaches may be required in conjunction with genome stabilizing therapy, perhaps also reducing drug dosages to a safer therapeutic window [288].

Several molecular approaches thus far have been proposed and are in various stages of *in vivo* and *in vitro* testing. An example of therapy that targets genomically unstable cells is the inhibition of centrosome clustering required for multipolar mitosis [289]. Additionally, the use of therapies such as $1,25(OH)_2D3$ that work at multiple regulatory levels [290-293] and potentiate the effects of many cytotoxic and antiproliferative therapies [290, 294] are promising therapeutic strategies according to the genome-centered paradigm. In a similar vein, randomized aptamer libraries have been suggested to reduce tumor cell heterogeneity by allowing cancer cells select for RNA sequences that are then conjugated to radionuclide and cytotoxic drugs and targeted to many tumor cells [295]. Therapies such as these are a particularly good match for the complex-systems adaptive view of cancer biology because they impact multiple regulatory levels and are therefore more likely to impact heterogeneous cell populations.

Implications of Incorporating the Genome-Centered Paradigm

In conclusion, the tremendous research efforts put forth in the past decades have demonstrated cancer to be extremely complicated. Cancer cell populations universally demonstrate karyotype alterations and genomic instability, signifying their ability to adapt by macroevolution. The implications of the paradigm shift to genome-centered thinking on cancer research and treatment are many as they alter the significance of a large amount of past and present clinical and basic research. Based on our data and the work of others demonstrating the extent and significance of cell-to-cell variability in cancer biology, it is anticipated that the measurement of outcome variables including intra-population variation will provide the context for data interpretation and inferences as to the repeatability of the experiments. Investigators are encouraged to consider the possibility of rapid macro-evolutionary change in cancer cell lines and tumors, particularly when the cell population incurs cell death, selection, or environmental stress. On beginning to explore how genomically unstable cell populations develop the emergent properties of complex systems, we are optimistic that investigators will

increase their utilization of genome-wide measurements of variables at multiple levels ranging from the gene sequence upwards to the cellular or tumor phenotype.

As technological advances permit the comparison of genomes, and interactome maps encompassing many pathways, high-throughput techniques may contribute to the systematic discounting of true biological noise. Particularly in genomically unstable systems, the possibility of rapid macro-evolutionary phenotypic shifts should be accounted for and evaluated by measuring the karyotypic make-up of the cell population in each condition. By explaining the complexity of cancer genomics and highlighting intra-tumor variability over inter-tumor variability, evidence is provided that supports a paradigm shift away from the current linear gene-centered approach. The genome-centered paradigm does not dispel the significance of gene-level change, but considers the genomic context and relative predictability of such changes. The genome-centered paradigm additionally allows for non-linearity and measurement of change considering cancer cell populations to be complex-adaptive systems. In this way the proposed paradigm incorporates the possibility of multi-level regulation and macro-evolutionary change. The genome-centered paradigm shift and the supporting data supply an improved conceptual framework for a field currently filled with exceptions and paradoxes. Moreover, the paradigm shift will alter experimental design and interpretations to support the advancement of cancer therapies and diagnostics, thereby improving future outcomes for cancer patients.

REFERENCES

1. National, I.C. *A snapshot of ovarian cancer*. Available from: http://planning.cancer.gov/disease/Ovarian-Snapshot.pdf.

2. Bonome, T., et al., *Expression profiling of serous low malignant potential, low-grade, and high-grade tumors of the ovary.* Cancer Res, 2005. **65**(22): p. 10602-12.

3. Bast, R.C., Jr., B. Hennessy, and G.B. Mills, *The biology of ovarian cancer: new opportunities for translation.* Nat Rev Cancer, 2009. **9**(6): p. 415-28.

4. Despierre, E., et al., *The molecular genetic basis of ovarian cancer and its roadmap towards a better treatment.* Gynecol Oncol, 2010. **117**(2): p. 358-65.

5. Prasad, M., et al., *High definition cytogenetics and oligonucleotide aCGH analyses of cisplatin-resistant ovarian cancer cells.* Genes Chromosomes Cancer, 2008. **47**(5): p. 427-36.

6. Shah, S.N., S.E. Hile, and K.A. Eckert, *Defective mismatch repair, microsatellite mutation bias, and variability in clinical cancer phenotypes.* Cancer Res, 2010. **70**(2): p. 431-5.

7. Barnholtz-Sloan, J.S., et al., *Ovarian cancer: changes in patterns at diagnosis and relative survival over the last three decades.* Am J Obstet Gynecol, 2003. **189**(4): p. 1120-7.

8. Arts, H.J., et al., *Options for modulation of drug resistance in ovarian cancer.* Int J Gynecol Cancer, 2000. **10**(S1): p. 47-52.

9. Eltabbakh, G.H. and C.S. Awtrey, *Current treatment for ovarian cancer.* Expert Opin Pharmacother, 2001. **2**(1): p. 109-24.

10. Anderson, E., et al., *Cytological observations of the ovarian epithelium in mammals during the reproductive cycle.* J Morphol, 1976. **150**(1): p. 135-65.

11. Auersperg, N., et al., *The biology of ovarian cancer.* Semin Oncol, 1998. **25**(3): p. 281-304.

12. Ozols, R.F., *Treatment goals in ovarian cancer.* Int J Gynecol Cancer, 2005. **15 Suppl 1**: p. 3-11.

13. Auersperg, N., T. Ota, and G.W. Mitchell, *Early events in ovarian epithelial carcinogenesis: progress and problems in experimental approaches.* Int J Gynecol Cancer, 2002. **12**(6): p. 691-703.

14. Sundfeldt, K., *Cell-cell adhesion in the normal ovary and ovarian tumors of epithelial origin; an exception to the rule.* Mol Cell Endocrinol, 2003. **202**(1-2): p. 89-96.

15. Feeley, K.M. and M. Wells, *Precursor lesions of ovarian epithelial malignancy.* Histopathology, 2001. **38**(2): p. 87-95.

16. Auersperg, N., et al., *Ovarian surface epithelium: biology, endocrinology, and pathology.* Endocr Rev, 2001. **22**(2): p. 255-88.

17. Morgan, R., et al., *Targeting HOX and PBX transcription factors in ovarian cancer.* BMC Cancer, 2010. **10**: p. 89.

18. Bahrani-Mostafavi, Z., et al., *Correlation analysis of HOX, ErbB and IGFBP family gene expression in ovarian cancer.* Cancer Invest, 2008. **26**(10): p. 990-8.

19. Yamashita, T., et al., *Suppression of invasive characteristics by antisense introduction of overexpressed HOX genes in ovarian cancer cells.* Int J Oncol, 2006. **28**(4): p. 931-8.

20. Fathalla, M.F., *Incessant ovulation--a factor in ovarian neoplasia?* Lancet, 1971. **2**(7716): p. 163.

21. Fleming, J.S., et al., *Incessant ovulation, inflammation and epithelial ovarian carcinogenesis: revisiting old hypotheses.* Mol Cell Endocrinol, 2006. **247**(1-2): p. 4-21.

22. Banks, E., ed. *The epidemiology of ovarian cancer*. In Ovarian Cancer Methods and Protocols, ed. J.M.S. Bartlett. 2000, Humana Press: Totowa. 3-11.

23. Choi, J.H., et al., *Gonadotropins upregulate the epidermal growth factor receptor through activation of mitogen-activated protein kinases and phosphatidyl-inositol-3-kinase in human ovarian surface epithelial cells.* Endocr Relat Cancer, 2005. **12**(2): p. 407-21.

24. Choi, J.H., et al., *Overexpression of follicle-stimulating hormone receptor activates oncogenic pathways in preneoplastic ovarian surface epithelial cells.* J Clin Endocrinol Metab, 2004. **89**(11): p. 5508-16.

25. Choi, K.C., et al., *Follicle-stimulating hormone activates mitogen-activated protein kinase in preneoplastic and neoplastic ovarian surface epithelial cells.* J Clin Endocrinol Metab, 2002. **87**(5): p. 2245-53.

26. Cramer, D.W. and W.R. Welch, *Determinants of ovarian cancer risk. II. Inferences regarding pathogenesis.* J Natl Cancer Inst, 1983. **71**(4): p. 717-21.

27. Cramer, D.W., et al., *Determinants of ovarian cancer risk. I. Reproductive experiences and family history.* J Natl Cancer Inst, 1983. **71**(4): p. 711-6.

28. Murdoch, W.J., R.S. Townsend, and A.C. McDonnel, *Ovulation-induced DNA damage in ovarian surface epithelial cells of ewes: prospective regulatory mechanisms of repair/survival and apoptosis.* Biol Reprod, 2001. **65**(5): p. 1417-24.

29. Godwin, A.K., J.R. Testa, and T.C. Hamilton, *The biology of ovarian cancer development.* Cancer, 1993. **71**(2 Suppl): p. 530-6.

30. Plaxe, S.C., et al., *Ovarian intraepithelial neoplasia demonstrated in patients with stage I ovarian carcinoma.* Gynecol Oncol, 1990. **38**(3): p. 367-72.

31. Resta, L., et al., *Morphologic precursors of ovarian epithelial tumors.* Obstet Gynecol, 1993. **82**(2): p. 181-6.

32. Wang, V., et al., *Ovarian cancer is a heterogeneous disease.* Cancer Genet Cytogenet, 2005. **161**(2): p. 170-3.

33. Melnick, R.L., et al., *Inhalation toxicology and carcinogenicity of 1,3-butadiene in B6C3F1 mice following 65 weeks of exposure.* Environ Health Perspect, 1990. **86**: p. 27-36.

34. Jacobs, A.J., et al., *Chemical induction of ovarian epithelial carcinoma in mice.* Gynecol Oncol, 1984. **18**(2): p. 177-80.

35. Orsulic, S., et al., *Induction of ovarian cancer by defined multiple genetic changes in a mouse model system.* Cancer Cell, 2002. **1**(1): p. 53-62.

36. Garson, K., et al., *Generation of tumors in transgenic mice expressing the SV40 T antigen under the control of ovarian-specific promoter 1.* J Soc Gynecol Investig, 2003. **10**(4): p. 244-50.

37. Liang, S., et al., *Expression of activated PIK3CA in ovarian surface epithelium results in hyperplasia but not tumor formation.* PLoS One, 2009. **4**(1): p. e4295.

38. Clark-Knowles, K.V., et al., *Conditional inactivation of Brca1, p53 and Rb in mouse ovaries results in the development of leiomyosarcomas.* PLoS One, 2009. **4**(12): p. e8534.

39. Dinulescu, D.M., et al., *Role of K-ras and Pten in the development of mouse models of endometriosis and endometrioid ovarian cancer.* Nat Med, 2005. **11**(1): p. 63-70.

40. Wu, R., et al., *Mouse model of human ovarian endometrioid adenocarcinoma based on somatic defects in the Wnt/beta-catenin and PI3K/Pten signaling pathways.* Cancer Cell, 2007. **11**(4): p. 321-33.

41. Pal, T., et al., *BRCA1 and BRCA2 mutations account for a large proportion of ovarian carcinoma cases.* Cancer, 2005. **104**(12): p. 2807-16.

42. Deng, C.X. and R.H. Wang, *Roles of BRCA1 in DNA damage repair: a link between development and cancer.* Hum Mol Genet, 2003. **12 Spec No 1**: p. R113-23.

43. MacLachlan, T.K., R. Takimoto, and W.S. El-Deiry, *BRCA1 directs a selective p53-dependent transcriptional response towards growth arrest and DNA repair targets.* Mol Cell Biol, 2002. **22**(12): p. 4280-92.

44. Welcsh, P.L., et al., *BRCA1 transcriptionally regulates genes involved in breast tumorigenesis.* Proc Natl Acad Sci U S A, 2002. **99**(11): p. 7560-5.

45. Bochar, D.A., et al., *BRCA1 is associated with a human SWI/SNF-related complex: linking chromatin remodeling to breast cancer.* Cell, 2000. **102**(2): p. 257-65.

46. Baer, R. and T. Ludwig, *The BRCA1/BARD1 heterodimer, a tumor suppressor complex with ubiquitin E3 ligase activity.* Curr Opin Genet Dev, 2002. **12**(1): p. 86-91.

47. Miki, Y., et al., *A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1.* Science, 1994. **266**(5182): p. 66-71.

48. Clark-Knowles, K.V., et al., *Conditional inactivation of Brca1 in the mouse ovarian surface epithelium results in an increase in preneoplastic changes.* Exp Cell Res, 2007. **313**(1): p. 133-45.

49. Tillmann, T., K. Kamino, and U. Mohr, *Incidence and spectrum of spontaneous neoplasms in male and female CBA/J mice.* Exp Toxicol Pathol, 2000. **52**(3): p. 221-5.

50. Walsh, K.M. and J. Poteracki, *Spontaneous neoplasms in control Wistar rats.* Fundam Appl Toxicol, 1994. **22**(1): p. 65-72.

51. Gregson, R.L., D.J. Lewis, and D.P. Abbott, *Spontaneous ovarian neoplasms of the laboratory rat.* Vet Pathol, 1984. **21**(3): p. 292-9.

52. Godwin, A.K., et al., *Spontaneous transformation of rat ovarian surface epithelial cells: association with cytogenetic changes and implications of repeated ovulation in the etiology of ovarian cancer.* J Natl Cancer Inst, 1992. **84**(8): p. 592-601.

53. Testa, J.R., et al., *Spontaneous transformation of rat ovarian surface epithelial cells results in well to poorly differentiated tumors with a parallel range of cytogenetic complexity.* Cancer Res, 1994. **54**(10): p. 2778-84.

54. Roby, K.F., et al., *Development of a syngeneic mouse model for events related to ovarian cancer.* Carcinogenesis, 2000. **21**(4): p. 585-91.

55. Roberts, P.C., et al., *Sequential molecular and cellular events during neoplastic progression: a mouse syngeneic ovarian cancer model.* Neoplasia, 2005. **7**(10): p. 944-56.

56. Greenaway, J., et al., *Epithelial-stromal interaction increases cell proliferation, survival and tumorigenicity in a mouse model of human epithelial ovarian cancer.* Gynecol Oncol, 2008. **108**(2): p. 385-94.

57. Urzua, U., et al., *Microarray comparative genomic hybridization profile of a murine model for epithelial ovarian cancer reveals genomic imbalances resembling human ovarian carcinomas.* Tumour Biol, 2005. **26**(5): p. 236-44.

58. Janat-Amsbury, M.M., et al., *Combination of local, non-viral IL12 gene therapy and systemic paclitaxel chemotherapy in a syngeneic ID8 mouse model for human ovarian cancer.* Anticancer Res, 2006. **26**(5A): p. 3223-8.

59. Janat-Amsbury, M.M., et al., *Comparison of ID8 MOSE and VEGF-modified ID8 cell lines in an immunocompetent animal model for human ovarian cancer.* Anticancer Res, 2006. **26**(4B): p. 2785-9.

60. Schwartz, D.R., et al., *Gene expression in ovarian cancer reflects both morphology and biological behavior, distinguishing clear cell from other poor-prognosis ovarian carcinomas.* Cancer Res, 2002. **62**(16): p. 4722-9.

61. Urzua, U., et al., *Transcriptomic analysis of an in vitro murine model of ovarian carcinoma: functional similarity to the human disease and identification of prospective tumoral markers and targets.* J Cell Physiol, 2006. **206**(3): p. 594-602.

62. Kiechle, M., et al., *Comparative genomic hybridization detects genetic imbalances in primary ovarian carcinomas as correlated with grade of differentiation.* Cancer, 2001. **91**(3): p. 534-40.

63. Gregoire, L., et al., *Spontaneous malignant transformation of human ovarian surface epithelial cells in vitro.* Clin Cancer Res, 2001. **7**(12): p. 4280-7.

64. Tsao, J., et al., *Telomerase activity in normal and neoplastic breast.* Clin Cancer Res, 1997. **3**(4): p. 627-31.

65. Jin, Y., et al., *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, 2004. **92**(1): p. 183-91.

66. Bayani, J., et al., *Distinct patterns of structural and numerical chromosomal instability characterize sporadic ovarian cancer.* Neoplasia, 2008. **10**(10): p. 1057-65.

67. Bernardini, M., et al., *High-resolution mapping of genomic imbalance and identification of gene expression profiles associated with differential chemotherapy response in serous epithelial ovarian cancer.* Neoplasia, 2005. **7**(6): p. 603-13.

68. Shridhar, V., et al., *Genetic analysis of early- versus late-stage ovarian tumors.* Cancer Res, 2001. **61**(15): p. 5895-904.

69. Taetle, R., et al., *Chromosome abnormalities in ovarian adenocarcinoma: II. Prognostic impact of nonrandom chromosome abnormalities in 244 cases.* Genes Chromosomes Cancer, 1999. **25**(1): p. 46-52.

70. Taetle, R., et al., *Chromosome abnormalities in ovarian adenocarcinoma: I. Nonrandom chromosome abnormalities from 244 cases.* Genes Chromosomes Cancer, 1999. **25**(3): p. 290-300.

71. Lengauer, C., K.W. Kinzler, and B. Vogelstein, *Genetic instability in colorectal cancers.* Nature, 1997. **386**(6625): p. 623-7.

72. Rajagopalan, H., et al., *The significance of unstable chromosomes in colorectal cancer.* Nat Rev Cancer, 2003. **3**(9): p. 695-701.

73. Cahill, D.P., et al., *Mutations of mitotic checkpoint genes in human cancers.* Nature, 1998. **392**(6673): p. 300-3.

74. Kops, G.J., B.A. Weaver, and D.W. Cleveland, *On the road to cancer: aneuploidy and the mitotic checkpoint.* Nat Rev Cancer, 2005. **5**(10): p. 773-85.

75. Nigg, E.A., *Centrosome aberrations: cause or consequence of cancer progression?* Nat Rev Cancer, 2002. **2**(11): p. 815-25.

76. Muleris, M., et al., *Chromosomal instability in near-diploid colorectal cancer: a link between numbers and structure.* PLoS ONE, 2008. **3**(2): p. e1632.

77. Bayani, J., et al., *Genomic mechanisms and measurement of structural and numerical instability in cancer cells.* Semin Cancer Biol, 2007. **17**(1): p. 5-18.

78. Stewart, S.A. and R.A. Weinberg, *Telomerase and human tumorigenesis.* Semin Cancer Biol, 2000. **10**(6): p. 399-406.

79. Roschke, A.V., et al., *Karyotypic complexity of the NCI-60 drug-screening panel.* Cancer Res, 2003. **63**(24): p. 8634-47.

80. Castro, M.A., et al., *Profiling cytogenetic diversity with entropy-based karyotypic analysis.* J Theor Biol, 2005. **234**(4): p. 487-95.

81. Pollack, J.R., et al., *Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors.* Proc Natl Acad Sci U S A, 2002. **99**(20): p. 12963-8.

82. Upender, M.B., et al., *Chromosome transfer induced aneuploidy results in complex dysregulation of the cellular transcriptome in immortalized and cancer cells.* Cancer Res, 2004. **64**(19): p. 6941-9.

83. Duesberg, P., *Chromosomal chaos and cancer.* Sci Am, 2007. **296**(5): p. 52-9.

84. FitzPatrick, D.R., et al., *Transcriptome analysis of human autosomal trisomy.* Hum Mol Genet, 2002. **11**(26): p. 3249-56.

85. Altug-Teber, O., et al., *Specific transcriptional changes in human fetuses with autosomal trisomies.* Cytogenet Genome Res, 2007. **119**(3-4): p. 171-84.

86. Heng, H.H., et al., *Genetic and epigenetic heterogeneity in cancer: A genome-centric perspective.* J Cell Physiol, 2009.

87. Ye, C.J., et al., *Genome based cell population heterogeneity promotes tumorigenicity: the evolutionary mechanism of cancer.* J Cell Physiol, 2009. **219**(2): p. 288-300.

88. Heng, H.H., L.C. Tsui, and P.B. Moens, *Organization of heterologous DNA inserts on the mouse meiotic chromosome core.* Chromosoma, 1994. **103**(6): p. 401-7.

89. Ye, C.J., et al., *The combination of SKY and specific loci detection with FISH or immunostaining.* Cytogenet Cell Genet, 2001. **93**(3-4): p. 195-202.

90. Liu, G., et al., *Spectral karyotyping of mouse cell line WMP2.* Cytogenet Cell Genet, 2000. **90**(3-4): p. 271-4.

91. Shannon, C.E., *The mathematical theory of communication. 1963.* MD Comput, 1997. **14**(4): p. 306-17.

92. Leek, J.T., et al., *EDGE: extraction and analysis of differential gene expression.* Bioinformatics, 2006. **22**(4): p. 507-8.

93. Storey, J.D., *The optimal discovery procedure: a new approach to simultaneous significance testing.* UW Biostatistics Working Paper Series Working Paper., 2005. **259**.

94. R:, *A language and environment for statistical computing R Development Core Team.* 2005, R Foundation for Statistical Computing: Vienna, Austria.

95. Ernst, J. and Z. Bar-Joseph, *STEM: a tool for the analysis of short time series gene expression data.* BMC Bioinformatics, 2006. **7**: p. 191.

96. Benjamini, Y. and Y. A Hochberg, *Controlling the false discovery rate: a practical and powerful approach to multiple testing*

J. Roy. Statist. Soc. Ser. B

57: p. 289-300.

97. Ashburner, M., et al., *Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.* Nat Genet, 2000. **25**(1): p. 25-9.

98. Staaf, J., et al., *Normalization of array-CGH data: influence of copy number imbalances.* BMC Genomics, 2007. **8**: p. 382.

99. Rieder, C.L. and R.E. Palazzo, *Colcemid and the mitotic cycle.* J Cell Sci, 1992. **102 (Pt 3)**: p. 387-92.

100. Vig-Varga, E., et al., *Alpha-lipoic acid modulates ovarian surface epithelial cell growth.* Gynecol Oncol, 2006. **103**(1): p. 45-52.

101. Murdoch, W.J. and A.C. McDonnel, *Roles of the ovarian surface epithelium in ovulation and carcinogenesis.* Reproduction, 2002. **123**(6): p. 743-50.

102.Coleman, D.L., et al., *Biology of the Laboratory Mouse*. SECOND EDITION ed, ed. E.L. GREEN. 1968, NEW YORK DOVER PUBLICATIONS, INC.

103.Baker, D.J., et al., *BubR1 insufficiency causes early onset of aging-associated phenotypes and infertility in mice.* Nat Genet, 2004. **36**(7): p. 744-9.

104.Smith, J.R. and O.M. Pereira-Smith, *Replicative senescence: implications for in vivo aging and tumor suppression.* Science, 1996. **273**(5271): p. 63-7.

105.Campisi, J., *Suppressing cancer: the importance of being senescent.* Science, 2005. **309**(5736): p. 886-7.

106.Campisi, J., *Senescent cells, tumor suppression, and organismal aging: good citizens, bad neighbors.* Cell, 2005. **120**(4): p. 513-22.

107.Rajaraman, R., et al., *Stem cells, senescence, neosis and self-renewal in cancer.* Cancer Cell Int, 2006. **6**: p. 25.

108.Kim, N.W., *Specific association of human telomerase activity with immortal cells and cancer.* Science, 1994. **266**: p. 2011-2015.

109.Prowse, K.R. and C.W. Greider, *Developmental and tissue-specific regulation of mouse telomerase and telomere length.* Proc. Natl Acad. Sci. USA, 1995. **92**: p. 4818-4822.

110.Artandi, S.E., *Telomere dysfunction promotes non-reciprocal translocations and epithelial cancers in mice.* Nature, 2000. **406**: p. 641-645.

111.Waterston, R.H., et al., *Initial sequencing and comparative analysis of the mouse genome.* Nature, 2002. **420**(6915): p. 520-62.

112.Boveri, T., *Zur Frage der Entstehung maligner Tumoren (The origin of malignant tumors)*, ed. Jena. 1914: Gustav Fischer.

113.Diaz-Rodriguez, E., et al., *Hec1 overexpression hyperactivates the mitotic checkpoint and induces tumor formation in vivo.* Proc Natl Acad Sci U S A, 2008. **105**(43): p. 16719-24.

114.Sotillo, R., et al., *Mad2 overexpression promotes aneuploidy and tumorigenesis in mice.* Cancer Cell, 2007. **11**(1): p. 9-23.

115.van Deursen, J.M., *Rb loss causes cancer by driving mitosis mad.* Cancer Cell, 2007. **11**(1): p. 1-3.

116.Strohmaier, H., et al., *Human F-box protein hCdc4 targets cyclin E for proteolysis and is mutated in a breast cancer cell line.* Nature, 2001. **413**(6853): p. 316-22.

117.Olesen, S.H., T. Thykjaer, and T.F. Orntoft, *Mitotic checkpoint genes hBUB1, hBUB1B, hBUB3 and TTK in human bladder cancer, screening for mutations and loss of heterozygosity.* Carcinogenesis, 2001. **22**(5): p. 813-5.

118.Dai, W., et al., *Slippage of mitotic arrest and enhanced tumor development in mice with BubR1 haploinsufficiency.* Cancer Res, 2004. **64**(2): p. 440-5.

119.Zhou, H., et al., *Tumour amplified kinase STK15/BTAK induces centrosome amplification, aneuploidy and transformation.* Nat Genet, 1998. **20**(2): p. 189-93.

120.Gisselsson, D., et al., *Abnormal nuclear shape in solid tumors reflects mitotic instability.* Am J Pathol, 2001. **158**(1): p. 199-206.

121.Thompson, S.L. and D.A. Compton, *Examining the link between chromosomal instability and aneuploidy in human cells.* J Cell Biol, 2008. **180**(4): p. 665-72.

122.Levan, A. and J.J. Biesele, *Role of chromosomes in cancerogenesis, as studied in serial tissue culture of mammalian cells.* Ann N Y Acad Sci, 1958. **71**(6): p. 1022-53.

123.Duesberg, P., *Does aneuploidy or mutation start cancer?* Science, 2005. **307**(5706): p. 41.

124.Duesberg, P., et al., *Aneuploidy and cancer: from correlation to causation.* Contrib Microbiol, 2006. **13**: p. 16-44.

125.Li, R., et al., *Aneuploidy vs. gene mutation hypothesis of cancer: recent study claims mutation but is found to support aneuploidy.* Proc Natl Acad Sci U S A, 2000. **97**(7): p. 3236-41.

126.Hansemann, D.v., *Uber asymmetrische zellteilung in epithelkrebsen und deren biologische bedeutung*. Virchow's Arch Path Anat Vol. 119. 1890.

127.Roschke, A.V., et al., *Stable karyotypes in epithelial cancer cell lines despite high rates of ongoing structural and numerical chromosomal instability.* Neoplasia, 2002. **4**(1): p. 19-31.

128.Heng, H.H., et al., *Cancer progression by non-clonal chromosome aberrations.* J Cell Biochem, 2006. **98**(6): p. 1424-35.

129.Heng, H.H., et al., *Clonal and non-clonal chromosome aberrations and genome variation and aberration.* Genome, 2006. **49**(3): p. 195-204.

130.Heng, H.H., et al., *Stochastic cancer progression driven by non-clonal chromosome aberrations.* J Cell Physiol, 2006. **208**(2): p. 461-72.

131.Zimonjic, D., et al., *Derivation of human tumor cells in vitro without widespread genomic instability.* Cancer Res, 2001. **61**(24): p. 8838-44.

132.Mitelman, F., *Recurrent chromosome aberrations in cancer.* Mutat Res, 2000. **462**(2-3): p. 247-53.

133.Kitayama, Y., et al., *Nonrandom chromosomal numerical abnormality predicting prognosis of gastric cancer: a retrospective study of 51 cases using pathology archives.* Lab Invest, 2003. **83**(9): p. 1311-20.

134.Kitayama, Y. and H. Sugimura, *Nonrandom chromosomal numerical abnormality as a new molecular cytogenetic tumor marker--a retrospective study of 60 gastric cancer cases.* Rinsho Byori, 2005. **53**(10): p. 881-6.

135.Jin, C., et al., *Nonrandom pattern of cytogenetic abnormalities in squamous cell carcinoma of the larynx.* Genes Chromosomes Cancer, 2000. **28**(1): p. 66-76.

136.Gorunova, L., et al., *Cytogenetic analysis of pancreatic carcinomas: intratumor heterogeneity and nonrandom pattern of chromosome aberrations.* Genes Chromosomes Cancer, 1998. **23**(2): p. 81-99.

137.Gorunova, L., et al., *Massive cytogenetic heterogeneity in a pancreatic carcinoma: fifty-four karyotypically unrelated clones.* Genes Chromosomes Cancer, 1995. **14**(4): p. 259-66.

138.Jin, Y., et al., *Nonrandom numerical chromosome abnormalities in basal cell carcinomas.* Cancer Genet Cytogenet, 1998. **103**(1): p. 35-42.

139.Duesberg, P., et al., *The chromosomal basis of cancer.* Cell Oncol, 2005. **27**(5-6): p. 293-318.

140.Sieber, O.M., K. Heinimann, and I.P. Tomlinson, *Genomic instability--the engine of tumorigenesis?* Nat Rev Cancer, 2003. **3**(9): p. 701-8.

141.Loeb, L.A., J.H. Bielas, and R.A. Beckman, *Cancers exhibit a mutator phenotype: clinical implications.* Cancer Res, 2008. **68**(10): p. 3551-7; discussion 3557.

142.Li, L., et al., *Cancer-causing karyotypes: chromosomal equilibria between destabilizing aneuploidy and stabilizing selection for oncogenic function.* Cancer Genet Cytogenet, 2009. **188**(1): p. 1-25.

143.Nowell, P.C., *The clonal evolution of tumor cell populations.* Science, 1976. **194**(4260): p. 23-8.

144.Parada, L.A., et al., *Cytogenetic abnormalities and clonal evolution in an adult hepatoblastoma.* Am J Surg Pathol, 1997. **21**(11): p. 1381-6.

145.Jin, C., et al., *Karyotypic heterogeneity and clonal evolution in squamous cell carcinomas of the head and neck.* Cancer Genet Cytogenet, 2002. **132**(2): p. 85-96.

146.Fabarius, A., et al., *Specific clones of spontaneously evolving karyotypes generate individuality of cancers.* Cancer Genet Cytogenet, 2008. **180**(2): p. 89-99.

147.Satagopan, J.M., et al., *Ovarian cancer risk in Ashkenazi Jewish carriers of BRCA1 and BRCA2 mutations.* Clin Cancer Res, 2002. **8**(12): p. 3776-81.

148.Brose, M.S., et al., *Cancer risk estimates for BRCA1 mutation carriers identified in a risk evaluation program.* J Natl Cancer Inst, 2002. **94**(18): p. 1365-72.

149.Boyd, J., *in Ovarian Cancer* ed. F. Sharp, Blackett, T., Berek, J. & Bast, R., Oxford: Isis Medical Media.

150.Moynahan, M.E., T.Y. Cui, and M. Jasin, *Homology-directed dna repair, mitomycin-c resistance, and chromosome stability is restored with correction of a Brca1 mutation.* Cancer Res, 2001. **61**(12): p. 4842-50.

151.Narod, S.A. and W.D. Foulkes, *BRCA1 and BRCA2: 1994 and beyond.* Nat Rev Cancer, 2004. **4**(9): p. 665-76.

152.Jazaeri, A.A., et al., *Gene expression profiles of BRCA1-linked, BRCA2-linked, and sporadic ovarian cancers.* J Natl Cancer Inst, 2002. **94**(13): p. 990-1000.

153.Kawakami, T., et al., *Characterization of loss-of-inactive X in Klinefelter syndrome and female-derived cancer cells.* Oncogene, 2004. **23**(36): p. 6163-9.

154.Sirchia, S.M., et al., *Loss of the inactive X chromosome and replication of the active X in BRCA1-defective and wild-type breast cancer cells.* Cancer Res, 2005. **65**(6): p. 2139-46.

155.Benoit, M.H., et al., *Global analysis of chromosome X gene expression in primary cultures of normal ovarian surface epithelial cells and epithelial ovarian cancer cell lines.* Int J Oncol, 2007. **30**(1): p. 5-17.

156.Eisen, M.B., et al., *Cluster analysis and display of genome-wide expression patterns.* Proc Natl Acad Sci U S A, 1998. **95**(25): p. 14863-8.

157.Tavazoie, S., et al., *Systematic determination of genetic network architecture.* Nat Genet, 1999. **22**(3): p. 281-5.

158.Tamayo, P., et al., *Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation.* Proc Natl Acad Sci U S A, 1999. **96**(6): p. 2907-12.

159.Kutalik, Z., et al., *Advanced significance analysis of microarray data based on weighted resampling: a comparative study and application to gene deletions in Mycobacterium bovis.* Bioinformatics, 2004. **20**(3): p. 357-63.

160.Jakobisiak, M. and J. Golab, *Potential antitumor effects of statins (Review).* Int J Oncol, 2003. **23**(4): p. 1055-69.

161.Keyomarsi, K., et al., *Synchronization of tumor and normal cells from G1 to multiple cell cycles by lovastatin.* Cancer Res, 1991. **51**(13): p. 3602-9.

162.Bouterfa, H.L., et al., *Inhibition of Ras farnesylation by lovastatin leads to downregulation of proliferation and migration in primary cultured human glioblastoma cells.* Anticancer Res, 2000. **20**(4): p. 2761-71.

163.Chapman, E.J. and M.A. Knowles, *Necdin: a multi functional protein with potential tumor suppressor role?* Mol Carcinog, 2009. **48**(11): p. 975-81.

164.Gao, C., et al., *Chromosome instability, chromosome transcriptome, and clonal evolution of tumor cell populations.* Proc Natl Acad Sci U S A, 2007. **104**(21): p. 8995-9000.

165.Dauphinot, L., et al., *The cerebellar transcriptome during postnatal development of the Ts1Cje mouse, a segmental trisomy model for Down syndrome.* Hum Mol Genet, 2005. **14**(3): p. 373-84.

166.Kahlem, P., et al., *Transcript level alterations reflect gene dosage effects across multiple tissues in a mouse model of down syndrome.* Genome Res, 2004. **14**(7): p. 1258-67.

167.Clarke, B., et al., *Intraepithelial T cells and prognosis in ovarian carcinoma: novel associations with stage, tumor type, and BRCA1 loss.* Mod Pathol, 2009. **22**(3): p. 393-402.

168.Rofstad, E.K., et al., *31P NMR spectroscopy and HbO2 cryospectrophotometry in prediction of tumor radioresistance caused by hypoxia.* Int J Radiat Oncol Biol Phys, 1989. **16**(4): p. 919-23.

169.Rofstad, E.K., B.M. Fenton, and R.M. Sutherland, *Intracapillary HbO2 saturations in murine tumours and human tumour xenografts measured by cryospectrophotometry: relationship to tumour volume, tumour pH and fraction of radiobiologically hypoxic cells.* Br J Cancer, 1988. **57**(5): p. 494-502.

170.Nelson, D.A., et al., *Hypoxia and defective apoptosis drive genomic instability and tumorigenesis.* Genes Dev, 2004. **18**(17): p. 2095-107.

171.Kops, G.J., D.R. Foltz, and D.W. Cleveland, *Lethality to human cancer cells through massive chromosome loss by inhibition of the mitotic checkpoint.* Proc Natl Acad Sci U S A, 2004. **101**(23): p. 8699-704.

172.Emens, L.A. and E.M. Jaffee, *Cancer vaccines: an old idea comes of age.* Cancer Biol Ther, 2003. **2**(4 Suppl 1): p. S161-8.

173.Cross, D. and J.K. Burmester, *Gene therapy for cancer treatment: past, present and future.* Clin Med Res, 2006. **4**(3): p. 218-27.

174.Ries, L., et al., *SEER Cancer Statistics Review, 1975-2004,* . 2004, National Cancer Institute. : Bethesda, MD.

175.Druker, B., et al., *Long-term benefits of imatinib (IM) for patients newly diagnosed with chronic myelogenous leukemia in chronic phase (CML-CP): The 5-year update from the IRIS study [abstract].* Journal of Clinical Oncology, 2006(24:18S. Abstract no. 6506).

176.Wooster, R., et al., *Identification of the breast cancer susceptibility gene BRCA2.* Nature, 1995. **378**(6559): p. 789-92.

177.Wooster, R. and M.R. Stratton, *Breast cancer susceptibility: a complex disease unravels.* Trends Genet, 1995. **11**(1): p. 3-5.

178.Boveri, T., *Concerning the Origin of Malignant Tumours*. 2008, The Company of Biologists Limited and Cold Spring Harbor Laboratory Press.

179.Kabil, A., E. Silva, and A. Kortenkamp, *Estrogens and genomic instability in human breast cancer cells--involvement of Src/Raf/Erk signaling in micronucleus formation by estrogenic chemicals.* Carcinogenesis, 2008. **29**(10): p. 1862-8.

180.Cavenee, W.K., et al., *Expression of recessive alleles by chromosomal mechanisms in retinoblastoma.* Nature, 1983. **305**(5937): p. 779-84.

181.Oppermann, H., et al., *Uninfected vertebrate cells contain a protein that is closely related to the product of the avian sarcoma virus transforming gene (src).* Proc Natl Acad Sci U S A, 1979. **76**(4): p. 1804-8.

182.Knudson, A.G., Jr., *Mutation and cancer: statistical study of retinoblastoma.* Proc Natl Acad Sci U S A, 1971. **68**(4): p. 820-3.

183.Vogelstein, B. and K.W. Kinzler, *The multistep nature of cancer.* Trends Genet, 1993. **9**(4): p. 138-41.

184.Hahn, W.C., et al., *Creation of human tumour cells with defined genetic elements.* Nature, 1999. **400**(6743): p. 464-8.

185.Kendall, S.D., et al., *A network of genetic events sufficient to convert normal human cells to a tumorigenic state.* Cancer Res, 2005. **65**(21): p. 9824-8.

186.Duesberg, P.H., *Are cancers dependent on oncogenes or on aneuploidy?* Cancer Genet Cytogenet, 2003. **143**(1): p. 89-91.

187.Ewald, D., et al., *Time-sensitive reversal of hyperplasia in transgenic mice expressing SV40 T antigen.* Science, 1996. **273**(5280): p. 1384-6.

188.Klein, A., et al., *Gene expression profiling: cell cycle deregulation and aneuploidy do not cause breast cancer formation in WAP-SVT/t transgenic animals.* J Mol Med, 2005. **83**(5): p. 362-76.

189.Jonkers, J. and A. Berns, *Oncogene addiction: sometimes a temporary slavery.* Cancer Cell, 2004. **6**(6): p. 535-8.

190.Monteiro, A.N.A., *BRCA1: the enigma of tissue-specific tumor development.* Trends in Genetics, 2003. **19**(6): p. 312-315.

191.Malone, K.E., et al., *BRCA1 mutations and breast cancer in the general population: analyses in women before age 35 years and in women before age 45 years with first-degree family history.* JAMA, 1998. **279**(12): p. 922-9.

192.Thompson, D. and D.F. Easton, *Cancer Incidence in BRCA1 mutation carriers.* J Natl Cancer Inst, 2002. **94**(18): p. 1358-65.

193.Heng, H.H., et al., *Imaging genome abnormalities in cancer research.* Cell Chromosome, 2004. **3**(1): p. 1.

194.Makino, S., *Further evidence favoring the concept of the stem cell in ascites tumors of rats.* Ann N Y Acad Sci, 1956. **63**(5): p. 818-30.

195.Hauschka, T.S., *The chromosomes in ontogeny and oncogeny.* Cancer Res, 1961. **21**: p. 957-74.

196.de Grouchy, J. and C. de Nava, *A chromosomal theory of carcinogenesis.* Ann Intern Med, 1968. **69**(2): p. 381-91.

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

198.Gonzalez-Garcia, I., R.V. Sole, and J. Costa, *Metapopulation dynamics and spatial heterogeneity in cancer.* Proc Natl Acad Sci U S A, 2002. **99**(20): p. 13085-9.

199.Harada, T., et al., *Interglandular cytogenetic heterogeneity detected by comparative genomic hybridization in pancreatic cancer.* Cancer Res, 2002. **62**(3): p. 835-9.

200.Murphy, D.S., et al., *Characterization of extensive genetic alterations in ductal carcinoma in situ by fluorescence in situ hybridization and molecular analysis.* J Natl Cancer Inst, 1995. **87**(22): p. 1694-704.

201.Ganem, N.J., Z. Storchova, and D. Pellman, *Tetraploidy, aneuploidy and cancer.* Curr Opin Genet Dev, 2007. **17**(2): p. 157-62.

202.Duesberg, P., A. Fabarius, and R. Hehlmann, *Aneuploidy, the primary cause of the multilateral genomic instability of neoplastic and preneoplastic cells.* IUBMB Life, 2004. **56**(2): p. 65-81.

203.Sunavala-Dossabhoy, G., et al., *A dominant negative mutant of TLK1 causes chromosome missegregation and aneuploidy in normal breast epithelial cells.* BMC Cell Biol, 2003. **4**: p. 16.

204.Smith, L., A. Plug, and M. Thayer, *Delayed replication timing leads to delayed mitotic chromosome condensation and chromosomal instability of chromosome translocations.* Proc Natl Acad Sci U S A, 2001. **98**(23): p. 13300-5.

205.Schvartzman, J.M., R. Sotillo, and R. Benezra, *Mitotic chromosomal instability and cancer: mouse modelling of the human disease.* Nat Rev Cancer, 2010. **10**(2): p. 102-15.

206.Heng, H.H., *The genome-centric concept: resynthesis of evolutionary theory.* Bioessays, 2009. **31**(5): p. 512-25.

207.Li, R., D. Rasnick, and P. Duesberg, *Correspondence re: D. Zimonjic et al., Derivation of human tumor cells in vitro without widespread genomic instability. Cancer Res., 61: 8838-8844, 2001.* Cancer Res, 2002. **62**(21): p. 6345-8; author reply 6348-9.

208.Mahale, A.M., et al., *Clonal selection in malignant transformation of human fibroblasts transduced with defined cellular oncogenes.* Cancer Res, 2008. **68**(5): p. 1417-26.

209.Kinsella, A.R., et al., *Introduction of the activated N-ras oncogene into human fibroblasts by retroviral vector induces morphological transformation and tumorigenicity.* Carcinogenesis, 1990. **11**(10): p. 1803-9.

210.Akagi, T., K. Sasai, and H. Hanafusa, *Refractory nature of normal human diploid fibroblasts with respect to oncogene-mediated transformation.* Proc Natl Acad Sci U S A, 2003. **100**(23): p. 13567-72.

211.Stindl, R., *Defining the steps that lead to cancer: replicative telomere erosion, aneuploidy and an epigenetic maturation arrest of tissue stem cells.* Med Hypotheses, 2008. **71**(1): p. 126-40.

212.Heng, H.H., et al., *Patterns of genome dynamics and cancer evolution.* Cell Oncol, 2008. **30**(6): p. 513-4.

213.Castro, M.A., et al., *Chromosome aberrations in solid tumors have a stochastic nature.* Mutat Res, 2006. **600**(1-2): p. 150-64.

214.Franklin, W.A., et al., *Widely dispersed p53 mutation in respiratory epithelium. A novel mechanism for field carcinogenesis.* J Clin Invest, 1997. **100**(8): p. 2133-7.

215.Habuchi, T., *Origin of multifocal carcinomas of the bladder and upper urinary tract: molecular analysis and clinical implications.* Int J Urol, 2005. **12**(8): p. 709-16.

216.Maley, C.C., et al., *Selectively advantageous mutations and hitchhikers in neoplasms: p16 lesions are selected in Barrett's esophagus.* Cancer Res, 2004. **64**(10): p. 3414-27.

217.Jonason, A.S., et al., *Frequent clones of p53-mutated keratinocytes in normal human skin.* Proc Natl Acad Sci U S A, 1996. **93**(24): p. 14025-9.

218.Brentnall, T.A., et al., *Mutations in the p53 gene: an early marker of neoplastic progression in ulcerative colitis.* Gastroenterology, 1994. **107**(2): p. 369-78.

219.Bielas, J.H., et al., *Human cancers express a mutator phenotype.* Proc Natl Acad Sci U S A, 2006. **103**(48): p. 18238-42.

220.Bielas, J.H. and L.A. Loeb, *Mutator phenotype in cancer: timing and perspectives.* Environ Mol Mutagen, 2005. **45**(2-3): p. 206-13.

221.Shaffer LG, T.N., *ISCN, An International System for Human Cytogenetic Nomenclature.*, ed. S.K.A. Basel. 2005.

222.Albertson, D.G. and D. Pinkel, *Genomic microarrays in human genetic disease and cancer.* Hum Mol Genet, 2003. **12 Spec No 2**: p. R145-52.

223.Albertson, D.G., *Profiling breast cancer by array CGH.* Breast Cancer Res Treat, 2003. **78**(3): p. 289-98.

224.Mitelman, F., et al., *Clinical significance of cytogenetic findings in solid tumors.* Cancer Genet Cytogenet, 1997. **95**(1): p. 1-8.

225.Sjoblom, T., et al., *The consensus coding sequences of human breast and colorectal cancers.* Science, 2006. **314**(5797): p. 268-74.

226.Vogelstein, B. and K.W. Kinzler, *Cancer genes and the pathways they control.* Nat Med, 2004. **10**(8): p. 789-99.

227.Weinstein, I.B., *Disorders in cell circuitry during multistage carcinogenesis: the role of homeostasis.* Carcinogenesis, 2000. **21**(5): p. 857-64.

228.Colomer, R., et al., *erbB-2 antisense oligonucleotides inhibit the proliferation of breast carcinoma cells with erbB-2 oncogene amplification.* Br J Cancer, 1994. **70**(5): p. 819-25.

229.Duesberg, P., R. Stindl, and R. Hehlmann, *Origin of multidrug resistance in cells with and without multidrug resistance genes: chromosome reassortments catalyzed by aneuploidy.* Proc Natl Acad Sci U S A, 2001. **98**(20): p. 11283-8.

230.Lehninger, A., *: The molecular logic of living organisms, Part 1 and Part 4*, in *Biochemistry. The Molecular Basis of Cell Structure and Function*. 1981, Worth, 1981: New York. p. 3–14, 17–360, 854–1056.

231.Darwin, C., *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life* 1859.

232.Waliszewski, P., *Complexity, dynamic cellular network, and tumorigenesis.* Pol J Pathol, 1997. **48**(4): p. 235-41.

233.Moser, J., *Convergent series expansion of quasi-periodic motions.* Math Ann, 1967. **169**: p. 163.

234.Karlsson, A., et al., *Genomically complex lymphomas undergo sustained tumor regression upon MYC inactivation unless they acquire novel chromosomal translocations.* Blood, 2003. **101**(7): p. 2797-803.

235.Navin, N., et al., *Inferring tumor progression from genomic heterogeneity.* Genome Res, 2010. **20**(1): p. 68-80.

236.Griffin, C.A., et al., *Consistent chromosome abnormalities in adenocarcinoma of the pancreas.* Cancer Res, 1995. **55**(11): p. 2394-9.

237.Heim, S., et al., *Marker ring chromosome--a new cytogenetic abnormality characterizing lipogenic tumors?* Cancer Genet Cytogenet, 1987. **24**(2): p. 319-26.

238.Kiechle-Schwarz, M., et al., *Recurrent cytogenetic aberrations in human ovarian carcinomas.* Cancer Detect Prev, 1995. **19**(3): p. 234-43.

239.Orndal, C., et al., *Supernumerary ring chromosomes in five bone and soft tissue tumors of low or borderline malignancy.* Cancer Genet Cytogenet, 1992. **60**(2): p. 170-5.

240.Pejovic, T., et al., *Chromosome aberrations in 35 primary ovarian carcinomas.* Genes Chromosomes Cancer, 1992. **4**(1): p. 58-68.

241.Roberts, C.G. and M.H. Tattersall, *Cytogenetic study of solid ovarian tumors.* Cancer Genet Cytogenet, 1990. **48**(2): p. 243-53.

242.Tredan, O., et al., *Drug resistance and the solid tumor microenvironment.* J Natl Cancer Inst, 2007. **99**(19): p. 1441-54.

243.Haskill, S., et al., *Detection of intratumor heterogeneity by simultaneous multiparameter flow cytometric analysis with enzyme and DNA markers.* Cancer Res, 1983. **43**(3): p. 1003-9.

244.Jones, A.M., et al., *Array-CGH analysis of microsatellite-stable, near-diploid bowel cancers and comparison with other types of colorectal carcinoma.* Oncogene, 2005. **24**(1): p. 118-29.

245.Kallioniemi, A., *CGH microarrays and cancer.* Curr Opin Biotechnol, 2008. **19**(1): p. 36-40.

246.Hu, W., et al., *A single nucleotide polymorphism in the MDM2 gene disrupts the oscillation of p53 and MDM2 levels in cells.* Cancer Res, 2007. **67**(6): p. 2757-65.

247.Tsao, J.L., et al., *Genetic reconstruction of individual colorectal tumor histories.* Proc Natl Acad Sci U S A, 2000. **97**(3): p. 1236-41.

248.Goel, A., et al., *Frequent inactivation of PTEN by promoter hypermethylation in microsatellite instability-high sporadic colorectal cancers.* Cancer Res, 2004. **64**(9): p. 3014-21.

249.Barrett, M.T., et al., *Evolution of neoplastic cell lineages in Barrett oesophagus.* Nat Genet, 1999. **22**(1): p. 106-9.

250.Stephens, P.J., et al., *Complex landscapes of somatic rearrangement in human breast cancer genomes.* Nature, 2009. **462**(7276): p. 1005-10.

251.Clark, M.J., et al., *U87MG Decoded: The Genomic Sequence of a Cytogenetically Aberrant Human Cancer Cell Line.* PLoS Genet, 2010. **6**(1): p. e1000832.

252.Perry, P. and S. Wolff, *New Giemsa method for the differential staining of sister chromatids.* Nature, 1974. **251**(5471): p. 156-8.

253.Fujiwara, T., et al., *Cytokinesis failure generating tetraploids promotes tumorigenesis in p53-null cells.* Nature, 2005. **437**(7061): p. 1043-7.

254.Kiechle, M., et al., *Genetic imbalances in precursor lesions of endometrial cancer detected by comparative genomic hybridization.* Am J Pathol, 2000. **156**(6): p. 1827-33.

255.Collins, F.S. and A.D. Barker, *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, 2007. **296**(3): p. 50-7.

256.Pleasance, E.D., et al., *A comprehensive catalogue of somatic mutations from a human cancer genome.* Nature, 2010. **463**(7278): p. 191-U73.

257.*Making the paper: Michael Stratton & Peter Campbell.* Nature, 2010. **463**(7278): p. 134-134.

258.Cai, L., N. Friedman, and X.S. Xie, *Stochastic protein expression in individual cells at the single molecule level.* Nature, 2006. **440**(7082): p. 358-62.

259.Levsky, J.M., et al., *Single-cell gene expression profiling.* Science, 2002. **297**(5582): p. 836-40.

260.Weinstein, I.B. and A. Joe, *Oncogene addiction.* Cancer Res, 2008. **68**(9): p. 3077-80; discussion 3080.

261.Cramer, K., et al., *BCR/ABL and other kinases from chronic myeloproliferative disorders stimulate single-strand annealing, an unfaithful DNA double-strand break repair.* Cancer Res, 2008. **68**(17): p. 6884-8.

262.Patel, H. and M.Y. Gordon, *Abnormal centrosome-centriole cycle in chronic myeloid leukaemia?* Br J Haematol, 2009. **146**(4): p. 408-17.

263.Khorashad, J.S., et al., *Multiple sub-microscopic genomic lesions are a universal feature of chronic myeloid leukaemia at diagnosis.* Leukemia, 2008. **22**(9): p. 1806-7.

264.Gorre, M.E., et al., *Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification.* Science, 2001. **293**(5531): p. 876-80.

265.Gorre, M.E. and C.L. Sawyers, *Molecular mechanisms of resistance to STI571 in chronic myeloid leukemia.* Curr Opin Hematol, 2002. **9**(4): p. 303-7.

266.Hochhaus, A., et al., *Molecular and chromosomal mechanisms of resistance to imatinib (STI571) therapy.* Leukemia, 2002. **16**(11): p. 2190-6.

267.Druker, B.J., *Circumventing resistance to kinase-inhibitor therapy.* N Engl J Med, 2006. **354**(24): p. 2594-6.

268.Druker, B.J., et al., *Activity of a specific inhibitor of the BCR-ABL tyrosine kinase in the blast crisis of chronic myeloid leukemia and acute lymphoblastic leukemia with the Philadelphia chromosome.* N Engl J Med, 2001. **344**(14): p. 1038-42.

269.Ottmann, O.G., et al., *A phase 2 study of imatinib in patients with relapsed or refractory Philadelphia chromosome-positive acute lymphoid leukemias.* Blood, 2002. **100**(6): p. 1965-71.

270.Talpaz, M., et al., *Imatinib induces durable hematologic and cytogenetic responses in patients with accelerated phase chronic myeloid leukemia: results of a phase 2 study.* Blood, 2002. **99**(6): p. 1928-37.

271.Druker, B.J., et al., *Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia.* N Engl J Med, 2001. **344**(14): p. 1031-7.

272.Sawyers, C.L., et al., *Imatinib induces hematologic and cytogenetic responses in patients with chronic myelogenous leukemia in myeloid blast crisis: results of a phase II study.* Blood, 2002. **99**(10): p. 3530-9.

273.Roumiantsev, S., et al., *Clinical resistance to the kinase inhibitor STI-571 in chronic myeloid leukemia by mutation of Tyr-253 in the Abl kinase domain P-loop.* Proc Natl Acad Sci U S A, 2002. **99**(16): p. 10700-5.

274.Wood, L.D., et al., *The genomic landscapes of human breast and colorectal cancers.* Science, 2007. **318**(5853): p. 1108-13.

275.Schimke, R.T., *Gene amplification, drug resistance, and cancer.* Cancer Res, 1984. **44**(5): p. 1735-42.

276.Trent, J.M., et al., *Cytologic evidence for gene amplification in methotrexate-resistant cells obtained from a patient with ovarian adenocarcinoma.* J Clin Oncol, 1984. **2**(1): p. 8-15.

277.Haapala, K., et al., *Androgen receptor amplification is associated with increased cell proliferation in prostate cancer.* Hum Pathol, 2007. **38**(3): p. 474-8.

278.Chen, G., et al., *Karyotypic change from heteroploidy to near diploidy associated with development of cisplatin resistance in a rat ovarian tumour cell line.* J Cancer Res Clin Oncol, 1991. **117**(6): p. 539-42.

279.Kotchetkov, R., et al., *Increased malignant behavior in neuroblastoma cells with acquired multi-drug resistance does not depend on P-gp expression.* Int J Oncol, 2005. **27**(4): p. 1029-37.

280.Michor, F., M.A. Nowak, and Y. Iwasa, *Evolution of resistance to cancer therapy.* Curr Pharm Des, 2006. **12**(3): p. 261-71.

281.Encarnacion, C.A., et al., *Measurement of steroid hormone receptors in breast cancer patients on tamoxifen.* Breast Cancer Res Treat, 1993. **26**(3): p. 237-46.

282.Daffada, A.A., et al., *Exon 5 deletion variant estrogen receptor messenger RNA expression in relation to tamoxifen resistance and progesterone receptor/pS2 status in human breast cancer.* Cancer Res, 1995. **55**(2): p. 288-93.

283.Johnston, S.R., et al., *Changes in estrogen receptor, progesterone receptor, and pS2 expression in tamoxifen-resistant human breast cancer.* Cancer Res, 1995. **55**(15): p. 3331-8.

284.Jin, Q. and F.J. Esteva, *Cross-talk between the ErbB/HER family and the type I insulin-like growth factor receptor signaling pathway in breast cancer.* J Mammary Gland Biol Neoplasia, 2008. **13**(4): p. 485-98.

285.Shou, J., et al., *Mechanisms of tamoxifen resistance: increased estrogen receptor-HER2/neu cross-talk in ER/HER2-positive breast cancer.* J Natl Cancer Inst, 2004. **96**(12): p. 926-35.

286.Sadovsky, Y. and S. Adler, *Selective modulation of estrogen receptor action.* J Clin Endocrinol Metab, 1998. **83**(1): p. 3-5.

287.Gascoigne, K.E. and S.S. Taylor, *Cancer cells display profound intra- and interline variation following prolonged exposure to antimitotic drugs.* Cancer Cell, 2008. **14**(2): p. 111-22.

288.Huang, H.C., et al., *Evidence that mitotic exit is a better cancer therapeutic target than spindle assembly.* Cancer Cell, 2009. **16**(4): p. 347-58.

289.Kwon, M., et al., *Mechanisms to suppress multipolar divisions in cancer cells with extra centrosomes.* Genes Dev, 2008. **22**(16): p. 2189-203.

290.Deeb, K.K., D.L. Trump, and C.S. Johnson, *Vitamin D signalling pathways in cancer: potential for anticancer therapeutics.* Nat Rev Cancer, 2007. **7**(9): p. 684-700.

291.Chatterjee, M., *Vitamin D and Genomic Stability.* Mutation Research, 2000. **475**: p. 69-88.

292.Chakraborty, T., et al., *Molecular basis of anticlastogenic potential of vanadium in vivo during the early stages of diethylnitrosamine-induced hepatocarcinogenesis in rats.* Mutat Res, 2006. **609**(2): p. 117-28.

293.Sarkar, A., et al., *Anticlastogenic potential of 1alpha,25-dihydroxyvitamin D3 in murine lymphoma.* Cancer Lett, 2000. **150**(1): p. 1-13.

294.Demasters, G., et al., *Potentiation of radiation sensitivity in breast tumor cells by the vitamin D3 analogue, EB 1089, through promotion of autophagy and interference with proliferative recovery.* Mol Cancer Ther, 2006. **5**(11): p. 2786-97.

295.Pienta, K.J., *Modeling Cancer as A Complex Adaptive System: Genetic Instability and Evolution* in *Complex Systems Science in Biomedicine*, D. T, Editor. 2006, Springer US. p. 537-556.

ABSTRACT

**TRACKING PROFILES OF GENOMIC INSTABILITY
IN SPONTANEOUS TRANSFORMATION AND TUMORIGENESIS**
by

LESLEY LAWRENSON

AUGUST 2010

Advisor: Dr. Wayne Lancaster

Major: Molecular Medicine and Genetics

Degree: Doctor of Philosophy

The dominant paradigm for cancer research focuses on the identification of specific genes for cancer causation and for the discovery of therapeutic targets. Alternatively, the current data emphasize the significance of karyotype heterogeneity in cancer progression over specific gene-based causes of cancer. Variability of a magnitude significant to shift cell populations from homogeneous diploid cells to a mosaic of structural and numerical chromosome alterations reflects the characteristic low-fidelity genome transfer of cancer cell populations. This transition marks the departure from micro-evolutionary gene-level change to macro-evolutionary change that facilitates the generation of many unique karyotypes within a cell population. Considering cancer cell populations to be complex-adaptive systems, multi-level analyses were performed longitudinally including whole genome microarray, population karyotype analysis, and determination of cell phenotype. As heterogeneity in ovarian cancer at each of these levels is linked to low survival, metastasis, and resistance to chemotherapy, a syngeneic model of spontaneous ovarian cancer development was employed. The significant findings of the current study are, 1) Genomic instability was apparent from the earliest stages of study, 2) Karyotypic heterogeneity was widespread, showed a pattern of expansion over time and preceded the acquisition of the transformed phenotype 3) a major karyotypic shift occurred between

transformed cells *in vivo* and tumors formed *in vitro*, documenting the formation of a new system induced by environmental change 4) Chromosome copy number has greater impact on gene expression in early-stage cell populations, where karyotypes are beginning to depart from the diploid genome. A genome-centered paradigm for transformation is emphasized through the discovery of early large-scale increases in karyotype heterogeneity. This occurred well before the appearance of the transformed phenotype, arose much faster in *Brca1* conditionally inactivated cells, was linked to the largest shift in gene expression, and was linked to the transition from *in vitro* to *in vivo* survival facilitating tumorigenesis. These data demonstrate the significance, methodologies and rationale for quantifying karyotype heterogeneity in transformation, tumorigenesis, and clinical cancers. Together, these findings support of a genome-centered evolutionary framework for cancer progression that emphasizes cell-to-cell genomic variability as the basis for macro-evolutionary selection and rapid phenotypic switching in response to new environments.

AUTOBIOGRAPHICAL STATEMENT

Biographical Sketch - Lesley Eileen Lawrenson, MS		POSITION/ TITLE – Md/PhD Candidate	
EDUCATION/TRAINING BY INSTITUTION AND LOCATION	DEGREE	YEAR(s)	FIELD OF STUDY
University of California San Diego, San Diego, CA	BS	2001	Biochemistry/ Cell Biology
University of California San Diego, San Diego, CA	MS	2003	Biology/ Physiology
Wayne State University School of Medicine, Detroit, Michigan	MD/PhD	current	Molecular Biol. / Genetics

A. Academic and Professional Honors

DATE	AWARD
2/2010	NCI Cancer Research Imaging Camp Scholarship Award
3/2010	Days of Molecular Medicine MD/PhD Scholarship Travel Award
2009	Research Excellence Award – Wayne State University School of Medicine
2008	DOD CDMRP Training Fellowship: Genomic Instability in Brest Cancer Progression
2004 – current	Wayne State University Board of Governors Merit Scholarship
2006	Wayne State University "Service to the City of Detroit" Award
2003	*American College of Sports Medicine* - National Graduate Student of the Year
2003	*Academic Senate* Distinguished Teaching Award
2002 – 2003	Provost's Honors (GPA 3.6 and above)

B. Peer Reviewed Journal Articles

1. Bailey DM, Young IS, McEneny J, **Lawrenson L**, Kim J, Barden J, Richardson RS. Regulation of free radical outflow from an isolated muscle bed in exercising humans. *American Journal of Heart and Circulatory Physiology.* 287: 1689-1699, 2004.
2. Bailey DM, **Lawrenson L**, Mceneny J, Young IS, James PE, Jackson SK, Henry R, Mathieu-Costello O, McCord J, Richardson RS. Electron paramagnetic spectroscopic evidence of exercise-induced free radical accumulation in human skeletal muscle. *Free Radical Research.* 2007 Feb; 41(2):182-90.
3. Barden J, **Lawrenson L**, Poole JG, Kim J, Wray D, Richardson RS. Limitations to maximal metabolic rate in trained human skeletal muscle: hypoxia, hyperoxia and beyond. *American Journal of Heart and Circulatory Physiology.* 2007 May;292(5):H2491-7. Epub 2007 Jan 26.
4. Chugani HC, **Lawrenson L**, Sood S, Musik O, Juhasz C. Increased Ipsilateral Striatal Serotonin Synthesis Following Cortical Resection. *Epilepsy Research* (In press).
5. Donato AJ, Uberoi A, Wray DW, Nishiyama S, **Lawrenson L**, Richardson RS. Differential effects of aging on limb blood flow in humans. Am J Physiol Heart Circ Physiol. 2006. Jan;290(1):H272-8.
6. Heng, H.H., J.B. Stevens, L. Lawrenson, G. Liu, K.J. Ye, S.W. Bremer, and C.J. Ye, *Patterns of genome dynamics and cancer evolution.* Cell Oncol, 2008. **30**(6): p. 513-4.
7. L**awrenson L**, Hoff J, Richardson RS. Aging attenuates vascular and metabolic plasticity but does not limit Improvement in Muscle VO_2max. *American Journal of Heart and Circulatory Physiology.* 286(4): 1565-1572, 2003.
8. **Lawrenson L**, Poole JG, Kim J, Brown CF, Richardson RS. Vascular and metabolic response to isolated small muscle mass exercise: the effect of age. *American Journal of Heart and Circulatory Physiology.* 285: 1023-1031, 2003.
9. Poole JG, **Lawrenson L**, Kim J, Brown C, Richardson RS. Vascular endothelial reactivity mediated by [HbO_2]: the effect of physical activity. *American Journal of Physiology. Submitted* 4/12/07.
10. Poole JG, **Lawrenson L**, Kim J, Brown C, Richardson RS. Vascular and metabolic response to cycle exercise in sedentary humans: effect of age. *American Journal of Heart and Circulatory Physiology.* 284: 1251-1259, 2003.
11. Wray, D.W., A. Uberoi, L. Lawrenson, D.M. Bailey, and R.S. Richardson, *Oral antioxidants and cardiovascular health in the exercise-trained and untrained elderly: a radically different outcome.* Clin Sci (Lond), 2009. **116**(5): p. 433-41.
12. Wray DW, Uberoi A, **Lawrenson L**, Richardson RS. Evidence of preserved endothelial function and vascular plasticity with age. Am J Physiol Heart Circ Physiol. 2006 Mar;290(3):H1271-7.
13. Wray DW, Uberoi A, **Lawrenson L**, Richardson RS. Heterogeneous limb vascular responsiveness to shear stimuli during dynamic exercise in humans. *Journal of Applied Physiology*. 99: 81-86, 2005.
14. Richardson, R.S., A.J. Donato, A. Uberoi, D.W. Wray, **L. Lawrenson**, S. Nishiyama, and D.M. Bailey, *Exercise-induced brachial artery vasodilation: role of free radicals.* Am J Physiol Heart Circ Physiol, 2007. **292**(3): p. H1516-22.
15. Wray, D.W., A. Uberoi, **L. Lawrenson**, D.M. Bailey, and R.S. Richardson, *Oral antioxidants and cardiovascular health in the exercise-trained and untrained elderly: a radically different outcome.* Clin Sci (Lond), 2009. **116**(5): p. 433-41.
16. 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, *Genome based cell population heterogeneity promotes tumorigenicity: the evolutionary mechanism of cancer.* J Cell Physiol, 2009. **219**(2): p. 288-300.