Defining the role of aneuploidy throughout tumorigenesis

by

Rebecca Estelle Silberman

B.A. Biology
Colgate University

Submitted to the Department of Biology
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2021

Signature of Author_____
Department of Biology
December 22nd, 2020

Certified by_____
Angelika Amon
Kathleen and Curtis Marble Professor of Cancer Research
Investigator, Howard Hughes Medical Institute
Thesis Supervisor

Accepted by_____
Amy Keating
Professor of Biology and Biological Engineering
Co-Director, Biology Graduate Committee

Defining the role of aneuploidy throughout tumorigenesis

by

Rebecca Estelle Silberman

Submitted to the Department of Biology on December 22nd, 2020 in Partial Fulfilment of the Requirements for the Degree of Doctor of Philosophy in Biology

Abstract

Aneuploidy is a state of genome imbalance which alters the copy number of whole chromosomes. While aneuploidy is rare in healthy tissues, it is one of the most common features of cancerous tumors. Studies of aneuploid yeast and aneuploid mammalian cells growing in culture revealed that aneuploidy induces cellular stress and slows proliferation. So it is surprising that aneuploidy is a hallmark of cancer, a disease of cellular over-proliferation and inappropriate cell survival. We sought to elucidate aneuploidy's role in tumorigenesis by defining the factors that affect the prevalence of aneuploid cells in normal, pre-cancerous, and cancerous tissues.

First, we investigated whether aneuploid mammalian cells experience fitness defects *in vivo*. We found that aneuploidy decreases hematopoietic stem cells' fitness and that aneuploid cells are selected against in normal, regenerating tissues *in vivo*. However, we also found that aneuploid cells can accumulate in the hematopoietic system when purifying selection is relaxed following bone marrow reconstitution. We then sought to extend our observations to the context of pre-cancerous tissues. We analyzed the prevalence of aneuploidy in the highly tumorigenic, but histologically normal tissues of women harboring heterozygous germline *BRCA2* mutations. Using single-cell sequencing, we revealed that breast cells from *BRCA2* mutation carriers lack aneuploidy but feature a distinct form of genome imbalance called sub-chromosomal copy number variants (CNVs), even before the initiation of tumorigenesis. We then analyzed the timing with which these two forms of genome imbalance—whole-chromosomal aneuploidy and sub-chromosomal CNVs—arise during tumorigenesis. We found that CNVs are present in the cells of early precursors of multiple cancers, but that whole-chromosomal aneuploidy arises late in tumorigenesis. Our findings propose that whole-chromosomal aneuploidy reduces cells' fitness in both normal and pre-cancerous tissues, and that aneuploidy is selected against throughout tumorigenesis. This has implications for the role of aneuploidy in cancer, suggesting that aneuploidy does not contribute to early tumorigenesis.

Thesis Supervisor: Angelika Amon
Title: Kathleen and Curtis Marble Professor of Cancer Research

Acknowledgments

Thank you to my mentor, Angelika. To praise her brilliance, her creativity, her enthusiasm, her unfailing integrity, and her bravado (both scientific and otherwise) still doesn't do her justice. She was an immensely generous person, who cared deeply about each member of her lab. She concerned herself not just with our science and our ambitions, but with our lives. I knew every day that I was in the lab—whether my experiments were going well or going poorly—that Angelika supported me and that we would find a way. It is a great honor to have been her student.

I would also like to thank Frank Solomon for his mentorship. Frank proves what I have long suspected: that people who love horses are the best people. Thank you to the members of my thesis committee Jackie Lees and Mike Hemann for their insight and guidance. I am grateful to have conducted this work alongside a network of generous collaborators both at MIT and beyond. I am indebted to the members of the KI histology core lab, KI flow cytometry core lab, and BioMicro Center who made critical contributions to many of the experiments presented in this thesis. Thank you also to the members of the Division of Comparative Medicine—especially the veterinarians and mouse technicians—who make it possible for us to work with laboratory animals.

I would like to thank the members of the Amon lab, past and present. It has been an honor to be a member of this group, who balance scientific rigor with joy. In particular, thank you to my fellow graduate students. Thank you to my spectacular undergraduates, Nicole and Emily. And thank you to Erica Burds for taking care of absolutely everything and always having emergency dog pictures at the ready.

Thank you to my friends for what I can only describe as everything. Thank you to Carolyn, who puts everything in perspective. Thank you to Kelsie, who always picks up the phone. Thank you to Cassie, who keeps me sane, laughing, and caffeinated. Thank you to Hannah, who always reminds me that the world exists beyond my lab bench. Thank you to Megan, who finds the humor in everything. Thank you to Emma, who is my own personal Meg Linehan. And, of course, thank you to the pets: Silas, Papaya, Duncan, Luna, Roka, Fox Princess, Nova, Dana Scully, Hestia, and Murder.

Finally, I would like to thank my family. Thank you all for your steadfast love, your patience, and your senses of humor. Thank you to Julia for never letting me take things too seriously. Thank you to Mom for sharing your school with me. Thank you to Dad for learning alongside me and always asking about my research.

Table of Contents

Contributions to Each Chapter

Chapter One

I wrote Chapter One with input and minor edits from Angelika Amon.

Chapter Two

I designed and performed experiments and analyses which appear in the following figures: Figure 2E, Figure 5D, Figure 6A-F, Sup. Figure S1A-D, Sup. Fig S2K, Sup. Figure S3V, Sup Figure S5I, Sup. Figure S6J, Sup. Figure S6T, Sup. Figure S7, and Sup. Figure S8A-H. Sarah Pfau and Angelika Amon wrote the manuscript with input from myself and others.

Chapter Three

I designed and performed experiments and analyses which appear in Figure 1a and Sup. Figure 13. Andrew Rozhok and James DeGregori wrote the manuscript with input from myself and others.

Chapter Four

I designed and performed experiments and analyses which appear in the following figures: Figure 1B-D, Figure 2A-C, and Figure 3C. Mihriban Karaayvaz and Leif Ellisen wrote the manuscript with input from myself and others.

Chapter Five

I designed, performed and analyzed all experiments and prepared all figures. Uri Ben-David assisted with e-karyotyping. Jatin Roper, Jonathan Braverman, and Shinya Imada aided with mouse experiments. Ömer Yilmaz, George Eng, James Stone, Dmitriy Kedrin, Shawn Demehri, Erik Schipperke, Leif Ellisen, and Mihriban Karaayvaz aided in the collection of human samples. I wrote the manuscript with input from Angelika Amon.

Chapter Six

I wrote Chapter Six.

Chapter One: Introduction

Introduction

Aneuploidy, a state of genome imbalance, is a nearly universal yet mysterious feature of cancer. During tumorigenesis, cells accumulate genetic changes which increase their proliferative capacity and survival, transforming them into cancer cells (Balmain et al. 2003; Vogelstein and Kinzler 2004). While aneuploidy is exceedingly rare in healthy human tissues, it is one of the most common features of cancer cells. For context, the most frequently altered gene in cancer, p53, is inactivated in about 50% of tumors (Vogelstein et al. 2000), but ~90% of tumors are aneuploid (Knouse et al. 2017). Beyond its ubiquity, aneuploidy is a striking feature of cancer cells because of its dramatic disruption of cell physiology; while many of the genetic changes that occur during tumorigenesis only impact the activity of a single gene, genome imbalances can affect thousands of genes at once.

But cataloging genetic changes is insufficient to understand cancer. Our current view of tumorigenesis includes the acquisition of epigenetic changes, alterations to the tumor microenvironment, metabolic reprogramming, and evasion of the immune system (Hanahan and Weinberg 2011). We view cancer not as monolith, but as a collection of diseases which vary by cell-of-origin, mutational spectrum, and patients' genetic background.

A simple view of aneuploidy is also insufficient. As I will discuss in this Introduction, in different contexts aneuploidy appears to both promote and suppress tumorigenesis, and its role appears to change throughout tumor evolution. First, I will define aneuploidy and describe the mechanisms which produce the aneuploid state. Second, I will detail how aneuploidy impacts cells' and organisms' fitness. Finally, I will relate what is currently known about the role of aneuploidy in tumorigenesis and tumor evolution.

Defining Aneuploidy

Aneuploidy is most precisely defined as a state in which a cell does not contain an exact multiple of its haploid chromosome complement. This definition draws an immediate distinction between any form of genome imbalance—which can range from small focal amplifications and deletions to whole chromosome gains and losses—and polyploidy. Though a polyploid cell may contain an aberrant number of chromosomes, these chromosomes compose an exact multiple of its haploid complement, so the cell's genome remains balanced. This distinction is important, as polyploid cells do not suffer from the fitness defects that characterize aneuploid cells (Storchova and Kuffer 2008). It is aneuploidy's imbalances that affect cell physiology, rather than the change in chromosome content.

While some definitions of aneuploidy might encompass all forms of genome imbalance, including sub-chromosomal amplifications and deletions, in this thesis I will consider sub-chromosomal amplifications and deletions to be a distinct form of genome imbalance (Fig. 1). As I will discuss, sub-chromosomal copy number variants (CNVs) arise through different mechanisms than whole chromosomal aneuploidy. Furthermore, sub-chromosomal CNVs and whole chromosomal aneuploidy are found with different prevalence in the human population and have different impacts on cellular and organismal physiology.

Defining aneuploidy by size

Since genome imbalance disrupts cellular function by altering the copy number of genes, the larger the genomic region affected by an imbalance, the greater the expected disruption to cell physiology. Support for this hypothesis comes from the prevalence of different types of aneuploidy in the human population. The only constitutional aneuploidies which survive to birth are trisomies of the three chromosomes which carry the fewest genes: chromosomes 13, 18, and 21. Of these, trisomy 21, the chromosome with the fewest genes, is both the most frequently

observed at 1 in 700 live births and carries the longest life expectancy (Hassold and Jacobs 1984).

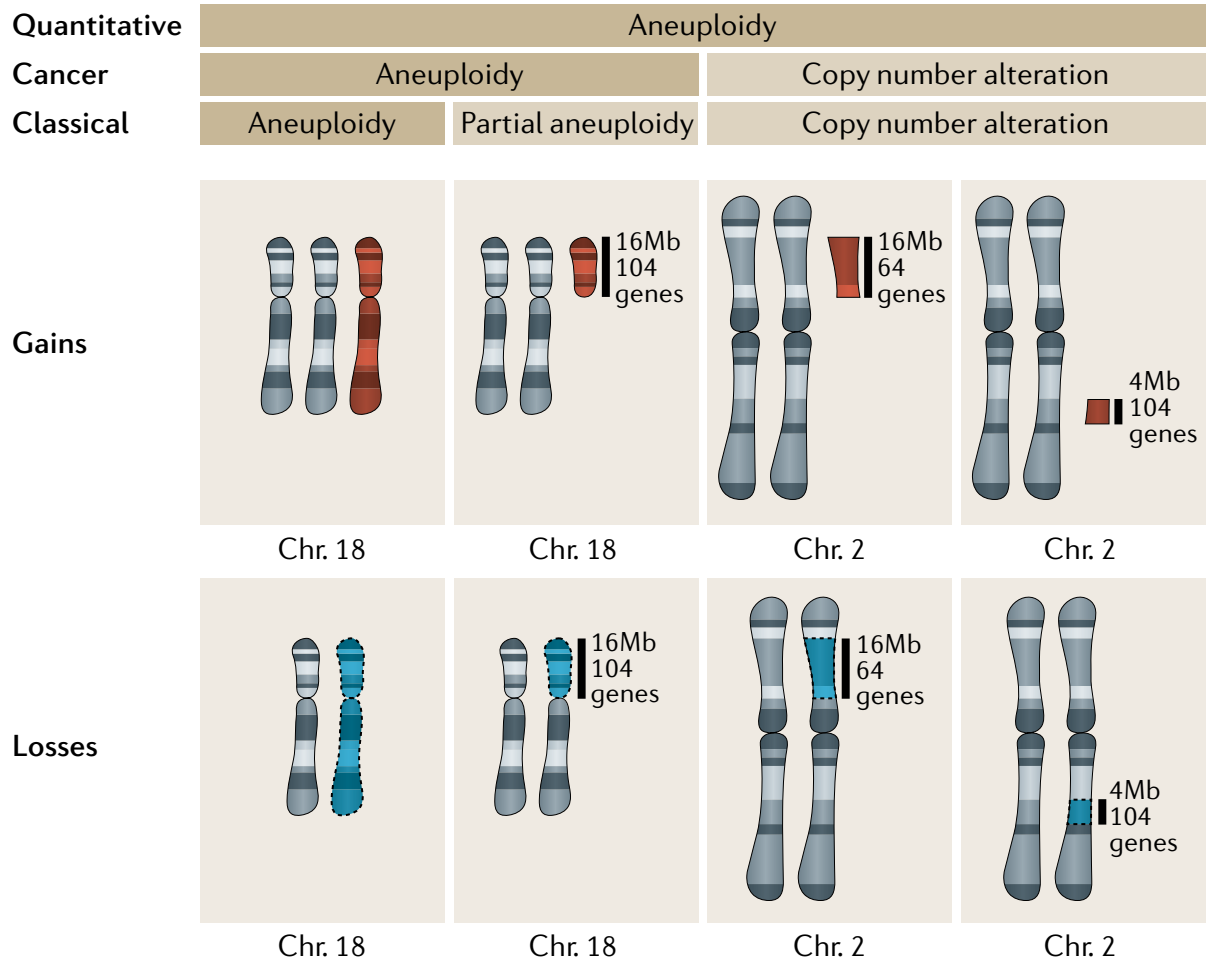


Figure 1: Differentiating between distinct types of genome imbalance. In this Introduction, I will follow the "classical" nomenclature. I will call gains and losses of whole chromosomes "aneuploidy," and I will refer to any sub-chromosomal copy number alterations as "sub-chromosomal copy number variants (CNVs)." The term sub-chromosomal CNVs will include "partial aneuploidies." Entire figure adapted from Ben-David & Amon (2020).

Sub-chromosomal CNVs are more common in humans, and also decrease in frequency with size. Sub-chromosomal CNVs can range in size from a few kilobases (kb), which can affect only one or a small number of genes, to over a hundred megabases (Mb), which can affect the expression of hundreds of genes (Henrichsen et al. 2009). While constitutional CNVs greater than 100 kb are present in 65% of individuals, constitutional CNVs exceeding 1 Mb are only

found in 1% of humans (Itsara et al. 2010). Moreover, large constitutional CNVs are associated with disease and face negative selection in the human population (Itsara et al. 2010; Girirajan et al. 2011).

Analyses of blood and skin have revealed that somatic Mb scale CNVs are present in 4% of individuals. Because these analyses were performed using bulk sequencing, these somatic CNVs must represent clones accounting for greater than 5% of the sampled tissue to have been detected (Forsberg et al. 2017; Jacobs et al. 2012; Laurie et al. 2012). Single-cell sequencing identified non-clonal, megabase-scale CNVs in 8-9% of healthy human brain and skin cells. Notably, over 85% of the CNVs reported in this study were smaller than 10 Mb in size (Knouse et al. 2016). For comparison, single-cell analysis revealed that healthy human brain, skin, and liver on average feature chromosomal aneuploidy in ~2% of cells with base tetraploid liver cells accounting for over half of the aneuploid cells (Knouse et al. 2014).

Together, these findings suggest that there are sub-types of genome imbalance with vastly different outcomes in humans, which can be stratified by size. Constitutive CNVs less than 1 Mb and small somatic CNVs could be considered common genome imbalances, which occurs in much of the human population and is seemingly unlinked to disease. Whole-chromosomal aneuploidy and large CNVs could be deemed pathologic genome imbalances, as these imbalances are associated with death and disease in humans.

Defining aneuploidy by mechanistic origin

Changes in the copy numbers of whole chromosomes are the result of chromosome mis-segregation, the mechanism behind chromosomal instability (CIN). If aneuploidy is a state, then CIN is the rate with which aneuploidy is generated. It is estimated that normal diploid cells mis-segregate a chromosome one out of every hundred divisions *in vivo* (Thompson and Compton 2008; Cimini et al. 2001). Chromosome segregation fidelity is maintained by two pathways. The

error correction pathway seeks to ensure that sister chromatids are bound at their kinetochores by microtubules emanating from opposite poles and the spindle assembly checkpoint (SAC) suspends cell division until all kinetochores have formed stable, amphipathic microtubule attachments (Cheeseman et al. 2002; DeLuca et al. 2006; Liu et al. 2009a).

The most common error in chromosome segregation is merotelic attachment, when a single kinetochore is bound by microtubules anchored at both poles (Fig. 2). Because merotely creates sufficient tension at the kinetochore to satisfy the SAC, it does not trigger a pause in cell division (Cimini et al. 2001). Some meroteley can be corrected by the error correction pathway (Cimini et al. 2003, 2006), but it can also result in the affected chromosome lagging behind during anaphase. This lagging chromosome can then be included in the wrong daughter cell, creating a mis-segregation event (Cimini et al. 2004). The lagging chromosome can also experience significant DNA damage if pulverized by the cleavage furrow, inducing double strand breaks. These breaks ultimately result in unbalanced translocations in the daughter cell, generating structural aberrations (Crasta et al. 2012; Janssen et al. 2011). Or, the lagging chromosome can become trapped in a micronucleus. Micronuclei are prone to membrane collapse, causing massive DNA damage (Hatch et al. 2013). While lagging chromosomes and micronuclei are rare in tissues, where tissue architecture enhances the correction of merotelic attachments (Knouse et al. 2018), lagging mitoses are frequently seen in cancer cell lines (Gordon et al. 2012).

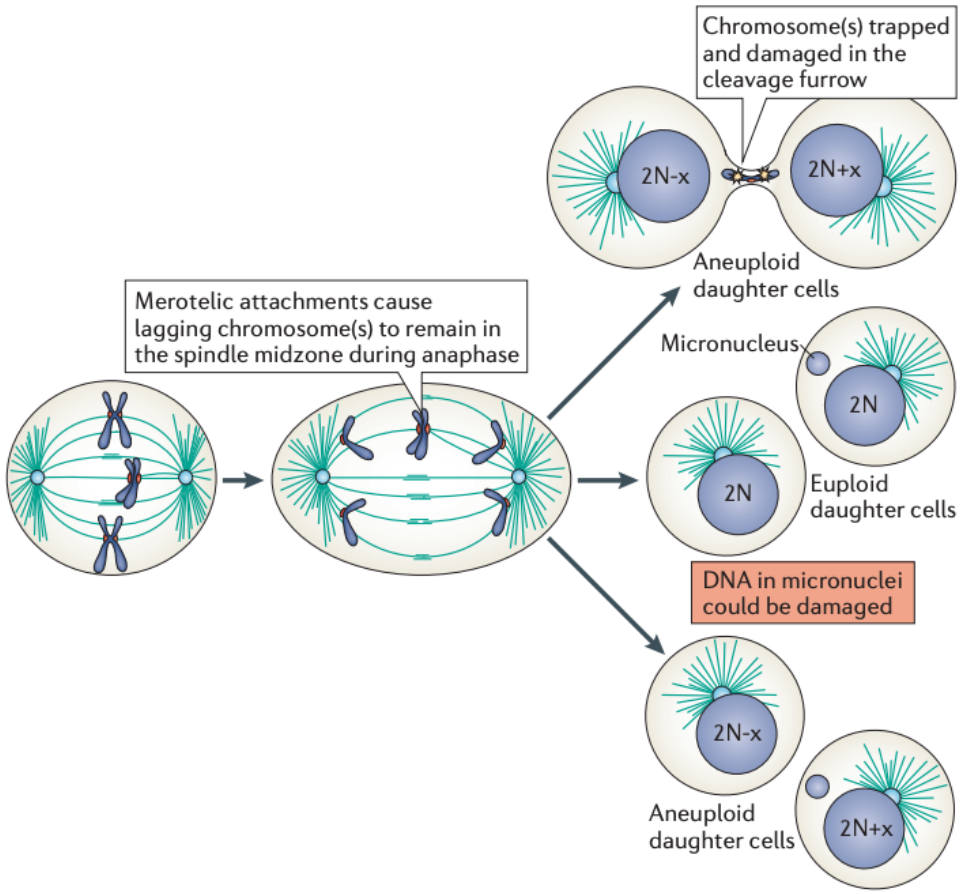


Figure 2: Merotely is the most common error in chromosome segregation. Because the inappropriate attachments create tension at the kinetochore, merotely does not cause the spindle assembly checkpoint to fire, allowing chromosome mis-segregation. Entire figure adapted from Santaguida & Amon (2015).

Beyond merotely, cells can become aneuploid through mutations in the SAC, the accumulation of supernumerary centrosomes, and through defects in chromosome cohesion. In cancer cells, mutations in the SAC are rare (Cahill et al. 1999; Haruki et al. 2001; Myrie et al. 2000; Wang et al. 2004), and most cancer cell lines with high CIN have functioning SACs. The presence of supernumerary centrosomes correlates strongly with increased CIN (Nigg 2002, 2006; Pihan et al. 1998, 2003) and is thought to affect segregation fidelity by increasing merotely (Ganem et al. 2009; Silkworth et al. 2009) rather than the creation of multipolar spindles, as was previously believed (Brinkley 2001; Ring et al. 1982; Lengauer et al. 1997). Inactivating

members of the cohesin complex can also cause chromosome segregation defects (Solomon et al. 2011).

CNVs arise through different mechanisms which occur throughout the cell cycle. These mechanisms include erroneous recombination (Sharp et al. 2005; Redon et al. 2006; Conrad et al. 2010) and failures during non-homologous end joining (Conrad et al. 2010). Errors during DNA replication can also produce CNVs through fork stalling and DNA breaks (Lee et al. 2007; Hastings et al. 2009). Chromothripsis can also generate sub-chromosomal imbalances as chromosomes are fragmented then reassembled incorrectly, such that the resulting chromosome contains copy number aberrations (He et al. 2019; Soto et al. 2018; Zhang et al. 2015). Because whole chromosomal aneuploidy and sub-chromosomal CNVs are produced by such different mechanisms and have different outcomes, it is important to distinguish between the two, even though they are both forms of genome imbalance.

Consequences of aneuploidy in non-transformed cells

Our understanding of the cellular consequences of genome imbalance comes almost entirely from studies of whole chromosomal aneuploidy as mis-segregating entire chromosomes routinely produces robust phenotypes. In most cases, gaining chromosomes uniformly increases mRNA production and protein levels across the amplified chromosome (Torres et al. 2007; Williams et al. 2008). Though there is some evidence that gene dosage is tempered in *Drosophila* and plants (Birchler et al. 2001; Birchler 2010), aneuploid yeast (Torres et al. 2007), mouse (Williams et al. 2008), and human (Pavelka et al. 2010) cells all show a corresponding change in gene expression following mis-segregation. Aneuploid cells also show altered global gene expression patterns (Upender et al. 2004).

This massive deregulation of gene expression results in both gene-specific and general effects. For example, increased expression of the amyloid precursor protein (APP) on chromosome 21 has long been implicated in the increased risk of early onset Alzheimer's disease in individuals with Down syndrome (McCarron et al. 2014). Clinical evidence has shown that individuals with partial chromosome 21 amplification that lack APP triplication do not experience neurodegeneration (Doran et al. 2017). However, recent work in mice demonstrated that triplication of most of the genes on human chromosome 21 excepting APP is also sufficient to induce amyloid-ß aggregation and other features of Alzheimer's disease, suggesting there is an additional role of mass deregulation in this phenomenon (Wiseman et al. 2018). Similarly, in aneuploid yeast, deleting copies of individual genes on gained chromosomes alleviates some of the growth defects associated with aneuploidy but is insufficient to alleviate the entire effect (Dodgson et al. 2016).

Aneuploid cells show many common phenotypes, regardless of which chromosomes are imbalanced (Fig. 3). Aneuploidy slows proliferation (Williams et al. 2008; Thorburn et al. 2013), alters metabolism (Torres et al. 2007), activates stress pathways (Sheltzer et al. 2012), causes defects in autophagy (Santaguida et al. 2015), and interferes with protein homeostasis (Tang et al. 2011; Torres et al. 2007, 2010; Oromendia et al. 2012).

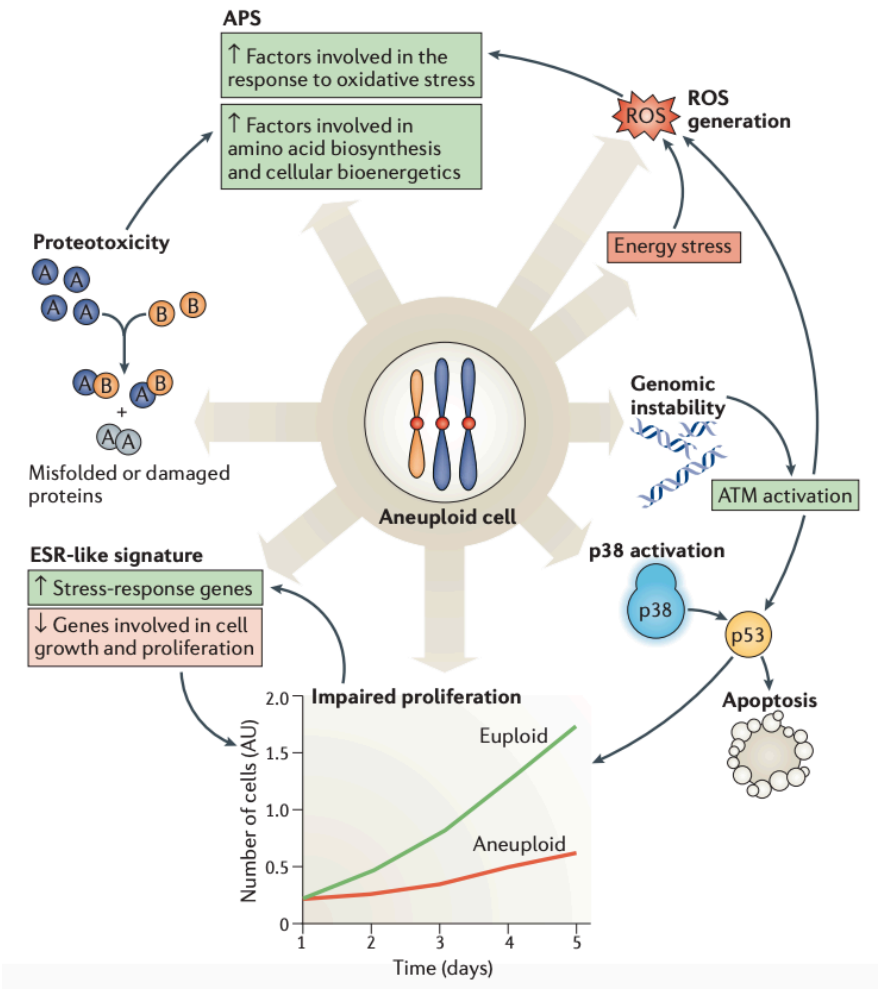


Figure 2: Aneuploidy induces common phenotypes. Entire figure adapted from Santaguida & Amon (2015).

Aneuploidy is not only the product of genome instability, but also causes it (Burrell et al. 2013; Lamm et al. 2018; Sheltzer et al. 2011). Aneuploid yeast are more likely to experience mis-segregation events and have defects in DNA repair (Sheltzer et al. 2011). Aneuploid human cells experience increased replication stress, and increased rates of abnormal mitoses (Santaguida et al. 2017; Passerini et al. 2016).

Aneuploidy decreases fitness

These cellular defects make aneuploid cells less fit. In growth competitions, aneuploid cells are routinely less proliferative than euploid cells (Williams et al. 2008; Thorburn et al.

2013). Aneuploid cells are also vulnerable to DNA damaging agents (Sheltzer et al. 2011) and inhibition of autophagy and protein aggregation (Tang et al. 2011).

Aneuploid organisms are also less fit. Aneuploidy is the leading cause of miscarriage in humans (Jia et al. 2015) and all constitutive trisomies are embryonic lethal in mice (Williams et al. 2008). However, not all aneuploid organisms are equally unfit. Trisomic mouse embryos show increased rates of resorption, but not all embryos are resorbed (Beach et al. 2017). Similarly, many trisomy 21 human embryos do not survive to birth and those individuals who are born experience a wide spectrum of physiological changes (Roper and Reeves 2006; Jia et al. 2015). The link between non-genetic variability and aneuploidy raises the possibility that a population of genetically identical aneuploid cells may contain some cells with greater fitness, which may have increased potential to undergo cancerous transformation. Future sections of this introduction will address the tumorigenic capacity of aneuploid cells.

Consequences of CNVs in non-transformed cells

Our current understanding of the impact of CNVs on cells emphasizes gene-specific effects. While large (>10 Mb) CNVs might induce some of the conserved effects of aneuploidy on cells, evidence for these phenotypes is lacking. Recent sequencing efforts have linked CNVs in the human population to disease susceptibility. For example, amplifications that include the gene *PRSS1* on chromosome 7 increase the risk of pancreatitis (Almal and Padh 2012). Conversely, there are examples of CNVs which are linked to complex phenotypes such as learning difficulties, where a direct mechanistic relationship with one or more amplified genes is not yet understood, leaving room for the possibility of mass effects affecting cell physiology (Iourov et al. 2015).

Aneuploidy and Tumorigenesis

This paradox—that a state that induces cellular stress, slows proliferation, and decreases cells' fitness could characterize a disease of hyper-proliferation and survival—remains one of the great mysteries in cancer biology. It has been over 100 years since Theodor and Marcella Boveri proposed that chromosome mis-segregation is the force behind cancerous transformation (Boveri 2008), yet aneuploidy's role in tumorigenesis remains to be completely understood. Evidence from aneuploid humans, experimental models, and analyses of cancer genomes suggests that aneuploidy can be both pro- and anti-tumorigenesis in a karyotype and tissue-dependent manner.

Tumor prevalence in aneuploid humans

In order to address these questions, researchers have evaluated the prevalence of cancer in aneuploid humans. While trisomy 21 individuals face a significant increase in their risk for developing leukemia, particularly acute lymphoblastic leukemia (ALL) and acute megakaryocytic leukemia (AMKL), they have a lower lifetime risk of developing solid tumors (Satgé et al. 1998). Only 10% of infants with trisomy 18 (Edwards syndrome) survive their first year of life (Lin et al. 2006). However, within this small pool of individuals, clinicians have observed an increase in their incidence of Wilms tumor and hepatocarcinoma (Ganmore et al. 2009). The short life expectancies for individuals with trisomy 18 do not allow for analysis of non-childhood cancers. Mosaic variegated aneuploidy (MVA) is an inherited disease in humans caused by loss of function of SAC components or centrosome constituents, increasing CIN (Snape et al. 2011; Hanks et al. 2004). Exceedingly rare in the human population, MVA is associated with increased risk of rhabdomyosarcoma and Wilms tumors during childhood (Plaja et al., 2001).

These data reveal both chromosome-specific and general impacts of aneuploidy on cancer. Chromosome 21 is frequently gained in non-Down Syndrome cases of AMKL (Hama et

al. 2008) and chromosome 18 is frequently gained in non-Edwards syndrome Wilms tumors (Betts et al. 1997). This indicates that chromosome specific effects may contribute to this increased cancer risk. The relationship between karyotype selection and tissue-specificity will continue to emerge as a theme in the role of aneuploidy in tumorigenesis. Here, there appears to be a relationship between each of these conditions and certain tumor types.

Individuals with Down Syndrome are less likely to develop cancer as adults, but are prone to childhood leukemia, which initiates during embryonic development (Lechman et al. 2014). This suggests that aneuploidy may have a different relationship with tumorigenesis at different stages of human development. However, since our evidence for aneuploidy suppressing tumorigenesis in adults is specific to one trisomy (which does contain known tumor suppressor genes), it is difficult to draw conclusions. To this point, it is known that many of the phenotypes of Down Syndrome cannot be attributed to the copy numbers of specific genes on chromosome 21 (Korbel et al. 2009), suggesting that the apparent tumor suppression of trisomy 21 could be another consequence of the aneuploid state itself.

Evaluating the tumorigenic capacity of aneuploid cells

Several studies have evaluated whether aneuploidy contributes to tumorigenesis by creating homogeneous cell populations which feature discreet gains or losses of chromosomes. Trisomic mouse embryonic fibroblasts (MEFs) transduced with a variety of oncogenes are no more tumorigenic than their euploid counterparts both *in vitro* and in a xenograft model (Sheltzer et al. 2017). A study of tetraploid, immortalized MEFs revealed that chromosome losses lead to decreased proliferative capacity *in vitro*, but increase growth in a xenograft model (Thomas et al. 2018). This finding is mirrored by observations in trisomic human cells (Rutledge et al. 2016) and yeast (Pavelka et al. 2010; Yona et al. 2012), where aneuploid cells have impaired proliferation under normal growth conditions, but thrive under stress conditions.

But studies of homogeneous aneuploid cell lines run counter to the nature of aneuploid cells, which is to be genomically unstable and heterogeneous. Both studies which examined the growth of aneuploid MEFs reported that the karyotypes of their cell lines evolved over the course of their experiments, producing novel karyotypes which were more proliferative (Sheltzer et al. 2017; Thomas et al. 2018). The idea that certain karyotypes might be pro-tumorigenic and others anti-tumorigenic is compatible with the observation that certain gains and losses co-occur both in the laboratory (Mayshar et al. 2010; Anders et al. 2009; Ravichandran et al. 2018) and in tumors (Taylor et al. 2018).

There is also evidence for aneuploid cells evolving *de novo* from euploid cells in culture. Trisomy 8 cells can spontaneously arise and accumulate within cultures of mouse embryonic stem cells, where aneuploidy appears to provide a selective advantage (Liu et al. 1997; Ben-David and Benvenisty 2012). Trisomy 12 commonly arises and spreads in cultured human cells where it is associated with increased proliferation and tumorigenicity (Ben-David et al. 2014). These findings suggest that certain karyotypes growing under the right conditions can have tumorigenic properties. But can aneuploid cells accumulate *in vivo*?

To explore whether aneuploidy can contribute to tumorigenesis *in vivo*, many labs have generated mice with genetic perturbations that increase CIN. Although these models diverge significantly from the process of tumorigenesis in healthy humans tissues, which have low CIN, they have shown that CIN can drive tumorigenesis (Foijer et al. 2014; Levine et al. 2017; Wijshake et al. 2012; Ricke et al. 2011; Li et al. 2009; Sotillo et al. 2010; Babu et al. 2003; Michel et al. 2001; Van Ree et al. 2010).

But this is not always true. Other mouse models have revealed that CIN can suppress tumorigenesis and alter the spectrum of tumor incidence. For example, although inhibition of the

SAC through depletion of CENP-E induces spontaneous lymphomas and lung adenomas, these animals are resistant to tumor induction by treatment with a carcinogen (Weaver et al. 2007). Moreover, mice featuring extremely high levels of chromosome mis-segregation, achieved by interfering with multiple components of the SAC, have decreased rates of tumor formation (Silk et al. 2013).

CIN and aneuploidy may have different effects on tumorigenesis. Mouse models of Down Syndrome, which are trisomic for much of the mouse equivalent of human chromosome 21, show decreased tumorigenesis (Baek et al. 2009; Reynolds et al. 2010; Yang and Reeves 2011). Whether this is a consequence of constitutive aneuploidy in the absence of high CIN, or a feature of trisomy 21, it furthers the observation that not all aneuploid mice are prone to tumor formation.

Lessons from tumor karyotypes

Another approach to uncovering the role of aneuploidy in tumorigenesis is analyzing the karyotypes of primary tumors in order to take a retrospective view of how aneuploidy might have contributed to tumorigenesis. These studies have revealed conserved patterns across cancer sub-types and the many ways that context shapes cancer genomes (Fig. 3).

The probability that a chromosome arm is gained or lost correlates with the number of coding regions on that arm (Duijf et al. 2013). This is in line with our understanding of aneuploidy as anti-tumorigenic, as karyotypes which represent milder copy number imbalances are enriched in cancer. However, in a typical cancer cell 25% of the genome has an altered copy number (Beroukhim et al. 2010; Zack et al. 2013), so any selection towards "milder" copy number variations does not prevent mass genome imbalance. For context, recall that 90% of humans harboring constitutional trisomy 18, which imbalances only 2.5% of the genome, do not survive infancy. One popular hypothesis for how aneuploidy contributes to tumorigenesis is by

altering the copy numbers of key genes. As this hypothesis predicts, chromosomes that are gained tend to be enriched for genes that enhance proliferation while chromosomes that are lost tend to contain genes which repress proliferation (Davoli et al. 2013).

The distribution of CNVs across cancer genomes also provides compelling evidence for this key genes hypothesis. In one pan-cancer analysis, pro- and anti-apoptotic BCL-2 gene family members were found to be over-represented in regions of the genome affected by focal amplifications and deletions, respectively. This indicates that CNVs could provide a mechanism for beneficially enriching pro-proliferation genes and eliminating copies of anti-proliferation genes. In this study, 33% of focal amplifications included known oncogenes and 11% of focal deletions contained known tumor suppressors (Beroukhim et al. 2010). Not only does this provide evidence that CNVs can shape cancer genomes to benefit tumors, but it suggests that recurrent CNVs may provide a method for detecting novel oncogenes and tumor suppressors. Alternatively, CNVs lacking for key genes may indicate that these genetic lesions are simply a consequence of instability and a passenger of tumorigenesis.

It is possible that CNVs and aneuploidy serve different roles and evolve differently over the course of tumor evolution. Not only do they arise by different mechanisms and carry different fitness consequences, but they have different effects on gene expression. Whole-chromosome gains or losses in diploid cells result in only a modest increase to 1.5 fold expression of affected genes, but focal amplifications can increase gene expression over a dozen fold (Nord et al. 2010).

While genome imbalances may alter the expression of key genes across the genome, mutations shape tumors' karyotypes. Evidence from genetically engineered mouse models of breast cancer indicates that different genetic drivers of tumorigenesis induce different karyotypes

(Ben-David et al. 2016). This observation also holds true in humans where distinct genetic subtypes of breast cancer have characteristic karyotypes (Gatza et al. 2014).

Though there are conserved elements of cancer karyotypes, there is no single karyotype or copy number change which is universally beneficial (Sack et al. 2018; Carter et al. 2012; Hoadley et al. 2018). Rather, karyotype selection appears to depend on tissue of origin and other disease-specific factors (Knouse et al. 2017). Evidence from mouse models has shown that CIN allows for the selection of different tumor karyotypes in different tissues (Foijer et al. 2017). In a recent meta-analysis, karyotypes were conserved amongst members of organ-specific groups—kidney cancers were more similar to other kidney cancers and gynecological cancers were more similar to gynecological cancers (Hoadley et al. 2018). However, squamous cell cancers are more similar to each other than to epithelial cancers arising from their same organ (Hoadley et al. 2018; Taylor et al. 2018).

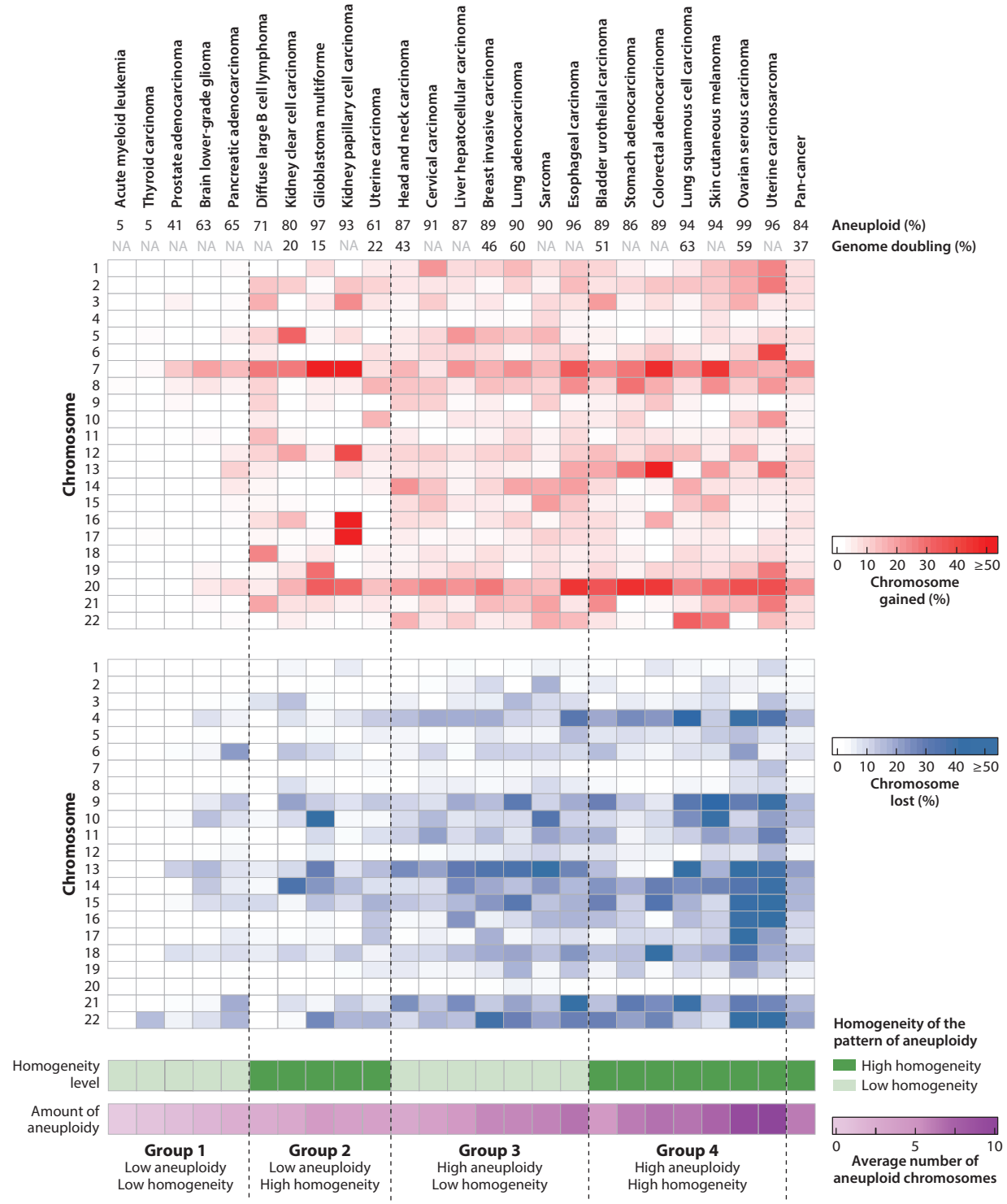


Figure 3: Comparing tumor karyotypes across cancers. Although some themes are consistent across cancers, the optimal karyotype for a given tumors depends on many factors—from tissue-of-origin to mutational drivers to individual genetic background. Entire figure adapted from Knouse et al. (2017).

Analyses of tumor genomes have proposed that whole genome duplication (WGD) can allow cancer cells to tolerate aneuploidy. Almost a third of tumors show evidence of genome doubling and WGD is associated with increased aneuploidy. WGD increases rates of chromosome loss, as additional copies of the genome can buffer against the deleterious effects of losing chromosomes (Bielski et al. 2018; Zack et al. 2013; Carter et al. 2012). It has been shown in both mouse and human cells that genome duplication can allow cells to tolerate aneuploidy and promote tumorigenesis (Thomas et al. 2018; Davoli and de Lange 2012).

Karyotype evolution during tumorigenesis

Aneuploid primary tumors arise from euploid healthy tissues, raising the question not just of why aneuploid cells dominate tumor populations, but when this transition occurs. If aneuploidy is an early event in tumorigenesis, this suggests that aneuploidy can contribute to this process—possibly by altering the copy numbers of key genes. We would then conclude that aneuploidy is pro-tumorigenic, even if it can be anti-tumorigenic in many contexts. If aneuploidy is a late event, then aneuploidy could represent a barrier to be overcome during tumorigenesis. Perhaps the accumulation of aneuploid cells requires that cells first develop aneuploidy tolerating mutations (Torres et al. 2010) or the right combination of driver mutations.

Although analyses of the karyotypes of primary tumors have offered snapshots of the selective pressures which can shape cancer genomes, most fail to elucidate how these karyotypes evolved. Because WGD affects the relative quantification of copy numbers throughout the genome, it is possible to determine which imbalances arose before and after WGD. One analysis described that genome doubling is a relatively early event, with most genome imbalances arising after duplication. The timing of duplication was consistent across tumors (Zack et al. 2013). This

is compatible with our understanding of genome doubling as a mechanism for aneuploidy tolerance and with the hypothesis aneuploidy is anti-tumorigenic early in tumor evolution.

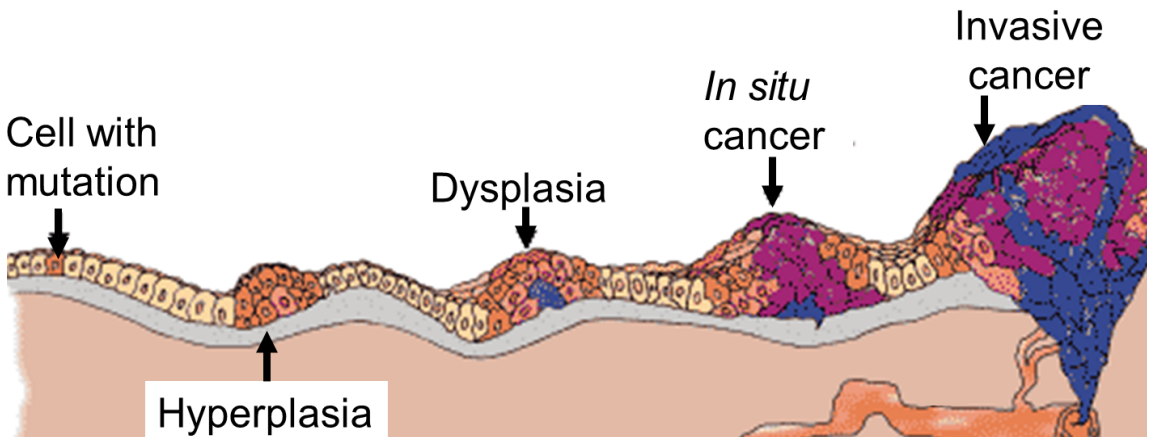


Figure 4: tumorigenesis progresses through successive pre-malignant stages. Figure adapted from the National Cancer Institute "Understanding Cancer" (science.education.nih.gov).

To assign the emergence of aneuploidy to one or more of the stages of tumorigenesis, researchers have analyzed samples of pre-malignant lesions. Studies linking tumor stage to aneuploidy have historically focused on the evolution of colorectal carcinoma, which progresses through well-defined stages (Fig. 4). Cumulatively, these studies concluded that aneuploidy arises early in tumorigenesis, as aneuploidy is present in early adenomas and increases progressively in late adenomas and carcinomas (Vogelstein et al. 1988, 1989; Ried et al. 1996; Meijer et al. 1998). A contemporary of these studies analyzed carcinoma of the uterine cervix and found that gain of chromosome 3p characterizes the transition from the latest pre-malignant stage to invasion and malignancy (Heselmeyer et al. 1996).

In more recent studies, genome sequencing and gene expression profiling have replaced the traditional—and less accurate—methods of fluorescence *in situ* hybridization and measuring DNA content by flow cytometry. Examining the last pre-malignant stages of lung cancer

(Teixeira et al. 2019) and breast cancer (Casasent et al. 2018), recent studies have argued that the critical transition to aneuploidy occurs at or before the carcinoma *in situ* stage. Retrospective studies which have ordered the evolution of the genetic components of breast cancer have supported this finding, arguing that aneuploidy arises in a brief, punctuated burst rather than gradually over the course of tumorigenesis (Gao et al. 2016b; Wang et al. 2014). This would place the accumulation of aneuploid cells relatively late in tumorigenesis, at the very end of the pre-malignant period, just before malignant invasion. These data are supported by studies of mouse models of cancer, which also concluded that aneuploidy arises during late tumorigenesis (Westcott et al. 2015; Nassar et al. 2015; Ben-David et al. 2016). However, we lack information about the karyotypes of earlier stages of pre-malignancy, before the cancer *in situ* stage, for most cancers.

Pre-cancerous lesions can spend decades slowly growing and transforming before achieving malignancy, and not all pre-cancerous lesions transform into cancer. Perhaps the best documentation of karyotype evolution during this transition has come from the study of Barrett's esophagus, a precursor to esophageal carcinoma. Not only does the accumulation of aneuploid cells accompany the transition from Barrett's esophagus to esophageal carcinoma (Ross-Innes et al. 2015), but higher levels of aneuploidy provide a biomarker for which Barrett patients will develop cancer (Bird-Lieberman et al. 2012; Martinez et al. 2018; Killcoyne et al. 2020).

There are some studies which have suggested that aneuploidy may arise earlier in tumorigenesis. Glioblastoma is characterized by the gain of chromosome 7 and this imbalance appears to be one of the earliest events in the generation of these cancers (Gerstung et al. 2020). Similarly, loss of chromosome 3p appears to be an initiating step in the generation of clear cell renal cancer, possibly arising 30-50 years before cancer diagnosis (Mitchell et al. 2018). Given

that these cancers feature highly recurrent, disease-specific karyotypes, it is possible that they follow a different evolutionary path than more heterogeneous diseases.

Despite studies which have shown that aneuploid cells accumulate by the last pre-malignant stage of tumorigenesis—at least, for some cancers—our understanding of this transition remains incomplete. These studies, by and large, have conducted bulk analyses and have neglected the earliest stages tumorigenesis. An examination of the dynamics of aneuploid cells growing at these stages, performed with single-cell precision, has yet to be performed and will be the subject of Chapter Five.

Aneuploidy and Tumor Evolution

So far this Introduction has focused on aneuploidy's relationship with tumorigenesis. But the formation of an early stage primary tumor is only an intermediate step in cancer. In the following sections, I will discuss the role of aneuploidy beyond the formation of a primary tumor, considering how aneuploidy may affect cancer prognosis and how aneuploidy may serve as a therapeutic target.

Aneuploidy and cancer prognosis

Beyond its ubiquity and dramatic cellular effects, aneuploidy is a notable feature of cancer because it impacts patients' survival. Many studies using a variety of methods to measure aneuploidy, studying either a small subset of patients with a single disease or performing metanalyses capturing thousands of tumors, have all concluded that aneuploidy is associated with worse prognosis (Auer et al. 1980; Steinbeck et al. 1994; Danielsen et al. 2016; Hieronymus et al. 2018; Vasudevan et al. 2020; Pinto et al. 2013; Gazic et al. 2008). However, this strong association may be confounded by tumor stage.

Since aneuploidy correlates with tumor stage (Laubert et al. 2015; Pinto et al. 2013; Gazic et al. 2008), it is possible that it is the higher tumor grade of aneuploid tumors which drives their deadliness, rather than the fact that they are aneuploid. However, there are three lines of evidence which refute the idea that tumor stage can account for all of aneuploidy's effect on prognosis. First, studies which have compared the prognosis of tumor samples at the same stage have also concluded that aneuploidy is associated with worse outcome (Kristensen et al. 2003; Macintyre et al. 2018). Tumor grade and aneuploidy can also be analyzed separately in prostate cancer, where aneuploidy is a better predictor of outcome than the Gleason score, a method of assigning tumor grade (Stopsack et al. 2019; Danielsen et al. 2016; Hieronymus et al. 2018). Second, evidence from CNVs provides gene-specific mechanisms for how genome imbalances can make tumors more aggressive, regardless of disease stage. Amplifications on chromosome 21 are linked to poor prognosis in acute myeloid leukemia, and are believed to primarily affect tumor growth by increasing the expression of the transcription factor *ERG* (Nibourel et al. 2017).

Third, not all karyotypes are equally lethal. Both myelodysplastic syndrome (MDS) and glioma can be stratified into high- and low-risk disease subtypes based on tumor karyotype. While MDS cases which lose one copy of chromosome 5 have a good prognosis, patients whose tumors feature monosomy 7 face a greater risk of dying (Kawankar and Vundinti 2011; Greenberg et al. 1997; Schanz et al. 2011). Glioma tumors monosomic for both chromosome 1p and 19q represent a unique sub-type of glioma (Idbaih et al. 2008; Brat et al. 2015; Wiestler et al. 2014) and carry a more favorable outcome than other sub-types, regardless of which treatment course patients receive (Weller et al. 2012; Wick et al. 2009; Cairncross et al. 2013; Van Den Bent et al. 2013). Together, these three lines of evidence suggest that aneuploidy can impact the prognosis of tumors—both negatively and positively.

As one would expect, CIN also correlates with poor prognosis. Colorectal carcinoma includes two distinct disease sub-types, each characterized by a particular form of genome instability: CIN and microsatellite instability (MIN). CIN tumors carry a worse prognosis than MIN tumors, and genomically stable, diploid tumors have the best prognosis of all (Hveem et al. 2014; Danielsen et al. 2016; Walther et al. 2008; Mouradov et al. 2013; Sinicrope et al. 2006). CIN has also been shown to increase lethality in gliomas (Nishizaki et al. 2002), glioblastomas (Crespo et al. 2015), osteosarcomas (Selvarajah et al. 2008), and acute myeloid leukemia (Heilig et al. 2010). In addition to very low levels of CIN correlating with better patient survival, tumors with very high levels of CIN have better outcomes and are more responsive to chemotherapy (Andor et al. 2016; Roylance et al. 2011; Birkbak et al. 2011).

Studies in animal models have arrived at divergent conclusions about the role of CIN in tumor evolution. Overexpression of the SAC component *MAD2* induces tumor growth in mice. However, unlike most oncogenic drivers, *MAD2* overexpression can be ceased without the *MAD2*-driven tumors collapsing into remission (Sotillo et al., 2007). In a mouse model of *KRAS*-driven lung cancer, *MAD2* overexpression causes tumor relapse after oncogene withdrawal forced the lung tumors into remission (Sotillo et al., 2010). These two studies suggest that CIN can maintain tumor growth and drive relapse. However, CIN induction by *MAD2* overexpression slowed tumorigenesis in a mouse model of breast cancer (Rowald et al. 2016). Furthermore, in a different mouse model of lung cancer, suppressing CIN increased the size and mitotic index of lung tumors. Decreasing CIN did not alter the incidence of tumors (Laucius et al. 2019), suggesting that in this model CIN is detrimental to tumors only in later stages of tumor growth.

These contradictory findings make sense if CIN contributes to tumor growth through karyotype selection. Because CIN produces random, heterogeneous karyotypes at some

frequency, it can accelerate adaptation by providing more opportunities for beneficial karyotypes to arise. Karyotype selection fuels adaptation in yeast confronted with drug treatments (Chen et al. 2015) and has been documented in CIN human cell lines (Wangsa et al. 2018; Ben-David et al. 2018). But, in the absence of a growth challenge or if CIN is too high, CIN depletes cancer cells from the growing tumor presumably by constantly creating novel, unfavorable karyotypes. A stochastic model of tumor growth weighed both the benefit of altering the copy number of key genes against the fitness penalty of gaining and losing chromosomes and the benefits of heterogeneity against the chaos of instability (Laughney et al. 2015). This computational model revealed that tumors must remain within a very narrow range of CIN to grow. These studies also suggest that different cancers will have different relationships with CIN. For diseases like MDS and glioma which are reliant on specific karyotypes, CIN could have a greater negative impact than in more karyotypically heterogeneous disease like breast cancer and lung cancer.

In light of the impact aneuploidy and CIN have on tumor evolution, two important questions are how genome imbalances affect tumors' responses to drug treatment and if genome imbalances play a role in metastasis. CIN can drive chemoresistance, but aneuploidy's relationship with drug treatment is complex. While highly aneuploid cancer cells are generally less sensitive to chemotherapies (Kuznetsova et al. 2015; Ben-David et al. 2017; Andor et al. 2016), not all karyotypes are chemoresistant. For example, loss of chromosomes 1p and 19p in glioma renders tumors more sensitive to chemotherapy (Cairncross et al. 2013; Van Den Bent et al. 2013; Weller et al. 2012; Wahl et al. 2017; Wick et al. 2009). Each possible karyotype, in each tumor, in each patient likely has its own suite of vulnerabilities and resistance mechanisms. While amplification of 14q32.33 is enriched in platinum-resistant ovarian cancer (Despierre et al. 2014), this mechanism will probably not prove universal to platinum-treated cancers from other

tissues because of the strong specificity between karyotype and tissue. Defining these karyotype-driven mechanisms remains a major challenge in understanding chemoresistance.

An emerging field of cancer research considers how aneuploidy's interactions with the immune system may enhance the growth and survival of aneuploid tumors. While aneuploidy correlates with immune evasion (Taylor et al. 2018; Buccitelli et al. 2017; Davoli et al. 2017), aneuploidy is also thought to attract the attention of the immune system. Highly aneuploid cells are targeted and destroyed by natural killer cells *in vitro* (Wang et al. 2020; Santaguida et al. 2017) and micronuclei—a hallmark of aneuploid cells—activate the cell intrinsic immune surveillance pathway via activation of the cGAS-STING pathway (Harding et al. 2017; MacKenzie et al. 2017). Further study is needed to elucidate the relationship between aneuploidy and the immune system, and how it changes over the course of tumorigenesis and tumor evolution.

Finally, it is not yet known whether aneuploidy promotes metastasis. There is some evidence that metastases can have distinct karyotypes, as monosomy for chromosome 9p is enriched in clear cell renal cancer metastases as compared to primary tumors of the same disease (Turajlic et al. 2018a). But, studies of pancreatic cancer and prostate cancer have reported that metastases' karyotypes do not diverge from their primary tumors (Makohon-Moore et al. 2017; Liu et al. 2009b).

There is some experimental evidence that aneuploidy may contribute to metastasis. Aneuploid cancer cell lines show behaviors characteristic of metastatic cells, such as increased mobility, adhesion, and migration (Roschke et al. 2008). Other studies have focused on how karyotypes evolve during the epithelial-to-mesenchymal transition (EMT). EMT can cause genome instability, leading to karyotype evolution (Comaills et al. 2016), and cells which

spontaneously undergo both EMT and the reverse process, reverting from the mesenchymal state to an epithelial one, experience karyotype evolution (Gao et al. 2016a).

In a xenograft model, CIN drove metastasis by activating the same cGAS-STING pathway which can trigger immune surveillance (Bakhoum et al. 2018). However, an analysis of 13 different congenic trisomic cell lines found that while aneuploidy does activate the cGAS-STING pathway it does not promote invasiveness (Vasudevan et al. 2020). Only one of the 13 trisomic cell lines in this study showed increased metastatic potential. These two studies suggest that—like so much of the relationship between aneuploidy and cancer—whether an aneuploid cell is more metastatic may depend on karyotype and cell type.

Aneuploidy as a therapeutic target

Having established the complex and often paradoxical nature of aneuploidy's relationship with cancer, I will close this Introduction with a brief discussion of aneuploidy's potential as a therapeutic target. Aneuploidy is an appealing target because it is common in tumors but rare in healthy tissues, is highly predictable in some cancers, and has substantial effects on cell physiology. Moreover, karyotype-specific treatments have been proven effective in the clinic. Lenalidomide is used in the treatment of MDS with loss of chromosome 5q and is the first treatment regimen used for a specific karyotype (Giagounidis 2012; List et al. 2018).

Tumors require continuous signaling from oncogenes to keep proliferating (Torti and Trusolino 2011). So the key genes hypothesis predicts that consistently gained chromosomes contain one or a few genes whose inhibition would halt tumor growth. Similarly, a consistently lost chromosome might contain repressors of a pathway which could be modulated to eliminate cancer cells. Finding these key genes, however, remains challenging. Metanalyses can identify the minimal region which is recurrently gained or lost to help identify the critical genes or

pathways to target (Beroukhim et al. 2010; Gatza et al. 2014; Curtis et al. 2012). But the benefit derived from gaining or losing a large section of the genome is most likely a net effect of many genes' changes in copy number (Davoli et al. 2013; Sack et al. 2018). Indeed, although p53 loss increases the proliferation of leukemias and lymphomas, when a 4 Mb region including p53 and its neighboring genes was deleted the tumors gained an even greater fitness benefit than simply deleting p53 alone (Liu et al. 2016).

Rather that isolating key genes, it might be beneficial to target the cumulative effects of altered gene expression. Because aneuploid cells share common cellular stresses, regardless of species and karyotype, these phenotypes may represent vulnerabilities which can be exploited for treatment. For example, due to the additional protein folding burden that follows gaining a chromosome, aneuploid cells are sensitive to compounds that interfere with protein folding and turnover (Tang et al., 2011; Torres et al., 2007). Even better, these vulnerabilities are additive. Simultaneously disrupting protein folding and inducing energy stress by treatment with 17-AAG and AICAR showed greater efficiency against aneuploid cells than either therapy alone (Tang et al., 2011).

Beyond inducing stress, aneuploidy increases populations' variability and induces genome instability, creating tumor heterogeneity. This tumor heterogeneity is believed to be a major driver of chemoresistance. But could aneuploidy's heterogeneity and adaptability be its Achilles heel? By studying how a heterogeneous population of yeast evolved their karyotypes in response to drug treatment, Chen et al. (2015) designed an evolutionary trap. Once they knew which karyotypes arose in response to treatment with one drug, the study designed a second treatment, tailored to eradicate cells selected to survive the first treatment.

In order to spring such a trap, a lot more must be learned about the role that aneuploidy

plays throughout tumor evolution. By creating a detailed history of how karyotypes evolve during tumorigenesis and tumor evolution, we can begin to understand what drives karyotype evolution and how this impacts tumor cells' survival. Eventually, we may develop the ability to predict how aneuploid cells will respond to treatment with respect to tissue-of-origin and patients' genetic background. Perhaps we can turn one of the great mysteries of tumor biology into an unrivaled tool for treating cancer.

Summary

Aneuploidy's relationship with cancer is paradoxical and context dependent. This thesis will describe my efforts to gain insight into what drives the accumulation of aneuploid cells during tumorigenesis. I have studied the dynamics of aneuploid cells in a variety of contexts, ranging from normal tissues to pre-malignant lesions to invasive tumors. Aneuploid cells were found to have fitness defects and be selected against in normal regenerating tissues *in vivo*. Modeling the dynamics of aneuploid cells in the hemopoietic system following bone marrow reconstitutions indicated that purifying selection determines the prevalence of aneuploid cells in regenerating tissues. Analyses of primary breast cells from women carrying *BRCA2* mutations have indicated that large CNVs are present in *BRCA2* carrier breast cells before the initiation of tumorigenesis.

To directly investigate the role of genome imbalances in tumorigenesis, I assessed the timing with which whole chromosomal aneuploidy and sub-chromosomal CNVs each arise during tumorigenesis. I have determined which, if any, genome imbalances are present throughout colorectal carcinogenesis and in cells from early-premalignant lesions from multiple tissues. I determined that large CNVs arise early in tumorigenesis, but whole chromosomal aneuploidy arises late in tumorigenesis.

This work proposes that while aneuploidy confers a fitness penalty in both normal and tumorigenic tissues and is selected against, CNVs carry a lesser fitness penalty. These findings also propose that the same purifying selection which purges aneuploid cells from normal, regenerating tissues, may be active during early tumorigenesis. Our work suggests that whole chromosomal aneuploidy has a limited role in early tumorigenesis and motivates future research into the role of aneuploidy in late tumorigenesis.

References

Almal SH, Padh H. 2012. Implications of gene copy-number variation in health and diseases. *J Hum Genet* **57**: 6–13.

Anders KR, Kudrna JR, Keller KE, Kinghorn BAA, Miller EM, Pauw D, Peck AT, Shellooe CE, Strong IJT. 2009. A strategy for constructing aneuploid yeast strains by transient nondisjunction of a target chromosome. *BMC Genet* **10**: 1–11.

Andor N, Graham TA, Jansen M, Xia LC, Aktipis CA, Petritsch C, Ji HP, Maley CC. 2016. Pan-cancer analysis of the extent and consequences of intratumor heterogeneity. *Nat Med* **22**: 105–113.

Auer GU, Caspersson TO, Wallgren AS. 1980. DNA content and survival in mammary carcinoma. *Anal Quant Cytol* **2**: 161–165.

Babu JR, Jeganathan KB, Baker DJ, Wu X, Kang-Decker N, Van Deursen JM. 2003. Rae1 is an essential mitotic checkpoint regulator that cooperates with Bub3 to prevent chromosome missegregation. *J Cell Biol* **160**: 341–353.

Baek KH, Zaslavsky A, Lynch RC, Britt C, Okada Y, Siarey RJ, Lensch MW, Park IH, Yoon SS, Minami T, et al. 2009. Down's syndrome suppression of tumour growth and the role of the calcineurin inhibitor DSCR1. *Nature* **459**: 1126–1130.

Bakhoum SF, Ngo B, Laughney AM, Cavallo JA, Murphy CJ, Ly P, Shah P, Sriram RK, Watkins TBK, Taunk NK, et al. 2018. Chromosomal instability drives metastasis through a cytosolic DNA response. *Nature* **553**: 467–472.

Balmain A, Gray J, Ponder B. 2003. The genetics and genomics of cancer. *Nat Genet* **33**: 238–244.

Beach RR, Ricci-Tam C, Brennan CM, Moomau CA, Hsu P hsin, Hua B, Silberman RE, Springer M, Amon A. 2017. Aneuploidy Causes Non-genetic Individuality. *Cell* **169**: 229-242.e21.

Ben-David U, Amon A. 2020. Context is everything: aneuploidy in cancer. *Nature Reviews Genetics* **21**:44-62.

Ben-David U, Arad G, Weissbein U, Mandefro B, Maimon A, Golan-Lev T, Narwani K, Clark AT, Andrews PW, Benvenisty N, et al. 2014. Aneuploidy induces profound changes in gene expression, proliferation and tumorigenicity of human pluripotent stem cells. *Nat Commun* **5**: 1–11.

Ben-David U, Benvenisty N. 2012. High prevalence of evolutionarily conserved and species-specific genomic aberrations in mouse pluripotent stem cells. *Stem Cells* **30**: 612–622.

Ben-David U, Ha G, Khadka P, Jin X, Wong B, Franke L, Golub TR. 2016a. The landscape of chromosomal aberrations in breast cancer mouse models reveals driver-specific routes to tumorigenesis. *Nat Commun* **7**: 1–13.

Ben-David U, Ha G, Tseng YY, Greenwald NF, Oh C, Shih J, McFarland JM, Wong B, Boehm JS, Beroukhim R, et al. 2017. Patient-derived xenografts undergo mouse-specific tumor evolution. *Nat Genet* **49**: 1567–1575.

Ben-David U, Siranosian B, Ha G, Tang H, Oren Y, Hinohara K, Strathdee CA, Dempster J, Lyons NJ, Burns R, et al. 2018. Genetic and transcriptional evolution alters cancer cell line drug response. *Nature* **560**: 325–330.

Beroukhim R, Mermel CH, Porter D, Wei G, Raychaudhuri S, Donovan J, Barretina J, Boehm JS, Dobson J, Urashima M, et al. 2010. The landscape of somatic copy-number alteration across human cancers. *Nature* **463**: 899–905.

Betts DR, Koesters R, Plüss HJ, Niggli FK. 1997. Routine karyotyping in Wilms tumor. *Cancer Genet Cytogenet* **96**: 151–156.

Bielski CM, Zehir A, Penson A V., Donoghue MTA, Chatila W, Armenia J, Chang MT, Schram AM, Jonsson P, Bandlamudi C, et al. 2018. Genome doubling shapes the evolution and prognosis of advanced cancers. *Nat Genet* **50**: 1189–1195.

Birchler JA. 2010. Reflections on studies of gene expression in aneuploids. *Biochem J* **426**: 119–123.

Birchler JA, Bhadra U, Bhadra MP, Auger DL. 2001. Dosage-dependent gene regulation in multicellular eukaryotes: Implications for dosage compensation, aneuploid syndromes, and quantitative traits. *Dev Biol* **234**: 275–288.

Bird-Lieberman EL, Dunn JM, Coleman HG, Lao-Sirieix P, Oukrif D, Moore CE, Varghese S, Johnston BT, Arthur K, McManus DT, et al. 2012. Population-based study reveals new risk-stratification biomarker panel for Barrett's esophagus. *Gastroenterology* **143**: 927-935.e3.

Birkbak NJ, Eklund AC, Li Q, McClelland SE, Endesfelder D, Tan P, Tan IB, Richardson AL, Szallasi Z, Swanton C. 2011. Paradoxical relationship between chromosomal instability and survival outcome in cancer. *Cancer Res* **71**: 3447–3452.

Boveri T. 2008. Concerning the origin of malignant tumours by Theodor Boveri. Translated and annotated by Henry Harris. *J Cell Sci* **121**: 1–84.

Brat DJ, Verhaak RGW, Aldape KD, Yung WKA, Salama SR, Cooper LAD, Rheinbay E, Miller CR, Vitucci M, Morozova O, et al. 2015. Comprehensive, integrative genomic analysis of diffuse lower-grade gliomas. *N Engl J Med* **372**: 2481–2498.

Brinkley BR. 2001. Managing the centrosome numbers game: From chaos to stability in cancer cell division. *Trends Cell Biol* **11**: 18–21.

Buccitelli C, Salgueiro L, Rowald K, Sotillo R, Mardin BR, Korbel JO. 2017. Pan-cancer analysis distinguishes transcriptional changes of aneuploidy from proliferation. *Genome Res* **27**: 501–511.

Burrell RA, McGranahan N, Bartek J, Swanton C. 2013. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature* **501**: 338–345.

Cahill DP, Da Costa LT, Carson-Walter EB, Kinzler KW, Vogelstein B, Lengauer C. 1999. Characterization of MAD2B and other mitotic spindle checkpoint genes. *Genomics* **58**: 181–187.

Cairncross G, Wang M, Shaw E, Jenkins R, Brachman D, Buckner J, Fink K, Souhami L, Laperriere N, Curran W, et al. 2013. Phase III trial of chemoradiotherapy for anaplastic oligodendroglioma: Long-term results of RTOG 9402. *J Clin Oncol* **31**: 337–343.

Carter SL, Cibulskis K, Helman E, McKenna A, Shen H, Zack T, Laird PW, Onofrio RC, Winckler W, Weir BA, et al. 2012. Absolute quantification of somatic DNA alterations in human cancer. *Nat Biotechnol* **30**: 413–421.

Casasent AK, Schalck A, Gao R, Sei E, Long A, Pangburn W, Casasent T, Meric-Bernstam F, Edgerton ME, Navin NE. 2018. Multiclonal Invasion in Breast Tumors Identified by Topographic Single Cell Sequencing. *Cell* **172**: 205-217.e12.

Cheeseman IM, Anderson S, Jwa M, Green EM, Kang J seog, Yates JR, Chan CSM, Drubin DG, Barnes G. 2002. Phospho-regulation of kinetochore-microtubule attachments by the Aurora kinase Ipl1p. *Cell* **111**: 163–172.

Chen G, Mulla WA, Kucharavy A, Tsai HJ, Rubinstein B, Conkright J, McCroskey S, Bradford WD, Weems L, Haug JS, et al. 2015. Targeting the adaptability of heterogeneous aneuploids. *Cell* **160**: 771–784.

Cimini D, Cameron LA, Salmon ED. 2004. Anaphase spindle mechanics prevent mis-segregation of merotelically oriented chromosomes. *Curr Biol* **14**: 2149–2155.

Cimini D, Howell B, Maddox P, Khodjakov A, Degrassi F, Salmon ED. 2001. Merotelic kinetochore orientation is a major mechanism of aneuploidy in mitotic mammalian tissue cells. *J Cell Biol* **152**: 517–527.

Cimini D, Moree B, Canman JC, Salmon ED. 2003. Merotelic kinetochore orientation occurs frequently during early mitosis in mammalian tissue cells and error correction is achieved by two different mechanisms. *J Cell Sci* **116**: 4213–4225.

Cimini D, Wan X, Hirel CB, Salmon ED. 2006. Aurora Kinase Promotes Turnover of Kinetochore Microtubules to Reduce Chromosome Segregation Errors. *Curr Biol* **16**: 1711–1718.

Comaills V, Kabeche L, Morris R, Buisson R, Yu M, Madden MW, LiCausi JA, Boukhali M, Tajima K, Pan S, et al. 2016. Genomic Instability Is Induced by Persistent Proliferation of Cells Undergoing Epithelial-to-Mesenchymal Transition. *Cell Rep* **17**: 2632–2647.

Conrad DF, Pinto D, Redon R, Feuk L, Gokcumen O, Zhang Y, Aerts J, Andrews TD, Barnes C, Campbell P, et al. 2010. Origins and functional impact of copy number variation in the human genome. *Nature* **464**: 704–712.

Crasta K, Ganem NJ, Dagher R, Lantermann AB, Ivanova E V., Pan Y, Nezi L, Protopopov A, Chowdhury D, Pellman D. 2012. DNA breaks and chromosome pulverization from errors in mitosis. *Nature* **482**: 53–58.

Crespo I, Vital AL, Gonzalez-Tablas M, Patino MDC, Otero A, Lopes MC, De Oliveira C, Domingues P, Orfao A, Tabernero MD. 2015. Molecular and Genomic Alterations in Glioblastoma Multiforme. *Am J Pathol* **185**: 1820–1833.

Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, Speed D, Lynch AG, Samarajiwa S, Yuan Y, et al. 2012. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**: 346–352.

Danielsen HE, Pradhan M, Novelli M. 2016. Revisiting tumour aneuploidy-the place of ploidy assessment in the molecular era. *Nat Rev Clin Oncol* **13**: 291–304.

Davoli T, de Lange T. 2012. Telomere-Driven Tetraploidization Occurs in Human Cells Undergoing Crisis and Promotes Transformation of Mouse Cells. *Cancer Cell* **21**: 765–776.

Davoli T, Uno H, Wooten EC, Elledge SJ. 2017. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science (80-)* **355**.

Davoli T, Xu AW, Mengwasser KE, Sack LM, Yoon JC, Park PJ, Elledge SJ. 2013. Cumulative Haploinsufficiency and Triplosensitivity Drive Aneuploidy Patterns and Shape the Cancer Genome. *Cell* **155**: 948–962.

DeLuca JG, Gall WE, Ciferri C, Cimini D, Musacchio A, Salmon ED. 2006. Kinetochore Microtubule Dynamics and Attachment Stability Are Regulated by Hec1. *Cell* **127**: 969–982.

Despierre E, Moisse M, Yesilyurt B, Sehouli J, Braicu I, Mahner S, Castillo-Tong DC, Zeillinger R, Lambrechts S, Leunen K, et al. 2014. Somatic copy number alterations predict response to platinum therapy in epithelial ovarian cancer. *Gynecol Oncol* **135**: 415–422.

Dodgson SE, Kim S, Costanzo M, Baryshnikova A, Morse DL, Kaiser CA, Boone C, Amon A. 2016. Chromosome-Specific and Global Effects of Aneuploidy in Saccharomyces cerevisiae. *Genetics* **202**: 1395–1409.

Doran E, Keator D, Head E, Phelan MJ, Kim R, Totoiu M, Barrio JR, Small GW, Potkin SG, Lott IT. 2017. Down Syndrome, Partial Trisomy 21, and Absence of Alzheimer's Disease: The Role of APP. *J Alzheimer's Dis* **56**: 459–470.

Duijf PHG, Schultz N, Benezra R. 2013. Cancer cells preferentially lose small chromosomes. *Int J Cancer* **132**: 2316–2326.

Foijer F, Albacker LA, Bakker B, Spierings DC, Yue Y, Xie SZ, Davis S, Lutum-Jehle A, Takemoto D, Hare B, et al. 2017. Deletion of the MAD2L1 spindle assembly checkpoint gene is tolerated in mouse models of acute T-cell lymphoma and hepatocellular carcinoma. *Elife* **6**.

Foijer F, Xie SZ, Simon JE, Bakker PL, Conte N, Davis SH, Kregel E, Jonkers J, Bradley A, Sorger PK. 2014. Chromosome instability induced by Mps1 and p53 mutation generates aggressive lymphomas exhibiting aneuploidy-induced stress. *Proc Natl Acad Sci U S A* **111**: 13427–13432.

Forsberg LA, Gisselsson D, Dumanski JP. 2017. Mosaicism in health and disease-clones picking up speed. *Nat Rev Genet* **18**: 128–142.

Ganem NJ, Godinho SA, Pellman D. 2009. A mechanism linking extra centrosomes to chromosomal instability. *Nature* **460**: 278–282.

Ganmore I, Smooha G, Izraeli S. 2009. Constitutional aneuploidy and cancer predisposition. *Hum Mol Genet* **18**: R84-93.

Gao C, Su Y, Koeman J, Haak E, Dykema K, Essenberg C, Hudson E, Petillo D, Khoo SK, Vande Woude GF. 2016a. Chromosome instability drives phenotypic switching to metastasis. *Proc Natl Acad Sci U S A* **113**: 14793–14798.

Gao R, Davis A, McDonald TO, Sei E, Shi X, Wang Y, Tsai PC, Casasent A, Waters J, Zhang H, et al. 2016b. Punctuated copy number evolution and clonal stasis in triple-negative breast cancer. *Nat Genet* **48**: 1119–1130.

Gatza ML, Silva GO, Parker JS, Fan C, Perou CM. 2014. An integrated genomics approach identifies drivers of proliferation in luminal-subtype human breast cancer. *Nat Genet* **46**: 1051–1059.

Gazic B, Pizem J, Bracko M, Cufer T, Borstnar S, Pohar-Marinsek Z, Us-Krasovec M. 2008. S-phase fraction determined on fine needle aspirates is an independent prognostic factor in breast cancer - A multivariate study of 770 patients. *Cytopathology* **19**: 294–302.

Gerstung M, Jolly C, Leshchiner I, Dentro SC, Gonzalez S, Rosebrock D, Mitchell TJ, Rubanova Y, Anur P, Yu K, et al. 2020. The evolutionary history of 2,658 cancers. *Nature* **578**: 122–128.

Giagounidis AAN. 2012. Lenalidomide for del(5q) and Non-del(5q) Myelodysplastic Syndromes. *Semin Hematol* **49**: 312–322.

Girirajan S, Campbell CD, Eichler EE. 2011. Human copy number variation and complex genetic disease. *Annu Rev Genet* **45**: 203–226.

Gordon DJ, Resio B, Pellman D. 2012. Causes and consequences of aneuploidy in cancer. *Nat Rev Genet* **13**: 189–203.

Greenberg P, Cox C, LeBeau MM, Fenaux P, Morel P, Sanz G, Sanz M, Vallespi T, Hamblin T, Oscier D, et al. 1997. International scoring system for evaluating prognosis in myelodysplastic syndromes. *Blood* **89**: 2079–2088.

Hama A, Yagasaki H, Takahashi Y, Nishio N, Muramatsu H, Yoshida N, Tanaka M, Hidaka H, Watanabe N, Yoshimi A, et al. 2008. Acute megakaryoblastic leukaemia (AMKL) in children: A comparison of AMKL with and without Down syndrome. *Br J Haematol* **140**: 552–561.

Hanahan D, Weinberg RA. 2011. Hallmarks of cancer: The next generation. *Cell* **144**: 646–674.

Hanks S, Coleman K, Reid S, Plaja A, Firth H, FitzPatrick D, Kidd A, Méhes K, Nash R, Robin N, et al. 2004. Constitutional aneuploidy and cancer predisposition caused by biallelic mutations in BUB1B. *Nat Genet* **36**: 1159–1161.

Harding SM, Benci JL, Irianto J, Discher DE, Minn AJ, Greenberg RA. 2017. Mitotic progression following DNA damage enables pattern recognition within micronuclei. *Nature* **548**: 466–470.

Haruki N, Saito H, Harano T, Nomoto S, Takahashi T, Osada H, Fujii Y, Takahashi T. 2001. Molecular analysis of the mitotic checkpoint genes BUB1, BUBR1 and BUB3 in human lung cancers. *Cancer Lett* **162**: 201–205.

Hassold TJ, Jacobs PA. 1984. Trisomy in man. *Annu Rev Genet* **18**: 69–97.

Hastings PJ, Ira G, Lupski JR. 2009. A Microhomology-Mediated Break-Induced Replication Model for the Origin of Human Copy Number Variation ed. I. Matic. *PLoS Genet* **5**: e1000327.

Hatch EM, Fischer AH, Deerinck TJ, Hetzer MW. 2013. Catastrophic Nuclear Envelope Collapse in Cancer Cell Micronuclei. *Cell* **154**: 47.

He B, Gnawali N, Hinman AW, Mattingly AJ, Osimani A, Cimini D. 2019. Chromosomes missegregated into micronuclei contribute to chromosomal instability by missegregating at the next division. *Oncotarget* **10**: 2660–2674.

Heilig CE, Löffler H, Mahlknecht U, Janssen JWG, Ho AD, Jauch A, Krämer A. 2010. Chromosomal instability correlates with poor outcome in patients with myelodysplastic syndromes irrespectively of the cytogenetic risk group. *J Cell Mol Med* **14**: 895–902.

Henrichsen CN, Vinckenbosch N, Zöllner S, Chaignat E, Pradervand S, Schütz F, Ruedi M, Kaessmann H, Reymond A. 2009. Segmental copy number variation shapes tissue transcriptomes. *Nat Genet* **41**: 424–429.

Heselmeyer K, Schröck E, Du Manoir S, Blegen H, Shah K, Steinbeck R, Auer G, Ried T. 1996. Gain of chromosome 3q defines the transition from severe dysplasia to invasive carcinoma of the uterine cervix. *Proc Natl Acad Sci U S A* **93**: 479–484.

Hieronymus H, Murali R, Tin A, Yadav K, Abida W, Moller H, Berney D, Scher H, Carver B, Scardino P, et al. 2018. Tumor copy number alteration burden is a pan-cancer prognostic factor associated with recurrence and death. *Elife* **7**.

Hoadley KA, Yau C, Hinoue T, Wolf DM, Lazar AJ, Drill E, Shen R, Taylor AM, Cherniack AD, Thorsson V, et al. 2018. Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. *Cell* **173**: 291-304.e6.

Hveem TS, Merok MA, Pretorius ME, Novelli M, Bævre MS, Sjo OH, Clinch N, Liestøl K, Svindland A, Lothe RA, et al. 2014. Prognostic impact of genomic instability in colorectal cancer. *Br J Cancer* **110**: 2159–2164.

Idbaih A, Marie Y, Lucchesi C, Pierron G, Manié E, Raynal V, Mosseri V, Hoang-Xuan K, Kujas M, Brito I, et al. 2008. BAC array CGH distinguishes mutually exclusive alterations that define clinicogenetic subtypes of gliomas. *Int J Cancer* **122**: 1778–1786.

Iourov IY, Vorsanova SG, Zelenova MA, Korostelev SA, Yurov YB. 2015. Genomic copy number variation affecting genes involved in the cell cycle pathway: Implications for somatic mosaicism. *Int J Genomics* **2015**.

Itsara A, Wu H, Smith JD, Nickerson DA, Romieu I, London SJ, Eichler EE. 2010. De novo rates and selection of large copy number variation. *Genome Res* **20**: 1469–1481.

Jacobs KB, Yeager M, Zhou W, Wacholder S, Wang Z, Rodriguez-Santiago B, Hutchinson A, Deng X, Liu C, Horner MJ, et al. 2012. Detectable clonal mosaicism and its relationship to aging and cancer. *Nat Genet* **44**: 651–658.

Janssen A, Van Der Burg M, Szuhai K, Kops GJPL, Medema RH. 2011. Chromosome segregation errors as a cause of DNA damage and structural chromosome aberrations. *Science (80-)* **333**: 1895–1898.

Jia CW, Wang L, Lan YL, Song R, Zhou LY, Yu L, Yang Y, Liang Y, Li Y, Ma YM, et al. 2015. Aneuploidy in early miscarriage and its related factors. *Chin Med J (Engl)* **128**: 2772–2776.

Kawankar N, Vundinti BR. 2011. Cytogenetic abnormalities in myelodysplastic syndrome: An overview. *Hematology* **16**: 131–138.

Killcoyne S, Gregson E, Wedge DC, Woodcock DJ, Eldridge MD, de la Rue R, Miremadi A, Abbas S, Blasko A, Kosmidou C, et al. 2020. Genomic copy number predicts esophageal cancer years before transformation. *Nat Med*.

Knouse KA, Davoli T, Elledge SJ, Amon A. 2017. Aneuploidy in Cancer: Seq-ing Answers to Old Questions. *Annu Rev Cancer Biol* **1**: 335–354.

Knouse KA, Lopez KE, Bachofner M, Amon A. 2018. Chromosome Segregation Fidelity in Epithelia Requires Tissue Architecture. *Cell* **175**: 200-211.e13.

Knouse KA, Wu J, Amon A. 2016. Assessment of megabase-scale somatic copy number variation using single-cell sequencing. *Genome Res* **26**: 376–384.

Knouse KA, Wu J, Whittaker CA, Amon A. 2014. Single cell sequencing reveals low levels of aneuploidy across mammalian tissues. *Proc Natl Acad Sci* **111**: 13409–13414.

Korbel JO, Tirosh-Wagner T, Urban AE, Chen XN, Kasowski M, Dai L, Grubert F, Erdman C, Gao MC, Lange K, et al. 2009. The genetic architecture of Down syndrome phenotypes revealed by high-resolution analysis of human segmental trisomies. *Proc Natl Acad Sci U S A* **106**: 12031–12036.

Kristensen GB, Kildal W, Abeler VM, Kaern J, Vergote I, Tropé CG, Danielsen HE. 2003. Large-scale genomic instability predicts long-term outcome for women with invasive stage I ovarian cancer. *Ann Oncol* **14**: 1494–1500.

Kuznetsova AY, Seget K, Moeller GK, de Pagter MS, de Roos JADM, Dürrbaum M, Kuffer C, Müller S, Zaman GJR, Kloosterman WP, et al. 2015. Chromosomal instability, tolerance of mitotic errors and multidrug resistance are promoted by tetraploidization in human cells. *Cell Cycle* **14**: 2810–2820.

Lamm N, Masamsetti VP, Read M, Biro M, Cesare A. 2018. ATR and mTOR regulate F-actin to alter nuclear architecture and repair replication stress. *bioRxiv* 451708.

Laubert T, Freitag-Wolf S, Linnebacher M, König A, Vollmar B, Habermann JK. 2015. Stage-specific frequency and prognostic significance of aneuploidy in patients with sporadic colorectal cancer—a meta-analysis and current overview. *Int J Colorectal Dis* **30**: 1015–1028.

Laucius CD, Orr B, Compton DA. 2019. Chromosomal instability suppresses the growth of K-Ras-induced lung adenomas. *Cell Cycle* **18**: 1702–1713.

Laughney AM, Elizalde S, Genovese G, Bakhoum SF. 2015. Dynamics of Tumor Heterogeneity Derived from Clonal Karyotypic Evolution. *Cell Rep* **12**: 809–820.

Laurie CC, Laurie CA, Rice K, Doheny KF, Zelnick LR, McHugh CP, Ling H, Hetrick KN, Pugh EW, Amos C, et al. 2012. Detectable clonal mosaicism from birth to old age and its relationship to cancer. *Nat Genet* **44**: 642–650.

Lechman ER, Hermans KG, Dobson SM, Gan OI, Kennedy JA, Chen J, Hitzler JK, Dick JE. 2014. Modeling the Multi-Step Pathogenesis of Acute Myeloid Leukemia of Down Syndrome. *Blood* **124**: 3579–3579.

Lee JA, Carvalho CMB, Lupski JR. 2007. A DNA Replication Mechanism for Generating Nonrecurrent Rearrangements Associated with Genomic Disorders. *Cell* **131**: 1235–1247.

Lengauer C, Kinzler KW, Vogelstein B. 1997. Genetic instability in colorectal cancers. *Nature* **386**: 623–627.

Levine MS, Bakker B, Boeckx B, Moyett J, Lu J, Vitre B, Spierings DC, Lansdorp PM, Cleveland DW, Lambrechts D, et al. 2017. Centrosome Amplification Is Sufficient to Promote Spontaneous Tumorigenesis in Mammals. *Dev Cell* **40**: 313-322.e5.

Li M, Fang X, Wei Z, York JP, Zhang P. 2009. Loss of spindle assembly checkpoint-mediated inhibition of Cdc20 promotes tumorigenesis in mice. *J Cell Biol* **185**: 983–94.

Lin HY, Lin SP, Chen YJ, Hung HY, Kao HA, Hsu CH, Chen MR, Chang JH, Ho CS, Huang FY, et al. 2006. Clinical characteristics and survival of trisomy 18 in a medical center in Taipei, 1988-2004. *Am J Med Genet* **140 A**: 945–951.

List A, Ebert BL, Fenaux P. 2018. A decade of progress in myelodysplastic syndrome with chromosome 5q deletion. *Leukemia* **32**: 1493–1499.

Liu D, Vader G, Vromans MJM, Lampson MA, Lens SMA. 2009a. Sensing Chromosome Bi-Orientation Kinase from Kinetochore Substrates. *Science* **323**: 1350–1353.

Liu W, Laitinen S, Khan S, Vihinen M, Kowalski J, Yu G, Chen L, Ewing CM, Eisenberger MA, Carducci MA, et al. 2009b. Copy number analysis indicates monoclonal origin of lethal metastatic prostate cancer. *Nat Med* **15**: 559–565.

Liu X, Wu H, Loring J, Hormuzdi S, Disteche CM, Bornstein P, Jaenisch R. 1997. Trisomy eight in ES cells is a common potential problem in gene targeting and interferes with germ line transmission. *Dev Dyn* **209**: 85–91.

Liu Y, Chen C, Xu Z, Scuoppo C, Rillahan CD, Gao J, Spitzer B, Bosbach B, Kastenhuber ER, Baslan T, et al. 2016. Deletions linked to TP53 loss drive cancer through p53-independent mechanisms. *Nature* **531**: 471–475.

Macintyre G, Goranova TE, De Silva D, Ennis D, Piskorz AM, Eldridge M, Sie D, Lewsley LA, Hanif A, Wilson C, et al. 2018. Copy number signatures and mutational processes in ovarian carcinoma. *Nat Genet* **50**: 1262–1270.

MacKenzie KJ, Carroll P, Martin CA, Murina O, Fluteau A, Simpson DJ, Olova N, Sutcliffe H, Rainger JK, Leitch A, et al. 2017. CGAS surveillance of micronuclei links genome instability to innate immunity. *Nature* **548**: 461–465.

Makohon-Moore AP, Zhang M, Reiter JG, Bozic I, Allen B, Kundu D, Chatterjee K, Wong F, Jiao Y, Kohutek ZA, et al. 2017. Limited heterogeneity of known driver gene mutations among the metastases of individual patients with pancreatic cancer. *Nat Genet* **49**: 358–366.

Martinez P, Mallo D, Paulson TG, Li X, Sanchez CA, Reid BJ, Graham TA, Kuhner MK, Maley CC. 2018. Evolution of Barrett's esophagus through space and time at single-crypt and whole-biopsy levels. *Nat Commun* **9**: 1–12.

Mayshar Y, Ben-David U, Lavon N, Biancotti JC, Yakir B, Clark AT, Plath K, Lowry WE, Benvenisty N. 2010. Identification and classification of chromosomal aberrations in human induced pluripotent stem cells. *Cell Stem Cell* **7**: 521–531.

McCarron M, McCallion P, Reilly E, Mulryan N. 2014. A prospective 14-year longitudinal follow-up of dementia in persons with Down syndrome. *J Intellect Disabil Res* **58**: 61–70.

Meijer GA, Hermsen MA, Baak JP, van Diest PJ, Meuwissen SG, Beliën JA, Hoovers JM, Joenje H, Snijders PJ, Walboomers JM. 1998. Progression from colorectal adenoma to carcinoma is associated with non-random chromosomal gains as detected by comparative genomic hybridisation. *J Clin Pathol* **51**: 901–9.

Michel LS, Liberal V, Chatterjee A, Kirchwegger R, Pasche B, Gerald W, Dobles M, Sorger PK, Murty VVVS, Benezra R. 2001. MAD2 haplo-insufficiency causes premature anaphase and chromosome instability in mammalian cells. *Nature* **409**: 355–359.

Mitchell TJ, Turajlic S, Rowan A, Nicol D, Farmery JHR, O'Brien T, Martincorena I, Tarpey P, Angelopoulos N, Yates LR, et al. 2018. Timing the Landmark Events in the Evolution of Clear Cell Renal Cell Cancer: TRACERx Renal. *Cell* **173**: 611-623.e17.

Mouradov D, Domingo E, Gibbs P, Jorissen RN, Li S, Soo PY, Lipton L, Desai J, Danielsen HE, Oukrif D, et al. 2013. Survival in stage II/III colorectal cancer is independently predicted by chromosomal and microsatellite instability, but not by specific driver mutations. *Am J Gastroenterol* **108**: 1785–1793.

Myrie KA, Percy MJ, Azim JN, Neeley CK, Petty EM. 2000. Mutation and expression analysis of human BUB1 and BUB1B in aneuploid breast cancer cell lines. *Cancer Lett* **152**: 193–199.

Nassar D, Latil M, Boeckx B, Lambrechts D, Blanpain C. 2015. Genomic landscape of carcinogen-induced and genetically induced mouse skin squamous cell carcinoma. *Nat Med* **21**: 946–954.

Nibourel O, Guihard S, Roumier C, Pottier N, Terre C, Paquet A, Peyrouze P, Geffroy S, Quentin S, Alberdi A, et al. 2017. Copy-number analysis identified new prognostic marker in acute myeloid leukemia. *Leukemia* **31**: 555–564.

Nigg EA. 2002. Centrosome aberrations: Cause or consequence of cancer progression? *Nat Rev Cancer* **2**: 815–825.

Nigg EA. 2006. Origins and consequences of centrosome aberrations in human cancers. *Int J Cancer* **119**: 2717–2723.

Nishizaki T, Harada K, Kubota H, Harada K, Furuya T, Suzuki M, Sasaki K. 2002. Chromosome instability in malignant astrocytic tumors detected by fluorescence in situ hybridization. *J Neurooncol* **56**: 159–165.

Nord H, Segersten U, Sandgren J, Wester K, Busch C, Menzel U, Komorowski J, Dumanski JP, Malmström PU, De Ståhl TD. 2010. Focal amplifications are associated with high grade and recurrences in stage Ta bladder carcinoma. *Int J Cancer* **126**: 1390–1402.

Oromendia AB, Dodgson SE, Amon A. 2012. Aneuploidy causes proteotoxic stress in yeast. *Genes Dev* **26**: 2696–2708.

Passerini V, Ozeri-Galai E, Pagter MS, Donnelly N, Schmalbrock S, Kloosterman WP, Kerem B, Storchova Z. 2016. The presence of extra chromosomes leads to genomic instability. *Nature Communications* **7**:1-12.

Pavelka N, Rancati G, Zhu J, Bradford WD, Saraf A, Florens L, Sanderson BW, Hattem GL, Li R. 2010a. Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. *Nature* **468**: 321–325.

Pihan GA, Purohit A, Wallace J, Knecht H, Woda B, Quesenberry P, Doxsey SJ. 1998. Centrosome defects and genetic instability in malignant tumors. *Cancer Res* **58**: 3974–3985.

Pihan GA, Wallace J, Zhou Y, Doxsey SJ. 2003. Centrosome abnormalities and chromosome instability occur together in pre-invasive carcinomas. *Cancer Res* **63**: 1398–1404.

Plaja, A., Perez, C., Miró, R. 2001. Chromosome aneuploidy and cancer: lessons from a chromosomal instability syndrome. *Cancer Genet Cytogenet* **131**: 144-145.

Pinto AE, Pereira T, Santos M, Branco M, Dias Â, Silva GL, Ferreira MC, André S. 2013. DNA ploidy is an independent predictor of survival in breast invasive ductal carcinoma: A long-term multivariate analysis of 393 patients. *Ann Surg Oncol* **20**: 1530–1537.

Ravichandran MC, Fink S, Clarke MN, Hofer FC, Campbell CS. 2018. Genetic interactions between specific chromosome copy number alterations dictate complex aneuploidy patterns. *Genes Dev* **32**: 1485–1498.

Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, et al. 2006. Global variation in copy number in the human genome. *Nature* **444**: 444–454.

Reynolds LE, Watson AR, Baker M, Jones TA, D'Amico G, Robinson SD, Joffre C, Garrido-Urbani S, Rodriguez-Manzaneque JC, Martino-Echarri E, et al. 2010. Tumour angiogenesis is reduced in the Tc1 mouse model of Downs syndrome. *Nature* **465**: 813–817.

Ricke RM, Jeganathan KB, van Deursen JM. 2011. Bub1 overexpression induces aneuploidy and tumor formation through Aurora B kinase hyperactivation. *J Cell Biol* **193**: 1049–1064.

Ried T, Knutzen R, Steinbeck R, Blegen H, Schröck E, Heselmeyer K, du Manoir S, Auer G. 1996. Comparative genomic hybridization reveals a specific pattern of chromosomal gains and losses during the genesis of colorectal tumors. *Genes, Chromosom Cancer* **15**: 234–245.

Ring D, Hubble R, Kirschner M. 1982. Mitosis in a cell with multiple centrioles. *J Cell Biol* **94**: 549–556.

Roper RJ, Reeves RH. 2006. Understanding the basis for Down syndrome phenotypes. *PLoS Genet* **2**: 0231–0236.

Roschke A V., Glebov OK, Lababidi S, Gehlhaus KS, Weinstein JN, Kirsch IR. 2008. Chromosomal instability is associated with higher expression of genes implicated in epithelial-mesenchymal transition, cancer invasiveness, and metastasis and with lower expression of genes involved in cell cycle checkpoints, DNA repair, and chromatin mai. *Neoplasia* **10**: 1222–1230.

Ross-Innes CS, Becq J, Warren A, Cheetham RK, Northen H, O'Donovan M, Malhotra S, di Pietro M, Ivakhno S, He M, et al. 2015. Whole-genome sequencing provides new insights into the clonal architecture of Barrett's esophagus and esophageal adenocarcinoma. *Nat Genet* **47**: 1038–1046.

Roylance R, Endesfelder D, Gorman P, Burrell RA, Sander J, Tomlinson I, Hanby AM, Speirs V, Richardson AL, Birkbak NJ, et al. 2011. Relationship of extreme chromosomal instability with long-term survival in a retrospective analysis of primary breast cancer. *Cancer Epidemiol Biomarkers Prev* **20**: 2183–2194.

Rutledge SD, Douglas TA, Nicholson JM, Vila-Casadesús M, Kantzler CL, Wangsa D, Barroso-Vilares M, Kale SD, Logarinho E, Cimini D. 2016. Selective advantage of trisomic human cells cultured in non-standard conditions. *Sci Rep* **6**: 1–12.

Sack LM, Davoli T, Li MZ, Li Y, Xu Q, Naxerova K, Wooten EC, Bernardi RJ, Martin TD, Chen T, et al. 2018. Profound Tissue Specificity in Proliferation Control Underlies Cancer Drivers and Aneuploidy Patterns. *Cell* **173**: 499-514.e23.

Santaguida, S. Amon, A. 2015. Short- and long-term effects of chromosome mis-segrgation and aneuploidy. *Nat Rev Mol Cell Bio* **16**: 473-485.

Santaguida S, Richardson A, Iyer DR, M'Saad O, Zasadil L, Knouse KA, Wong YL, Rhind N, Desai A, Amon A. 2017. Chromosome Mis-segregation Generates Cell-Cycle-Arrested Cells with Complex Karyotypes that Are Eliminated by the Immune System. *Dev Cell* **41**: 638-651.e5.

Satgé D, Sommelet D, Geneix A, Nishi M, Malet P, Vekemans M. 1998. A tumor profile in Down syndrome. *Am J Med Genet* **78**: 207–216.

Schanz J, Steidl C, Fonatsch C, Pfeilstöcker M, Nösslinger T, Tuechler H, Valent P, Hildebrandt B, Giagounidis A, Aul C, et al. 2011. Coalesced multicentric analysis of 2,351 patients with myelodysplastic syndromes indicates an underestimation of poor-risk cytogenetics of myelodysplastic syndromes in the International Prognostic Scoring System. *J Clin Oncol* **29**: 1963–1970.

Selvarajah S, Yoshimoto M, Ludkovski O, Park PC, Bayani J, Thorner P, Maire G, Squire JA, Zielenska M. 2008. Genomic signatures of chromosomal instability and osteosarcoma progression detected by high resolution array CGH and interphase FISH. *Cytogenet Genome Res* **122**: 5–15.

Sharp AJ, Locke DP, McGrath SD, Cheng Z, Bailey JA, Vallente RU, Pertz LM, Clark RA, Schwartz S, Segraves R, et al. 2005. Segmental duplications and copy-number variation in the human genome. *Am J Hum Genet* **77**: 78–88.

Sheltzer JM, Blank HM, Pfau SJ, Tange Y, George BM, Humpton TJ, Brito IL, Hiraoka Y, Niwa O, Amon A. 2011. Aneuploidy Drives Genomic Instability in Yeast. *Science (80-)* **333**: 1026–1030.

Sheltzer JM, Ko JH, Replogle JM, Habibe Burgos NC, Chung ES, Meehl CM, Sayles NM, Passerini V, Storchova Z, Amon A. 2017. Single-chromosome Gains Commonly Function as Tumor Suppressors. *Cancer Cell* **31**: 240–255.

Sheltzer JM, Torres EM, Dunham MJ, Amon A. 2012. Transcriptional consequences of aneuploidy. *Proc Natl Acad Sci* **109**: 12644–12649.

Silk AD, Zasadil LM, Holland AJ, Vitre B, Cleveland DW, Weaver BA. 2013. Chromosome missegregation rate predicts whether aneuploidy will promote or suppress tumors. *Proc Natl Acad Sci* **110**: E4134–E4141.

Silkworth WT, Nardi IK, Scholl LM, Cimini D. 2009. Multipolar Spindle Pole Coalescence Is a Major Source of Kinetochore Mis-Attachment and Chromosome Mis-Segregation in Cancer Cells ed. K.G. Hardwick. *PLoS One* **4**: e6564.

Sinicrope FA, Rego RL, Halling KC, Foster N, Sargent DJ, La Plant B, French AJ, Laurie JA, Goldberg RM, Thibodeau SN, et al. 2006. Prognostic Impact of Microsatellite Instability and DNA Ploidy in Human Colon Carcinoma Patients. *Gastroenterology* **131**: 729–737.

Snape K, Hanks S, Ruark E, Barros-Núñez P, Elliott A, Murray A, Lane AH, Shannon N, Callier P, Chitayat D, et al. 2011. Mutations in CEP57 cause mosaic variegated aneuploidy syndrome. *Nat Genet* **43**: 527–529.

Solomon DA, Kim T, Diaz-Martinez LA, Fair J, Elkahloun AG, Harris BT, Toretsky JA, Rosenberg SA, Shukla N, Ladanyi M, et al. 2011. Mutational inactivation of STAG2 causes aneuploidy in human cancer. *Science (80-)* **333**: 1039–1043.

Sotillo R, Hernando E, Díaz-Rodríguez E, Teruya-Feldstein J, Cordón-Cardo C, Lowe SW, Benezra R. 2007. Mad2 Overexpression Promotes Aneuploidy and Tumorigenesis in Mice. *Cancer Cell* **11**: 9–23.

Sotillo R, Schvartzman J-M, Socci ND, Benezra R. 2010. Mad2-induced chromosome instability leads to lung tumour relapse after oncogene withdrawal. *Nature* **464**: 436–440.

Soto M, García-Santisteban I, Krenning L, Medema RH, Raaijmakers JA. 2018. Chromosomes trapped in micronuclei are liable to segregation errors. *J Cell Sci* **131**.

Steinbeck RG, Heselmeyer KM, Auer GU. 1994. DNA ploidy in human colorectal adenomas. *Anal Quant Cytol Histol* **16**: 196–202.

Stopsack KH, Whittaker CA, Gerke TA, Loda M, Kantoff PW, Mucci LA, Amon A. 2019. Aneuploidy drives lethal progression in prostate cancer. *Proc Natl Acad Sci U S A* **166**: 11390–11395.

Storchova Z, Kuffer C. 2008. The consequences of tetraploidy and aneuploidy. *J Cell Sci* **121**: 3859–3866.

Tang Y-C, Williams BR, Siegel JJ, Amon A. 2011a. Identification of Aneuploidy-Selective Antiproliferation Compounds. *Cell* **144**: 499–512.

Taylor AM, Shih J, Ha G, Gao GF, Zhang X, Berger AC, Schumacher SE, Wang C, Hu H, Liu J, et al. 2018. Genomic and Functional Approaches to Understanding Cancer Aneuploidy. *Cancer Cell* **33**: 676-689.e3.

Teixeira VH, Pipinikas CP, Pennycuick A, Lee-Six H, Chandrasekharan D, Beane J, Morris TJ, Karpathakis A, Feber A, Breeze CE, et al. 2019. Deciphering the genomic, epigenomic, and transcriptomic landscapes of pre-invasive lung cancer lesions. *Nat Med* **25**: 517–525.

Thomas R, Marks DH, Chin Y, Benezra R. 2018. Whole chromosome loss and associated breakage–fusion–bridge cycles transform mouse tetraploid cells. *EMBO J* **37**: 201–218.

Thompson SL, Compton DA. 2008. Examining the link between chromosomal instability and aneuploidy in human cells. *J Cell Biol* **180**: 665–672.

Thorburn RR, Gonzalez C, Brar GA, Christen S, Carlile TM, Ingolia NT, Sauer U, Weissman JS, Amon A. 2013. Aneuploid yeast strains exhibit defects in cell growth and passage through START. *Mol Biol Cell* **24**: 1274–1289.

Torres EM, Dephoure N, Panneerselvam A, Tucker CM, Whittaker CA, Gygi SP, Dunham MJ, Amon A. 2010. Identification of Aneuploidy-Tolerating Mutations. *Cell* **143**: 71–83.

Torres EM, Sokolsky T, Tucker CM, Chan LY, Boselli M, Dunham MJ, Amon A. 2007. Effects of aneuploidy on cellular physiology and cell division in haploid yeast. *Science* **317**: 916–924.

Torti D, Trusolino L. 2011. Oncogene addiction as a foundational rationale for targeted anti-cancer therapy: Promises and perils. *EMBO Mol Med* **3**: 623–636.

Turajlic S, Xu H, Litchfield K, Rowan A, Chambers T, Lopez JI, Nicol D, O'Brien T, Larkin J, Horswell S, et al. 2018a. Tracking Cancer Evolution Reveals Constrained Routes to Metastases: TRACERx Renal. *Cell* **173**: 581-594.e12.

Turajlic S, Xu H, Litchfield K, Rowan A, Horswell S, Chambers T, O'Brien T, Lopez JI, Watkins TBK, Nicol D, et al. 2018b. Deterministic Evolutionary Trajectories Influence Primary Tumor Growth: TRACERx Renal. *Cell* **0**: 1–16.

Upender MB, Habermann JK, McShane LM, Korn EL, Barrett JC, Difilippantonio MJ, Ried T. 2004. Chromosome transfer induced aneuploidy results in complex dysregulation of the cellular transcriptome in immortalized and cancer cells. *Cancer Res* **64**: 6941–6949.

Van Den Bent MJ, Brandes AA, Taphoorn MJB, Kros JM, Kouwenhoven MCM, Delattre JY, Bernsen HJJA, Frenay M, Tijssen CC, Grisold W, et al. 2013. Adjuvant procarbazine, lomustine, and vincristine chemotherapy in newly diagnosed anaplastic oligodendroglioma: Long-term follow-up of EORTC brain tumor group study 26951. *J Clin Oncol* **31**: 344–350.

Van Ree JH, Jeganathan KB, Malureanu L, Van Deursen JM. 2010. Overexpression of the E2 ubiquitin-conjugating enzyme UbcH10 causes chromosome missegregation and tumor formation. *J Cell Biol* **188**: 83–100.

Vasudevan A, Baruah PS, Smith JC, Wang Z, Sayles NM, Andrews P, Kendall J, Leu J, Chunduri NK, Levy D, et al. 2020. Single-Chromosomal Gains Can Function as Metastasis Suppressors and Promoters in Colon Cancer. *Dev Cell* **52**: 413-428.e6.

Vogelstein B, Fearon ER, Hamilton SR, Kern SE, Preisinger AC, Leppert M, Smits AMM, Bos JL. 1988. Genetic Alterations during Colorectal-Tumor Development. *N Engl J Med* **319**: 525–532.

Vogelstein B, Fearon ER, Kern SE, Hamilton SR, Preisinger AC, Nakamura Y, White R. 1989. Allelotype of colorectal carcinomas. *Science* **244**: 207–11.

Vogelstein B, Kinzler KW. 2004. Cancer genes and the pathways they control. *Nat Med* **10**: 789–799.

Vogelstein B, Lane D, Levine AJ. 2000. Surfing the p53 network. *Nature* **408**: 307–310.

Wahl M, Phillips JJ, Molinaro AM, Lin Y, Perry A, Haas-Kogan DA, Costello JF, Dayal M, Butowski N, Clarke JL, et al. 2017. Chemotherapy for adult low-grade gliomas: Clinical outcomes by molecular subtype in a phase II study of adjuvant temozolomide. *Neuro Oncol* **19**: 242–251.

Walther A, Houlston R, Tomlinson I. 2008. Association between chromosomal instability and prognosis in colorectal cancer: A meta-analysis. *Gut* **57**: 941–950.

Wang RW, Vigano' S, Ben-David U, Amon A, Santaguida S. 2020. Aneuploid cells activate NF-κB to promote their immune clearance by NK cells. *bioRxiv* 2020.06.25.172239.

Wang Y, Waters J, Leung ML, Unruh A, Roh W, Shi X, Chen K, Scheet P, Vattathil S, Liang H, et al. 2014. Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature* **512**: 155–160.

Wang Z, Cummins JM, Shen D, Cahill DP, Jallepalli P V., Wang TL, Parsons DW, Traverso G, Awad M, Silliman N, et al. 2004. Three Classes of Genes Mutated in Colorectal Cancers with Chromosomal Instability. *Cancer Res* **64**: 2998–3001.

Wangsa D, Braun R, Schiefer M, Gertz EM, Bronder D, Quintanilla I, Padilla-Nash HM, Torres I, Hunn C, Warner L, et al. 2018. The evolution of single cell-derived colorectal cancer cell lines is dominated by the continued selection of tumor-specific genomic imbalances, despite random chromosomal instability. *Carcinogenesis* **39**: 993–1005.

Weaver BAA, Silk AD, Montagna C, Verdier-Pinard P, Cleveland DW. 2007. Aneuploidy Acts Both Oncogenically and as a Tumor Suppressor. *Cancer Cell* **11**: 25–36.

Weller M, Stupp R, Hegi ME, Van Den Bent M, Tonn JC, Sanson M, Wick W, Reifenberger G. 2012. Personalized care in neuro-oncology coming of age: Why we need MGMT and 1p/19q testing for malignant glioma patients in clinical practice. *Neuro Oncol* **14**: iv100–iv108.

Westcott PMK, Halliwill KD, To MD, Rashid M, Rust AG, Keane TM, Delrosario R, Jen KY, Gurley KE, Kemp CJ, et al. 2015. The mutational landscapes of genetic and chemical models of Kras-driven lung cancer. *Nature* **517**: 489–492.

Wick W, Hartmann C, Engel C, Stoffels M, Felsberg J, Stockhammer F, Sabel MC, Koeppen S, Ketter R, Meyermann R, et al. 2009. NOA-04 randomized phase III trial of sequential radiochemotherapy of anaplastic glioma with Procarbazine, Lomustine, and Vincristine or Temozolomide. *J Clin Oncol* **27**: 5874–5880.

Wiestler B, Capper D, Sill M, Jones DTW, Hovestadt V, Sturm D, Koelsche C, Bertoni A, Schweizer L, Korshunov A, et al. 2014. Integrated DNA methylation and copy-number profiling identify three clinically and biologically relevant groups of anaplastic glioma. *Acta Neuropathol* **128**: 561–571.

Wijshake T, Malureanu LA, Baker DJ, Jeganathan KB, van de Sluis B, van Deursen JM. 2012. Reduced Life- and Healthspan in Mice Carrying a Mono-Allelic BubR1 MVA Mutation ed. P. Hasty. *PLoS Genet* **8**: e1003138.

Williams BR, Prabhu VR, Hunter KE, Glazier CM, Whittaker CA, Housman DE, Amon A. 2008. Aneuploidy Affects Proliferation and Spontaneous Immortalization in Mammalian Cells. *Science (80-)* **322**: 703–709.

Wiseman FK, Pulford LJ, Barkus C, Liao F, Portelius E, Webb R, Chávez-Gutiérrez L, Cleverley K, Noy S, Sheppard O, et al. 2018. Trisomy of human chromosome 21 enhances amyloid-b deposition independently of an extra copy of APP. *Brain* **141**: 2457–2474.

Yang A, Reeves RH. 2011. Increased survival following tumorigenesis in Ts65Dn mice that model down syndrome. *Cancer Res* **71**: 3573–3581.

Yona AH, Manor YS, Herbst RH, Romano GH, Mitchell A, Kupiec M, Pilpel Y, Dahan O. 2012. Chromosomal duplication is a transient evolutionary solution to stress. *Proc Natl Acad Sci U S A* **109**: 21010–21015.

Zack TI, Schumacher SE, Carter SL, Cherniack AD, Saksena G, Tabak B, Lawrence MS, Zhang CZ, Wala J, Mermel CH, et al. 2013. Pan-cancer patterns of somatic copy number alteration. *Nat Genet* **45**: 1134–1140.

Zhang CZ, Spektor A, Cornils H, Francis JM, Jackson EK, Liu S, Meyerson M, Pellman D. 2015. Chromothripsis from DNA damage in micronuclei. *Nature* **522**: 179–184.

Chapter Two: Aneuploidy impairs hematopoietic stem cell fitness and is selected against in regenerating tissues in vivo.

Pfau SJ, Silberman RE, Knouse KK, Amon A. 2016. Aneuploidy impairs hematopoietic stem cell fitness and is selected against in regenerating tissues *in vivo. Genes & Development* **30**:1395-1408.

Abstract

Aneuploidy, an imbalanced karyotype, is a widely observed feature of cancer cells that has long been hypothesized to promote tumorigenesis. Here we evaluate the fitness of cells with constitutional trisomy or chromosomal instability (CIN) *in vivo* using hematopoietic reconstitution experiments. We did not observe cancer, but instead find that aneuploid hematopoietic stem cells (HSCs) exhibit decreased fitness. This reduced fitness is due at least in part to the decreased proliferative potential of aneuploid hematopoietic cells. Analyses of mice with CIN caused by a hypomorphic mutation in the gene *Bub1b* further support the finding that aneuploidy impairs cell proliferation *in vivo*. Whereas non-regenerating adult tissues are highly aneuploid in these mice, HSCs and other regenerative adult tissues are largely euploid. These findings indicate that *in vivo*, mechanisms exist to select against aneuploid cells.

Introduction

More than 90 percent of human solid tumors are aneuploid (Rajagopalan and Lengauer 2004). It was initially hypothesized that aneuploidy may play a causal role in tumorigenesis in 1914 (Boveri 1914), and recent analyses of tumor genomes support this idea, finding that aneuploidies observed in cancer can in part be explained by loss of tumor suppressor genes and gain of oncogenes (Davoli et al. 2013). Despite the prevalence of imbalanced karyotypes in cancer, *in vitro* studies of aneuploid cells have not provided decisive evidence that aneuploidy alone can promote tumorigenesis. Trisomic human (Segal and McCoy 1974; Stingele et al. 2012) and mouse (Williams et al. 2008) cells exhibit decreased proliferation; and cells harboring mutations that promote CIN either proliferate normally (Babu et al. 2003; Jeganathan et al. 2007; Weaver et al. 2007) or more slowly, indicating that aneuploidy is, at best, fitness neutral. Furthermore, studies in mouse models of CIN and Down Syndrome (DS) have demonstrated that these conditions can both promote and inhibit tumorigenesis *in vivo* (Pfau and Amon 2012). These observations make clear the need for additional experimental systems to assess the effects of aneuploidy *per se* on cell proliferation *in vivo*. However, systematic evaluation of aneuploid cell fitness *in vivo* is difficult because most autosomal aneuploidies are embryonic lethal in mammals.

To bypass the embryonic lethality of aneuploidy in mice, we performed transplantation experiments with hematopoietic stem cells (HSCs) isolated from mouse embryos. HSCs provide a cellular system that is amenable for systematic study of the effects of aneuploidy *in vivo* because HSCs give rise to all differentiated blood cell types (Kondo et al. 2003). Whereas HSCs reside in the bone marrow of adult mice*, they can be found in the fetal liver between embryonic stages E12.5 and E15.5* (Orkin and Zon 2008), *allowing the isolation of HSCs from fetal livers of aneuploid embryos that survive to this embryonic age.*

High doses of irradiation ablate the hematopoietic function of HSCs; however, transplantation of HSCs from a non-irradiated donor can reconstitute the hematopoietic system of an irradiated recipient. Thus, aneuploid fetal liver HSCs (FL-HSCs) can be used to reconstitute the blood of a lethally irradiated wild type recipient, giving rise to an otherwise wild type mouse with aneuploid blood. Hematopoietic reconstitutions performed with FL-HSCs derived from euploid wild type littermates allow for direct comparison of aneuploid and euploid HSC fitness. Thus, the hematopoietic compartment represents an ideal in vivo system to assess the fitness of aneuploid karyotypes that cause embryonic lethality. Analyzing HSC fitness in vivo not only permits evaluation of aneuploid cell proliferative capacity, but also provides insight into how aneuploidy affects stem cell potential, as previous studies have demonstrated that aneuploidy can perturb stem cell viability and differentiation (Adorno et al. 2013; Gogendeau et al. 2015). *Finally, this system also offers a unique model for studying aneuploid cells in the context of an otherwise euploid environment, a setting typical of tumorigenesis in vivo.*

We chose three mouse models of aneuploidy to evaluate the fitness of aneuploid HSCs: constitutional trisomy 16, constitutional trisomy 19 and mice harboring a hypomorphic allele of the gene Bub1b, which encodes the spindle assembly checkpoint protein BUBR1 ($Bub1b^{H/H}$) (Baker et al. 2004). Mouse chromosome 16 is the closest whole chromosome homolog to human chromosome 21 and thus a model of DS. *This model is of particular interest because individuals with DS often show perturbations in the hematopoietic lineage* (Henry et al. 2007; Choi 2008) *and have a greatly increased risk of developing childhood leukemia* (Satge et al. 1998). Mouse chromosome 19 is the smallest mouse autosome, with homology mainly to human chromosomes 9, 10 and 11. While mouse chromosome 19 is not known to harbor genes typically associated

with acute myeloid leukemia, activating mutations in JAK2 (encoded by mouse chromosome 19) are frequently associated with myeloproliferative neoplasms (Kiladjian 2012). The Bub1b$^{H/H}$ CIN model produces aneuploid cells with mostly single chromosome gains or losses, thus representing a variety of aneuploid chromosomes due to random mis-segregation events (Baker et al. 2004). *Bub1b$^{H/H}$* mice survive to adulthood, permitting comparison of both fetal liver and adult bone marrow *Bub1b$^{H/H}$* HSCs to constitutional trisomic FL-HSCs. *Bub1b$^{H/H}$* mice develop progeria-like symptoms and have a decreased lifespan but do not develop cancer (Baker et al. 2004).

Comparison of HSCs from these three models has revealed a range of responses to aneuploidy in the blood and permitted differentiation between chromosome-specific and general effects of aneuploidy *in vivo*. We find that while some aneuploidies can be well-tolerated in the hematopoietic lineage, aneuploidy generally causes a decrease in HSC fitness. This decreased fitness is at least partially due to the decreased proliferative potential of aneuploid hematopoietic cells. Additional analyses of *Bub1b$^{H/H}$* CIN mice show that aneuploidy is tolerated in this strain during periods of rapid hematopoietic population expansion. However, single cell sequencing of tissues from adult *Bub1b$^{H/H}$* mice revealed that aneuploidy is not uniformly tolerated across different adult tissue types. While tissues that are largely non-proliferative in the adult display high levels of aneuploidy, regenerative tissues harbor few, if any, aneuploid cells. These data provide evidence that aneuploidy-selective mechanisms eliminate aneuploid cells during adult hematopoiesis and likely in other tissues that regenerate during adulthood.

Results

Aneuploidy decreases HSC competitive fitness in vivo

To determine the effect of aneuploidy on cell fitness *in vivo,* we first employed competitive reconstitution assays to evaluate the fitness of aneuploid FL-HSCs. In this assay, two populations of HSCs are co-injected into a lethally irradiated recipient, and the relative contributions of each population to the hematopoietic compartment are evaluated over time by analysis of the peripheral blood. To ensure equal numbers of cells were being competed, we first measured HSC levels. Quantification by flow cytometry revealed no significant differences in the HSC levels in trisomy 16 or trisomy 19 fetal livers (Fig. 1A). Because *Bub1b*$^{H/H}$ animals are viable, we quantified HSC levels in the adult and found them to be similar to those of their wild type littermates (Supplemental Fig. S5I). Thus, we concluded that HSC levels are similar in aneuploid and euploid donors.

To assess the fitness of aneuploid HSCs, we injected equal numbers of live aneuploid or euploid littermate control fetal liver cells into a lethally irradiated euploid recipient together with the same number of live fetal liver cells from a common euploid competitor of the same embryonic age (henceforth common WT; Fig. 1B). To distinguish between experimental HSCs and the common WT competitor, each donor was tracked using a different isoform of the pan-leukocyte cell surface marker CD45, which can be distinguished by isoform-specific antibodies (CD45.1 and CD45.2). Aneuploid donors and their wild type littermates expressed the CD45.2 isoform, whereas the common WT competitor expressed the CD45.1 isoform. We chose to utilize a CD45.1 common donor because previous studies had shown that CD45.1 HSCs exhibit decreased fitness when compared to CD45.2 HSCs in competition assays (Waterstrat et al. 2010), thus giving the CD45.2 aneuploid donors a slight advantage in these experiments. Additionally, we used CD45.1 recipients to unambiguously quantify the contribution from

aneuploid and euploid wild type littermate donors. We further note that in this experimental setup, a small population of recipient-derived memory T cells remains in the recipient peripheral blood after reconstitution despite lethal irradiation (Frasca et al. 2000). This recipient-derived memory cell population is evident in our data 3 weeks after transfer (Fig. 1C-E). In these samples, an increased proportion of the blood is comprised of CD45.1 positive cells because donor-derived peripheral blood cells are present at low levels at this early stage of reconstitution, making the contribution from recipient-derived memory cells more prominent.

Evaluation of competitive reconstitution assays over time revealed a range of aneuploid cell fitness phenotypes. Trisomy 16 FL-HSCs were much less fit than FL-HSCs from their wild type littermates and were nearly completely outcompeted by the common WT donor (Fig. 1C, F). Trisomy 19 FL-HSCs exhibited slightly reduced relative fitness (Fig. 1D, F). *Bub1b*$^{H/H}$ FL-HSCs showed no relative fitness defects in this assay (Fig. 1E, F).

Because we did not observe decreased fitness in *Bub1b*$^{H/H}$ competitive reconstitutions, we assessed the level of aneuploidy in these CIN cells. We isolated peripheral white blood cells derived from *Bub1b*$^{H/H}$ FL-HSC donor cells from a recipient mouse 16 weeks after transfer. *Bub1b*$^{H/H}$ CD45.2 cells were collected by FACS, and their karyotype was determined by single cell sequencing. Analysis of 18 cells revealed 7 to be aneuploid (~38.9%; Fig. 1G; Supplemental Fig. S7A). This level of aneuploidy is on par with previously reported levels of aneuploidy for the brain (38.1%) and the liver (18.8%) (Knouse et al. 2014), as well as for stimulated splenocytes (3-33%) in adult *Bub1b*$^{H/H}$ mice (Baker et al. 2004). Therefore *Bub1b*$^{H/H}$ FL-HSCs show similar fitness to euploid wild type controls in this assay despite the prevalence of aneuploid cells in the peripheral blood.

In summary, when compared to the average fitness of euploid wild type littermates after 18 weeks, the average relative fitness of trisomy 16, trisomy 19 and *Bub1b*$^{H/H}$ FL-HSCs was 0.08, 0.84 and 1.06, respectively (Fig. 1F). Thus, in this assay some aneuploidies confer decreased HSC fitness whereas others are fitness neutral. These findings suggest that either the observed fitness decreases are due to chromosome-specific effects or that a certain level of aneuploidy is tolerated in the blood.

Decreased fitness of aneuploid HSCs is due to decreased proliferation

To determine why trisomy 16 and trisomy 19 FL-HSCs exhibited decreased relative fitness in competitive reconstitution assays, we first evaluated the ability of hematopoietic cells from each aneuploid donor to home to the bone marrow niche, a property which is essential to restore hematopoiesis during HSC transplantation (Lapidot 2005). Fetal liver cells were labeled with a fluorescent cell surface dye and injected into irradiated recipients. The number of labeled cells in the bone marrow was quantified after 24 hours. We observed no significant decrease in the homing efficiency of trisomy 16, trisomy 19 and *Bub1b*$^{H/H}$ fetal liver cells (Fig. 2A; Student's t-test), suggesting that a defect in homing is not responsible for the decreased ability of aneuploid FL-HSCs to reconstitute the hematopoietic system of lethally irradiated recipients.

We next performed colony forming unit spleen (CFU-S) assays to determine whether the proliferation of aneuploid FL-HSC-derived progenitor cells was impaired during early stages of hematopoietic reconstitution (Purton and Scadden 2007). Irradiated mice were injected with fetal liver cells, and spleen were analyzed 7 or 8 days later (Till and McCulloch 1961; Fig. 2B). Aneuploid and euploid fetal liver cells formed similar numbers of colonies in the spleen (Fig. 2C). However, quantification of the area of each CFU-S colony revealed that the average size of colonies produced by trisomy 16 and trisomy 19 donor cells was reduced compared to colonies

formed by cells from their wild type littermates (Fig. 2D). The average size of *Bub1b*$^{H/H}$ CFU-S colonies was similar to those formed by wild type littermate fetal liver cells (Fig. 2D).

To further evaluate the proliferative potential of aneuploid FL-HSCs, we measured EdU incorporation in the bone marrow of irradiated recipient mice reconstituted with trisomy 16 FL-HSCs. The number of EdU positive donor-derived cells was significantly reduced compared to wild type controls 6 days after reconstitution (Fig. 2E), indicating that trisomy 16 interferes with proliferation of FL-HSCs and progenitor cells. Consistent with this finding is the observation that trisomy 16 and trisomy 19 fetal livers were significantly smaller than fetal livers of their wild type littermates (Supplemental Fig. S1). Interestingly, *Bub1b*$^{H/H}$ fetal liver-derived cells did not exhibit decreased proliferation as evaluated by EdU incorporation 6 days after reconstitution; rather there appeared to be a (not statistically significant) trend towards increased proliferation in these cells compared to wild type control cells (Fig. 2E). The analyses of aneuploidy levels in *Bub1b*$^{H/H}$ tissues described below provide a potential explanation for this observation. Taken together, our data demonstrate that aneuploid FL-HSCs can home effectively but, in the cases of trisomies 16 and 19, show impaired proliferative potential.

Trisomy 16 recipients exhibit peripheral blood defects and decreased survival

Whereas trisomy 16 FL-HSCs exhibited a severe fitness defect in competitive reconstitution experiments, the fitness decrease was more subtle for trisomy 19 FL-HSCs and not evident for *Bub1b*$^{H/H}$ FL-HSCs. This suggests that either low levels of aneuploidy do not strongly impair HSC fitness or that HSCs were not challenged sufficiently in these assays to reveal decreased cellular fitness. To address the latter possibility, we evaluated the long-term fitness of HSCs from each aneuploid strain individually by serially transferring HSCs from primary to secondary and in some cases to tertiary and quaternary recipients (Fig. 3A). Serial transfer poses a significant challenge to the replicative potential of HSCs and their progeny,

causing the eventual exhaustion of even wild type HSCs (Harrison and Astle 1982). Trisomy 16 and 19 FL-HSCs have been evaluated previously in primary transplantation (Herbst and Winking 1991); however, these studies were not performed on an isogenic background, and therefore it is difficult to determine whether the observed phenotypes were due to aneuploidy or factors such as graft rejection (Gropp et al. 1983).

We transferred either CD45.2 trisomy 16 or CD45.2 euploid wild type littermate fetal liver cells into lethally irradiated CD45.1 primary recipient mice (Fig. 3A). Peripheral blood analyses of primary recipients performed periodically between 4 and 16 weeks after transfer revealed that trisomy 16 FL-HSCs harbor significant fitness defects when transferred individually (Fig. 3, Supplemental Fig. S2). These defects are consistent with decreased proliferation in the hematopoietic lineage. For example, trisomy 16 FL-HSCs contributed relatively fewer peripheral white blood cells than wild type HSCs (Fig. 3B). Trisomy 16 primary recipients also exhibited leukopenia – reduced white blood cell counts – when compared to primary recipients reconstituted with cells from their euploid littermates (Fig. 3C). This seemed to be largely due to low numbers of B cells (Supplemental Fig. S2A, B, I). Furthermore, complete blood cell counts showed that trisomy 16 primary recipients have comparatively fewer red blood cells and macrocytic anemia, a reduced red blood cell count accompanied by increased red blood cell volume and decreased blood hemoglobin concentration (Fig. 3D; Supplemental Fig. S2C-G; 2J). These results are in line with previous *in vivo* characterizations of trisomy 16 FL-HSCs (Herbst and Winking 1991). In addition to peripheral blood defects, trisomy 16 primary recipients exhibited decreased survival compared to primary recipient mice reconstituted with HSCs from their euploid wild type littermates (Fig. 3E).

As the ultimate test of HSC potential, secondary transfers were performed with bone marrow cells from trisomy 16 primary recipients that survived to 16 weeks. Although long-term reconstituted primary recipients had comparable HSC levels in the bone marrow (Supplemental Fig. S2K), transfer of wild type primary recipient bone marrow cells efficiently reconstituted secondary recipients while trisomy 16 primary recipient HSCs failed to reconstitute secondary recipients (Fig. 3F). We conclude that trisomy 16 FL-HSCs have a substantial fitness defect that is at least in part due to decreased proliferative potential, leading to HSC exhaustion upon secondary transfer.

Trisomy 19 is better tolerated in the blood than trisomy 16

In contrast to trisomy 16 FL-HSCs, trisomy 19 FL-HSCs contributed effectively to the white blood cell lineages in primary recipients (Fig. 4A; Supplemental Fig. S3A-J). Recipients reconstituted with trisomy 19 FL-HSCs showed normal white blood cell counts when compared to recipients reconstituted with wild type littermate FL-HSCs (Fig. 4B), exhibited normal hematocrit levels (Fig. 4C) and did not show either red blood cell defects (Supplemental Fig. S3C-G) or macrocytic anemia (Supplemental Fig. S3J). However, trisomy 19 primary recipients harbored fewer platelets compared to wild type primary recipients at early time points (Supplemental Fig. S3H). Further, the proportion of the blood comprised of each white blood cell type in trisomy 19 primary recipients was not significantly perturbed (Supplemental Fig. S3A, B, I). Consistent with trisomy 19 having little or no effect on HSC fitness in primary reconstitutions, we found that all mice transplanted with trisomy 19 HSCs survived both primary transfer of fetal liver cells and secondary transfer of bone marrow cells isolated from primary recipients (data not shown). Furthermore, secondary recipient mice showed no obvious peripheral blood defects (Supplemental Fig. S3K-U), except for a slightly weaker initial engraftment after 4 weeks. However, this delay did not persist. In fact, the engraftment at 8

weeks was stronger in the trisomy 19 HSC secondary recipients than in wild type recipients (Fig. 4D; Student's two-tailed t-test, $p<0.05$). Upon tertiary transfer, however, we began to see evidence of decreased fitness in some recipients of trisomy 19 bone marrow cells (Fig. 4E), even though HSC levels were similar in the bone marrow of trisomy 19 and wild type long-term reconstituted secondary recipients (Supplemental Fig. S3V). While there were no obvious peripheral blood defects when compared to wild type tertiary recipients (Supplemental Fig. S4A-K), weaker engraftment of trisomy 19 cells was evident in some tertiary recipients (Fig. 4E), and long-term survival of tertiary recipients was slightly decreased (Supplemental Fig. S4L). Trisomy 19 fitness defects became even more pronounced in quaternary transfers (Fig. 4F). This defect was not due to fewer HSCs being transferred. Trisomy 19 and wild-type long-term reconstituted tertiary recipients had comparable levels of HSCs (Supplemental Fig. S4M), but the majority of trisomy 19 quaternary recipient mice showed very low levels of engraftment compared to wild type quaternary recipients (Fig. 4F). Lethality was high in both wild type and trisomy 19 quaternary recipients (Supplemental Fig. S4O). Sequencing of the CD45.2 cells in two trisomy 19 quaternary recipients 14 weeks after transfer confirmed that the trisomic chromosome was retained through all serial transfers (Supplemental Fig. S4N). We conclude that trisomy 19 is better tolerated in the hematopoietic lineage than trisomy 16, with no difference in survival and little difference in peripheral blood cell counts in both primary and secondary recipient mice. However, decreased reconstitution potential in tertiary and quaternary recipients indicates that aneuploidy of even the smallest mouse autosome will eventually lead to reduced fitness.

Bub1bH/H HSCs undergo stem cell exhaustion upon serial transplantation

We next sought to evaluate the fitness of the *Bub1b*$^{H/H}$ CIN model, a model in which random combinations of whole chromosomal aneuploidies can be continuously generated by

chromosome mis-segregation. To evaluate the long-term regenerative potential of *Bub1b*$^{H/H}$ HSCs, we utilized bone marrow HSCs (BM-HSCs; Fig. 5A), because *Bub1b*$^{H/H}$ mice survive to adulthood (Baker et al. 2004) and show no obvious peripheral blood defects (Supplemental Fig. S5A-H). Like *Bub1b*$^{H/H}$ FL-HSCs, *Bub1b*$^{H/H}$ BM-HSCs showed no fitness defects in competitive reconstitution assays and were found at similar levels as their wild type littermates (Supplemental Fig. S5I-L). As BM-HSCs have already migrated from the fetal liver to populate the bone marrow of an adult mouse, we considered bone marrow transfer a more significant proliferative challenge than transfer of fetal liver cells. Also, FL-HSCs have been shown to reconstitute irradiated recipients more effectively than BM-HSCs (Harrison et al. 1997; Morrison et al. 1995; Ema and Nakauchi 2000). However, *Bub1b*$^{H/H}$ BM-HSCs contributed most of the peripheral white blood cells in primary recipient mice (Fig. 5B). The animals were mildly leukopenic but exhibited no other obvious blood defects (Supplemental Fig. S6A-I).

To further challenge *Bub1b*$^{H/H}$ HSCs, we performed serial bone marrow transfers (Fig. 5A). Despite similar total HSC levels in the bone marrow of long-term reconstituted *Bub1b*$^{H/H}$ and wild type primary recipients (Supplemental Fig. S6J), bone marrow cells from primary *Bub1b*$^{H/H}$ recipients repopulated the hematopoietic compartment of secondary recipients less efficiently than wild type primary recipient bone marrow cells (Fig. 5C). While the donor contributed most of the peripheral white blood cells in mice transferred with wild type cells (on average 87.3% at 12 weeks, n=10), variable contribution was observed in *Bub1b*$^{H/H}$ secondary recipients. Although complete blood cell counts of *Bub1b*$^{H/H}$ secondary recipients were similar to those of wild type secondary recipients (Supplemental Fig. S6K-S), some *Bub1b*$^{H/H}$ secondary recipient mice had strong donor contribution (~88% at 12 weeks, n=6), one recipient mouse had weak donor contribution (18.2% at 12 weeks) and some recipient mice had an intermediate level

of donor contribution (~70% at 12 weeks, n=3). Peripheral blood analyses revealed that the blood of the animal with weak contribution of *Bub1b*$^{H/H}$ HSCs was largely comprised of CD45.1 cells, presumably descendants from rare recipient HSCs that survived irradiation (Fig. 5C). These data indicate that *Bub1b*$^{H/H}$ BM-HSCs lose their regenerative potential with serial transfer. To test this further, we performed a tertiary transfer with bone marrow cells derived from secondary recipients with high, intermediate, and low peripheral blood contribution from *Bub1b*$^{H/H}$ HSCs. Despite similar total HSC levels in the bone marrow of long-term reconstituted *Bub1b*$^{H/H}$ and wild type secondary recipients (Supplemental Fig. S6T), mice that received bone marrow from secondary recipient donors with high *Bub1b*$^{H/H}$ peripheral blood contribution did not survive the tertiary transfer (0% survival after 4 weeks, n=10). Mice that received bone marrow from donors with intermediate or low *Bub1b*$^{H/H}$ peripheral blood contribution showed increased survival (60% and 80%, respectively after 16 weeks, n=5 for each condition). However, peripheral blood analyses of all mice that survived tertiary transfer revealed that virtually no cells originated from the original CD45.2 *Bub1b*$^{H/H}$ donor, but rather were derived from some surviving CD45.1 HSCs from the recipients (Fig. 5D). This is in contrast to the tertiary recipients that received bone marrow from secondary recipients reconstituted with wild type bone marrow, which had on average a 71.6% contribution from the wild type donor after 16 weeks (n=7; Fig. 5D). Thus, *Bub1b*$^{H/H}$ BM-HSCs lose the potential to reconstitute hematopoiesis upon serial transfer.

Bub1bH/H blood cells become progressively less aneuploid during hematopoietic reconstitution

Initial characterization of the *Bub1b*$^{H/H}$ mouse model hypothesized that the aging-associated phenotypes observed in this mouse were due to progressive accumulation of aneuploid cells in adult tissues (Baker et al. 2004). Thus, *Bub1b*$^{H/H}$ BM-HSCs could be losing the potential to reconstitute the hematopoietic system upon serial transfer because the *Bub1b*$^{H/H}$

HSC pool becomes progressively more aneuploid with each successive proliferative challenge, such that the level of aneuploidy in *Bub1b*$^{H/H}$ HSCs is too high to support effective proliferation. To test this possibility, we determined the karyotype of donor-derived peripheral white blood cells from one *Bub1b*$^{H/H}$ secondary recipient by single cell sequencing (57 weeks after transfer; 72% CD45.2 at time of sequencing). Surprisingly, all 17 cells sequenced were euploid (Fig. 6A; black triangle; Supplemental Fig. S7B). Thus, the failure of *Bub1b*$^{H/H}$ BM-HSCs to reconstitute a tertiary recipient was not simply due to increased aneuploidy in blood cells derived from *Bub1b*$^{H/H}$ HSCs. Rather, the absence of aneuploid descendants of *Bub1b*$^{H/H}$ BM-HSCs suggests that aneuploid cells derived from these HSCs cannot effectively contribute to the peripheral blood of secondary recipients and that aneuploid cells are selected against. This inability of *Bub1b*$^{H/H}$ aneuploid BM-HSCs to contribute to the peripheral blood makes them unable to reconstitute tertiary recipients.

The lack of aneuploidy observed in the peripheral blood of a secondary recipient mouse of *Bub1b*$^{H/H}$ BM-HSCs is in contrast to what we observed when we sequenced peripheral blood cells of a recipient mouse from a competitive reconstitution assay, in which 38.9% of *Bub1b*$^{H/H}$ FL-HSC-derived peripheral blood cells or 19.45% of the total peripheral blood was aneuploid (Fig. 1G, 6A; gray triangle, Supplemental Fig. S7A). We hypothesized that differences in either the source of the HSCs (fetal liver or adult bone marrow) or the time of sampling could be responsible for the different levels of aneuploidy observed in the two analyses. Single cell sequencing of *Bub1b*$^{H/H}$ FL-HSCs and BM-HSCs revealed that both of these cell types do not harbor any aneuploidies (n=19 cells each, Supplemental Fig. S7C, D), suggesting that the latter possibility was more likely responsible for the differences in aneuploidy observed. The sample from the competitive reconstitution assay that revealed high levels of aneuploidy was obtained

16 weeks after transfer of fetal liver donor cells, at a time during the reconstitution process when long-term hematopoiesis is just being established after a period a rapid expansion (Purton and Scadden 2007). In contrast, the sample from the secondary bone marrow recipient mouse that harbored euploid peripheral blood was obtained about a year after transfer of donor cells, when the bone marrow more closely resembles the steady state observed in an adult mouse.

The hypothesis that time after HSC transplantation can affect the degree of aneuploidy in the peripheral blood of recipient mice was informed by an evolutionary principle known as the "population flush" effect (Carson 1968). This principle states that in rapidly expanding populations—as occurs immediately following transfer of donor HSCs to a lethally irradiated recipient—purifying selection is relaxed such that less fit individuals can survive and significantly contribute to the population. A prediction of this hypothesis is that aneuploid *Bub1b*$^{H/H}$ peripheral blood cells would be more readily observed earlier during hematopoietic reconstitution, when the donor HSCs are rapidly proliferating to establish stable, long-term hematopoiesis and less fit cells—such as those generated by random chromosome mis-segregation in the *Bub1b*$^{H/H}$ adult mouse—would be tolerated. However, when populations reach steady state once long-term hematopoiesis has been established—as at 57 weeks after secondary transplantation—purifying selection forces become relatively stronger, selecting against less fit aneuploid cells. The observation that BM-HSCs are not aneuploid (0% aneuploidy) is consistent with this principle. BM-HSCs divide rarely, relying more on the division of progenitor cells to produce blood cells (Busch et al. 2015; Sun et al. 2015). Thus in steady state, BM-HSCs likely experience strong purifying selection. FL-HSCs proliferate at E14.5, but this cell population remains relatively small even after expansion (Morrison et al. 1995; Ema and Nakauchi 2000). Additionally, FL-HSCs must still home to their niches before adult hematopoiesis commences

(Orkin and Zon 2008), making it likely that this stem cell population is also under relatively strong purifying selection.

To directly test whether the population flush hypothesis can explain the degrees of aneuploidy observed in hematopoietic reconstitutions, we performed a time course, transferring *Bub1b*$^{H/H}$ fetal liver or bone marrow cells into lethally irradiated recipient mice and performing single cell sequencing of peripheral blood cells 3, 6, 13 and 36 weeks after transfer to determine the levels of autosomal aneuploidy at these times. Single cell sequencing of *Bub1b*$^{H/H}$ FL-HSCs and BM-HSCs, where no aneuploidy was observed, was used as the baseline level of aneuploidy in each population (0 time point in Fig. 6A; Supplemental Fig. S7C, D). In *Bub1b*$^{H/H}$ FL-HSC recipients, we found 8.3% of peripheral blood cells to be aneuploid 3 weeks post transplantation (n=12 cells; Fig. 6A, Supplemental Fig. S8A). This percentage increased to 24% aneuploid cells 6 weeks after transplantation (n=25 cells; Fig. 6A, Supplemental Fig. S8B). At 13 weeks, 25% of peripheral blood cells were aneuploid (n=16 cells; Fig. 6A, Supplemental Fig. S8C). However, 36 weeks after transplantation the proportion of aneuploid peripheral blood cells dropped to 8.3% (n=12 cells; Fig. 6A; Supplemental Fig. S8D). Similar results were obtained in *Bub1b*$^{H/H}$ BM-HSC recipients. We found 5.6% of peripheral blood cells to be aneuploid 3 weeks post transplantation (n=18 cells; Fig. 6A, Supplemental Fig. S8E). At 6 weeks, 24% of peripheral blood cells were aneuploid (n=25 cells; Fig. 6A, Supplemental Fig. S8F), and at 13 weeks 18.8% cells (n=16 cells; Fig. 6A, Supplemental Fig. S8G). However 34 weeks after transplantation, the proportion of aneuploid peripheral blood cells dropped to 15.4% (n=13; Fig. 6A; Supplemental Fig. S8H).

Taken together, our findings indicate that a higher level of aneuploidy is tolerated in the peripheral blood of *Bub1b*$^{H/H}$ recipient mice when the HSC pool is still expanding to establish

long term hematopoiesis after irradiation; however, these aneuploid cells are depleted from the peripheral blood of recipients once the HSC population is no longer rapidly expanding and the hematopoietic compartment reaches a steady state. Further, we conclude that HSC source, either fetal liver or adult bone marrow, does not affect the degree of aneuploidy observed in the peripheral blood.

Aneuploidy is selected against in Bub1bH/H regenerating tissues

The idea that rapidly expanding cell populations are more tolerant of aneuploidy than tissues that have reached a proliferative steady state further predicts that the prevalence of aneuploidy in adult *Bub1b*$^{H/H}$ tissues should depend on the proliferative capacity of tissues. Tissues that form during embryogenesis – when cells are rapidly proliferating – and that are largely non-proliferative in the adult should harbor higher levels of aneuploidy. In contrast, tissues that undergo self-renewal and are thus under more stringent purifying selection in the adult should harbor lower levels of aneuploidy. To test this prediction, we sequenced single cells from tissues that self renew in adult *Bub1b*$^{H/H}$ mice (peripheral blood cells, keratinocytes and intestinal crypt cells) and compared the observed degree of autosomal aneuploidy to that of tissues that divide primarily during embryogenesis (liver, brain). Liver and brain are formed during embryogenesis, and hepatocytes and neurons are largely non-proliferative in the adult (Zimmermann 2004; Campisi and d'Adda di Fagagna 2007). Our previous studies showed that aneuploid cells were prevalent in the liver (18.8%) and in the brain (38.1%) of 8-week old and 12-week old *Bub1b*$^{H/H}$ animals, respectively (Knouse et al. 2014). Sequencing 4 additional hepatocytes from a 30-week old *Bub1b*$^{H/H}$ mouse and 4 additional neurons from a 28-week old *Bub1b*$^{H/H}$ mouse estimated the levels of aneuploidy to be 14.3% and 40%, respectively (Fig. 6B, Supplemental Fig. S9D, E).

BM-HSCs and peripheral white blood cells both obtained from the same 19-week old *Bub1b*$^{H/H}$ mouse exhibited 0 and 4.3% aneuploidy, respectively (n=19 and 23 cells sequenced; Fig. 6B; Supplemental Fig. S7D, S9A). A previous study reported higher levels of aneuploidy in splenocytes of 3- and 5-month old *Bub1b*$^{H/H}$ animals (9% and 15%) (Baker et al. 2004). However, metaphase spreads, which tend to overestimate aneuploidy (Knouse et al. 2014), were employed to evaluate cell karyotype in this study. It is also possible that aneuploid blood cells have a higher survival rate in the spleen than in the bone marrow or peripheral blood. Keratinocytes and intestinal cells enriched for crypts also showed low levels of aneuploidy: 4.3% of keratinocytes isolated from a 15-week old *Bub1b*$^{H/H}$ mouse were aneuploid (n=23 cells sequenced; Fig. 6B; Supplemental Fig. S9B) and 4.8% aneuploidy was observed in intestinal crypt cells from an 18- and a 30-week old *Bub1b*$^{H/H}$ mouse (n=18 and 3 cells sequenced, respectively; Fig. 6B; Supplemental Fig. S9C). We conclude that tissues that regenerate during adulthood harbor lower levels of aneuploidy than tissues that proliferate predominantly during embryogenesis.

Direct comparison of degree of aneuploidy between adult tissues further revealed that adult non-proliferative tissues (brain, liver) are more tolerant of aneuploidy than self-renewing adult tissues (peripheral blood, BM-HSCs, skin, intestines) and blood cells derived from fetal liver or bone marrow transfers (Fig. 6C). Further analyses of specific chromosome gains and losses revealed that non-proliferative tissues (brain, liver) harbored cells with more complex aneuploidies than cells from *Bub1b*$^{H/H}$ HSC recipients (Fig. 6D; Supplemental Fig. S10). Of the 3 aneuploid cells observed in tissues that regenerate in the adult, 2 harbored a single chromosome gain and 1 harbored a single chromosome loss (Supplemental Fig. S9). Thus, we conclude that aneuploidy is selected against in tissues that regenerate in the adult.

Discussion

Here we describe a system that has permitted direct comparison of aneuploid mammalian cell fitness to the fitness of isogenic euploid cells *in vivo*. We found that constitutional trisomy and aneuploidy resulting from CIN negatively affect HSC fitness. Further, we have obtained evidence that aneuploidy is selected against in adult regenerating tissues. These findings greatly inform our understanding of the role of aneuploidy in cancer and aging.

The effects of aneuploidy and CIN on HSC fitness and hematopoiesis

The hematopoietic reconstitution system we developed here has allowed us to investigate whether or not aneuploidy provides a proliferative advantage to cells. Our findings demonstrate that constitutional trisomy and aneuploidy generated by CIN both adversely affect the fitness of HSCs *in vivo*. Defects specific to certain aneuploidies as well as defects observed in all aneuploid strains analyzed were evident in our aneuploid HSCs. The lineage-specific defects such as reduced B cell number observed in trisomy 16 FL-HSCs reconstitutions are likely due to chromosome-specific effects, as B cell differentiation defects have also been observed in DS (Lane et al. 2014; Roy et al. 2012). Further, we observed decreased proliferative potential in all aneuploidy models that we examined. Trisomy 16 FL-HSCs are much less fit than euploid HSCs and show phenotypes characteristic of proliferation defects when challenged to reconstitute irradiated euploid recipients. Trisomy 19 and *Bub1b*$^{H/H}$ HSCs exhibit less severe fitness defects, indicating that lower levels of aneuploidy are better tolerated in the hematopoietic lineage. Recent studies have shown that murine HSCs divide very infrequently and utilize numerous, more differentiated progenitor cells to sustain long-term hematopoiesis (Sun et al. 2015; Busch et al. 2015). This provides a potential explanation as to why repeated challenges are needed to reveal the fitness defects of aneuploidy in *Bub1b*$^{H/H}$ HSCs. In the *Bub1b*$^{H/H}$ HSC pool, only a subpopulation of cells are likely aneuploid and selection for euploid cells likely occurs, given

that we did not observe aneuploidy in *Bub1b*$^{H/H}$ HSCs. In the same vein, a proliferative defect in trisomy 19 HSCs began to emerge after tertiary bone marrow transfer and was obvious only after quaternary transfer. Thus, *in vivo* analyses of primary aneuploid cells demonstrate that aneuploidy in the hematopoietic lineage reduces cellular fitness and proliferative capacity. While some aneuploidies, such as trisomy 19, are better tolerated than others, such as trisomy 16, reconstitutions with the *Bub1b*$^{H/H}$ CIN model reveal that when euploid cells can be generated, aneuploidy is selected against in the hematopoietic lineage.

How aneuploidy leads to reduced fitness in HSCs remains to be determined. Cell culture studies of yeast and mammalian cells have shown that the proteomic imbalances caused by aneuploidy lead to proteotoxic stress, metabolic alterations, increased ROS production and cell cycle delays (Santaguida and Amon 2015). These general characteristics of aneuploid cells likely contribute to the decreased proliferative potential and eventual exhaustion of aneuploid HSCs that is observed.

While increased proliferation was not observed in transplantation of aneuploid HSCs alone, it will be interesting to determine whether particular aneuploidies can contribute to tumorigenesis in specific oncogenic contexts or in conjunction with aneuploidy-tolerating mutations. For example, aneuploidy has been shown to increase tumor burden and decrease survival in animals lacking the tumor suppressor p53 (Li et al. 2010). Loss of p53 could permit the survival of cells with abnormal karyotypes and/or DNA damage that ensues when chromosomes are mis-segregated (Crasta et al. 2012), allowing for the selection of malignant karyotypes. Particular aneuploidies in conjunction with other specific genomic alterations could also facilitate the development of specific cancers. For example, analyses of individuals with DS demonstrate that trisomy 21 is sufficient to bias differentiation in the human fetal liver, leading

to a relative expansion of myeloid cells (Chou et al. 2008; Tunstall-Pedoe et al. 2008). This bias can progress to a transient myeloproliferative disorder (TMD) (Gamis and Smith 2012), and eventually AML if a cooperating *GATA1* mutation is also present (Mundschau et al. 2003; Hitzler 2003). Furthermore, trisomy 21 is also observed in non-DS hematologic cancers (Mitelman et al. 1990; Cheng et al. 2009). The model system we have developed here will permit effective molecular dissection of the effect of specific chromosomal abnormalities, and the role of aneuploidy and CIN in general on the development of leukemias and lymphomas.

Aneuploidy is selected against in regenerating tissues

Our finding that aneuploidy impairs rather than promotes proliferation *in vivo* is supported by our tissue analyses in *Bub1b*$^{H/H}$ mice. Tissues that form largely during embryogenesis harbored high levels of aneuploid cells, whereas bone marrow-derived HSCs and adult peripheral blood, skin and intestinal crypt cells harbor few aneuploid cells. Importantly, we also found that FL-HSCs were euploid, suggesting that mechanisms that eliminate aneuploid cells are not unique to adult tissues. Rather, we favor the idea that the varying degree of aneuploidy observed in different tissues reflects differences in the strength of purifying selection across tissues. During rapid population expansion a "population flush" effect (Carson 1968) can occur, in which genetically less fit individuals are able to survive and contribute to the population in a more substantive manner. Once populations reach steady state, purifying selection forces are relatively stronger and select against the survival of less fit individuals. Thus, tissues that form during embryogenesis, a period of rapid cell expansion, and that are largely non-proliferative and are not maintained by stem cells in the adult (brain and liver) harbor high levels of aneuploidy whereas adult regenerating tissues (blood, skin and intestine) harbor few aneuploid cells. Our analysis of aneuploidy dynamics in *Bub1b*$^{H/H}$ HSC-derived blood cells during hematopoietic reconstitution further indicates that the way in which tissues are generated

and maintained determines the prevalence of aneuploidy in those tissues. Both BM- and FL-HSCs are euploid. Yet, when HSCs are challenged to rapidly expand to reconstitute an irradiated recipient, aneuploidy becomes more prevalent in the peripheral blood of primary recipients during early stages of hematopoietic reconstitution. Aneuploidy levels then decline as the tissue reaches steady state.

A previous evaluation of aneuploidy in aged mice that harbored increased expression of the BUBR1 protein in all tissues throughout life found that cells that overexpress BUBR1 and additionally all wild type aged adult stem cells show low levels of aneuploidy (Baker et al. 2012). We found that even in a mutant where the BUBR1 checkpoint protein is compromised, aneuploidy is selected against in tissues that are maintained by adult stem cells. Further investigation is needed to determine how aneuploid cells are culled in adult regenerating tissues. Previous cell culture studies have shown that chromosome mis-segregation leads to p53 activation (Thompson and Compton 2010; Li et al. 2010), raising the possibility that aneuploid cells are culled in tissues through apoptosis. We have not observed high levels of apoptosis in the skin or bone marrow of adult *Bub1b*$^{H/H}$ mice by standard assays (data not shown), suggesting that other mechanisms are responsible for eliminating aneuploid cells in adult regenerating tissues.

Aneuploidy and aging

Bub1b$^{H/H}$ mice have a decreased lifespan and show a number of progeroid phenotypes such as dwarfism, cataracts, loss of subcutaneous fat and impaired wound healing (Baker et al. 2004). It has been proposed that these early aging phenotypes are caused by the progressive accumulation of aneuploid cells in these animals. However, our findings that serial bone marrow transfer leads to *Bub1b*$^{H/H}$ HSC exhaustion and that adult *Bub1b*$^{H/H}$ regenerating tissues contain very few aneuploid cells suggest an alternative explanation. The decreased lifespan and early aging phenotypes exhibited by *Bub1b*$^{H/H}$ mice could be due to the exhaustion of adult stem cell

pools that must proliferate more than wild type stem cells to produce euploid cells. It is also possible that the strength of purifying selection in adult regenerative tissues is relaxed as organisms age and stem cell fitness declines (Rozhok and DeGregori 2015). Such decreased purifying selection would cause tissue function decline due to the accumulation of aneuploid and hence less fit cells with time. Further investigation is needed to determine whether decreased purifying selection can also cause the accumulation of cells that are genetically unstable and karyotypically abnormal (Sheltzer et al. 2011; Zhu et al. 2012; Nicholson et al. 2015), as such a population from which rare cells with high proliferative and hence tumorigenic potential may arise could cause cancer as individuals age.

Materials and Methods

Mouse strains

The mouse strains utilized previously (Williams et al. 2008) were backcrossed for at least 10 generations into the C57BL/6J background (Jackson Laboratory) to generate congenic strains. Strains used to generate trisomic embryos were B6.Cg-Rb(6.16)24Lub/JAmonJ or B6.Cg-Rb(13.16)1Mpl/JAmonJ and B6.Cg-Rb(16.17)7Bnr/JAmonJ (Trisomy 16) and B6.Cg-Rb(5.19)1Wh/JAmonJ and B6Ei.Cg-Rb(9.19)163H/J (Trisomy 19). All male compound heterozygous mice were mated with C57BL/6J females to generate trisomic embryos, although at a lower frequency than previously reported most likely due to backcrossing (Pfau and Amon 2016). *Bub1b*$^{H/H}$ mice were a generous gift from Dr. J.M. van Deursen. Embryos from all strains were collected at embryonic days E14.5 – E15.5 by timed matings. Recipient mice were 6-8 week old B6.SJL-*Ptprc*^{a}*Pepc*^{b}/BoyJ (CD45.1) female mice from Jackson Laboratory. Fetal livers were prepared and karyotyped as described in Supplemental Methods. All animal studies and procedures were approved by the MIT Institutional Animal Care and Use Committee.

Hematopoietic reconstitutions

For all reconstitution experiments, mice were closely monitored for signs of bone marrow failure and rapid weight loss. All protocols for treating irradiated mice were outlined by MIT's Division of Comparative Medicine. Whole body irradiation was performed using a 137^{Cs} irradiator (γ cell 40) at a dose rate of ~ 100 cGy/min.

For competitive reconstitution assays and hematopoietic reconstitution experiments, recipient mice were irradiated with a 12 Gy total dose administered as a split dose of 8 Gy followed by a second dose of 4 Gy 3 hours later. Fetal liver cells were thawed in Iscove's modified Dulbecco's medium (IMDM) supplemented with 2% FBS and counted on a Cellometer Auto T4 automated hemacytometer (Nexelcom). Viability was assessed by propidium iodide exclusion using a FACSCalibur flow cytometer (Becton Dickinson). 10^6 live cells were injected intravenously in Hank's Balanced Salt Solution (HBSS). In one replicate of the trisomy 16 competition assay (n=5 trisomy 16, n=7 wild type), CD45.2/CD45.1 aneuploid donors and CD45.2 common WT donors were transferred into CD45.1 irradiated recipient mice.

Recipient mice for colony forming unit spleen (CFU-S) assays were irradiated with 9.5 Gy administered as a single dose. Fetal liver cells were then prepared and 10^6 live cells were injected as described above. 7-8 days after injection, spleens were harvested from recipients and then fixed overnight in Bouin's fixative. Spleens were sectioned into 5 μm slices and stained with hematoxylin and eosin. Slides were then scanned on a Leica Aperio slide scanner, and colony area was measured using ImageJ. To control for small residual white blood cell nodules in the spleen, sections of spleens from irradiated mice that were not transferred with cells were also analyzed. We determined the average background colony size to be 0.336% total spleen area, and colonies larger than this average were considered CFU-S colonies.

Homing assays were performed as described (Gilbert and Hemann 2012) and are detailed in the Supplemental Methods.

For EdU incorporation assays, recipient mice were irradiated with 9.5 Gy, administered as a single dose. Fetal liver cells were thawed, and viability was determined by PI exclusion as described above. 10^6 live fetal liver cells were injected intravenously in HBSS. 6 days later, mice were injected intraperitoneally with 1.25 mg EdU in PBS. 20 hours after EdU injection, recipient mice were euthanized. Bone marrow was harvested by flushing the long bones 24 hours after injection. Red blood cells were lysed in ACK Lysing Buffer, then samples were washed in IMDM containing 2% FBS. EdU incorporation was detected using the Click-iT plus EdU Alexa Fluor 488 flow cytometry kit (Thermo-Fisher) on an LSR II flow cytometer (Becton Dickinson). EdU-positive recipient-derived bone marrow cells were excluded using a CD45.1 antibody conjugated to PE (Biolegend; clone A20).

Recipient mice for bone marrow transfers were irradiated with 9.5 Gy, administered as a single dose. Bone marrow cells were isolated and red blood cells were lysed as described above. White blood cells were then counted, and recipient mice were reconstituted by intravenous injection of 10^6 cells in HBSS.

Peripheral blood analysis

For complete blood cell counts, peripheral blood was collected with heparinized capillary tubes into EDTA-coated Microvette 100 tubes (Sarstedt) and analyzed on a HemaVet 950FS (Drew Scientific). Peripheral blood for flow cytometry analysis was collected with heparinized capillary tubes into sodium heparin diluted in PBS. Red blood cells were lysed in ACK Lysing Buffer, then washed in HBSS containing 2% FBS. Cells were then incubated with antibodies according to the manufacturer's specifications, and analyzed with a FACSCalibur or LSR II flow cytometer

(Becton Dickinson). The following antibodies were used and obtained from BioLegend: CD45.1 (A20), CD45.2 (104), CD45R/B220 (RA3-6B2), CD90.2/Thy-1.2 (53-2.1).

HSC quantification

Fetal livers and bone marrow cells were harvested, genotyped or karyotyped, processed and counted as described above for hematopoietic reconstitutions. FL-HSCs were quantified using markers described previously (Kim et al. 2006)**.** BM-HSCs were quantified as described previously (Kiel et al. 2005). More detailed procedures are included in the Supplemental Methods. The following antibodies were used and obtained from BioLegend: CD150 (TC15-12F12.2), CD48 (HM48-1), Sca-1 (E13-161.7), CD117 (2B8). The following antibodies were used in a lineage cocktail for fetal liver cells and were obtained from BioLegend: CD3ε (145-2C11), Ly-6G/Ly-6C (Gr-1) (RB6-8C5), B220/CD45R (RA3-6B2), Ter-119, CD5 (53-7.3), CD8a (53-6.7).

Single cell sequencing and trisomy 19 sequencing

Single cell sequencing was performed and analyzed as described in (Knouse et al., 2014). Single cell isolation protocols for each cell type are described in detail in the Supplemental Methods. CD45.2 trisomy 19 peripheral blood was prepared as above for other blood cells and sorted on a FACSAria (Becton Dickinson). Genomic DNA was isolated from collected cells using the DNeasy Blood and Tissue kit (Qiagen), amplified with GenomePlex Complete WGA Kit (Sigma-Aldrich), and sequenced and analyzed in the same way as amplified single cells. The sequences reported in this paper have been deposited in the National Center for Biotechnology Information Sequence Read Archive (accession no. TBD).

Acknowledgments

We would like to thank the Swanson Biotechnology Center at the Koch Institute for Integrative Cancer Research; Paul Danielian and Christian Pallasch for advice and technical assistance; the MIT Division of Comparative Medicine diagnostic lab for technical assistance; Jie Wu for help

with bioinformatics analysis; Jan van Deursen for generously providing the *Bub1b*$^{H/H}$ mouse strain; David Housman, Michael Hemann, James DeGregori, Andrii Rozhok, Jianzhu Chen and Ömer Yilmaz for discussions and advice; and Jianzhu Chen, Ömer Yilmaz, and members of the Amon lab for critical reading of this manuscript. This work was supported by the National Institutes of Health (CA206157) to A.A., the Kathy and the Curt Marble Cancer Research Fund and in part by the Koch Institute Support Grant P30-CA14051. S.J.P. is supported by an NSF Graduate Research Fellowship, and K.A.K. is supported by the NIGMS Training grant T32GM007753. A.A is an investigator of the Howard Hughes Medical Institute and the Glenn Foundation for Biomedical Research. The authors declare no conflict of interest.

Figures

Figure 1

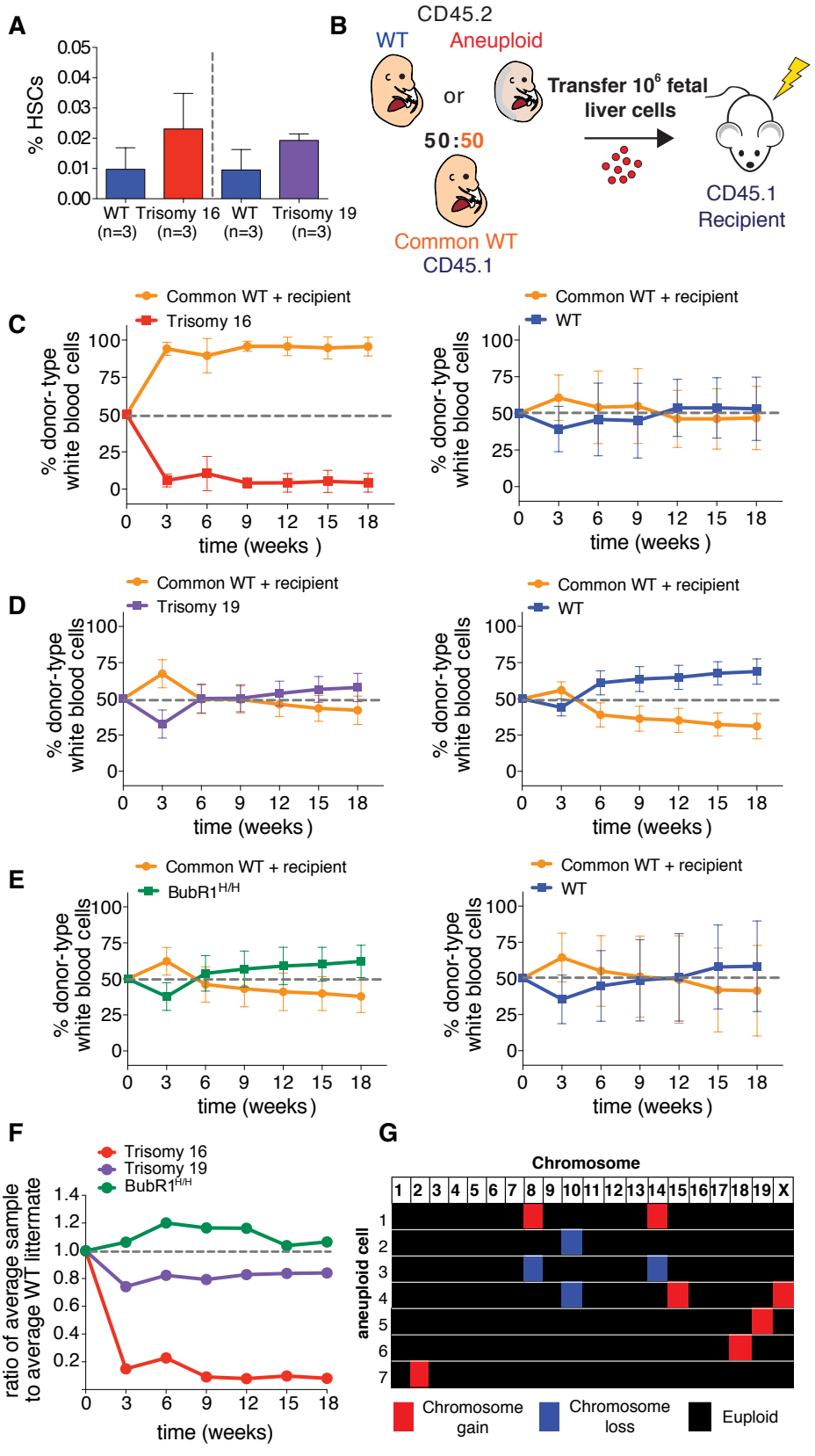


Figure 1. *Aneuploidy decreases HSC competitive fitness in vivo*
(A) The percentage of HSCs ($CD150^+$ $CD48^-$ $Sca\text{-}1^+$ lin^- cells) found in trisomy 16 and 19 fetal livers was quantified by flow cytometry. Data are shown as mean ± standard deviation. (B) Schematic of competitive reconstitution experiments. (C–E) CD42.5 fetal liver cells from wild type or aneuploid E14.5 littermates were co-injected into a lethally irradiated CD45.1 recipient with an equal number of fetal liver cells from a CD45.1 common WT donor of the same age derived from a separate mating. Peripheral blood was sampled at indicated times. The percentage of the white blood cell population contributed by each donor was quantified by flow cytometry with isoform-specific antibodies against CD45.1 and CD45.2 for recipients of common WT cells and trisomy 16 fetal liver cells (C, left graph, n=17), trisomy 19 fetal liver cells (D, left graph, n=10) and $Bub1b^{H/H}$ fetal liver cells (E, left graph, n=10). The contribution of WT littermates when competed to the common WT for all aneuploidies was quantified at the same time (C–E; right graphs). Data are represented as mean ± standard deviation for each time point. (F) Ratios of the average percentage of the peripheral blood reconstituted by the aneuploid fetal liver cells to the average percentage of the peripheral blood reconstituted by wild type littermate fetal liver cells at the indicated times are shown. (G) Single cell sequencing of white blood cells from a mouse competitively reconstituted with CD45.2 $Bub1b^{H/H}$ and CD45.1 euploid FL-HSCs at 16 weeks after transplantation (Fig. 1E) revealed that 7 of 18 CD45.2 $Bub1b^{H/H}$ cells analyzed (~39%) are aneuploid. Karyotypes of the 7 aneuploid cells are shown with chromosome gains in red, chromosome losses in blue, and euploidy shown in black. Segmentation plots of all sequenced cells are shown in Supplemental Fig. S7A.

Figure 2

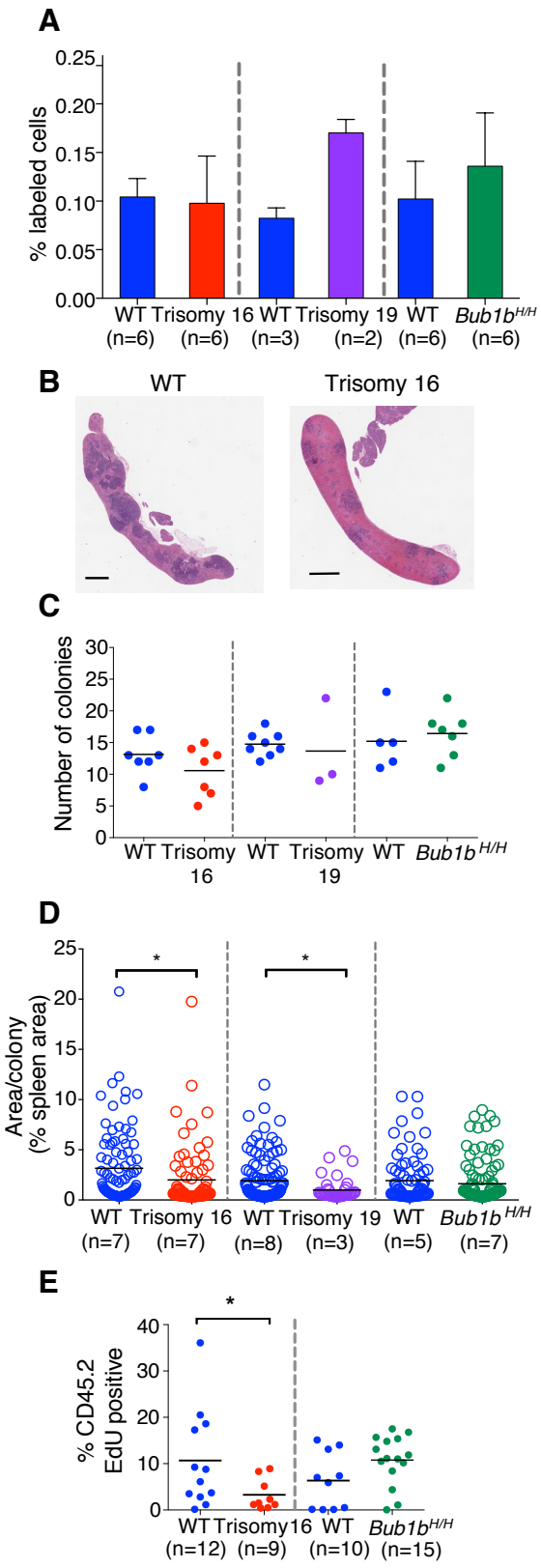

A
% labeled cells
0.25
0.20
0.15
0.10
0.05
0.00
WT (n=6)
Trisomy 16 (n=6)
WT (n=3)
Trisomy 19 (n=2)
WT (n=6)
Bub1b H/H (n=6)
B
WT
Trisomy 16
C
Number of colonies
30
25
20
15
10
5
0
WT
Trisomy 16
WT
Trisomy 19
WT
Bub1b H/H
D
Area/colony (% spleen area)
25
20
15
10
5
0
*
*
WT (n=7)
Trisomy 16 (n=7)
WT (n=8)
Trisomy 19 (n=3)
WT (n=5)
Bub1b H/H (n=7)
E
% CD45.2 EdU positive
40
30
20
10
0
*
WT (n=12)
Trisomy16 (n=9)
WT (n=10)
Bub1b H/H (n=15)

Figure 2. *Proliferation but not homing ability is reduced in trisomy 16 and trisomy 19 reconstitutions*
(A) DiI-labeled fetal liver cells were injected into irradiated recipient mice. The percentage of DiI-positive cells in the bone marrow of recipient mice was measured 24 hours after injection. Data are shown as mean ± standard deviation. (B) Representative images of sections of spleens isolated from mice transferred with WT or trisomy 16 fetal liver cells 8 days after reconstitution. Scale bar represents 1 mm. (C) Quantification of CFU-S colonies from spleen sections of recipient mice of trisomy 16 or wild type littermate fetal liver cells 8 days after injection and trisomy 19, *Bub1b*$^{H/H}$ or wild type littermate fetal liver cells 7 days after injection. The bar represents the mean value for each condition. (D) Quantification of the mean size of each colony in Fig. 2B as determined by percent of total spleen area. Measurements from all individuals for each condition were pooled. The bar represents the mean value for each population. Populations were compared by Student's t-test; * indicates $p<0.05$. (E) Trisomy 16 or *Bub1b*$^{H/H}$ fetal liver cells or cells from their wild type littermates were injected into lethally irradiated recipients. Mice were injected with EdU 6 days later, and the level of EdU incorporation in CD45.2 positive donor-derived bone marrow cells was evaluated by flow cytometry 24 hours later. Data are shown as mean ± standard deviation. * indicates $p<0.05$ by Student's t-test.

Figure 3

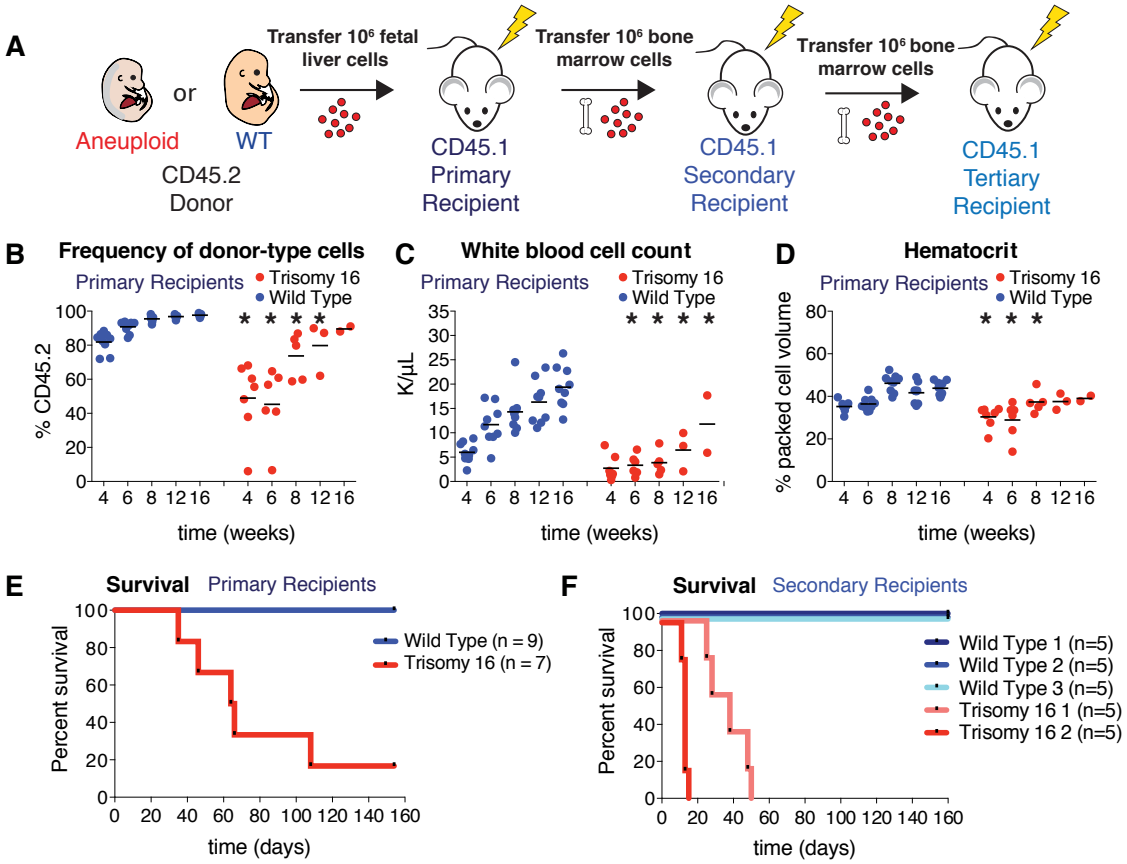


Figure 3. *Trisomy 16 causes peripheral blood defects and decreases HSC reconstitution potential*

(A) Schematic of serial reconstitution experiments. (B–D) For primary reconstitutions, fetal liver cells from a CD45.2 trisomic embryo or its wild type littermate were injected into lethally irradiated CD45.1 recipients. Bone marrow cells from primary recipients were injected into secondary CD45.1 recipients to assess serial reconstitution capacity. Peripheral blood of primary recipients of trisomic fetal liver cells or their wild type littermates was sampled at the indicated times. The percentage of CD45.2 positive cells in the blood of trisomy 16 or wild type primary recipients was determined by flow cytometry (B) and white blood cell count (C) and hematocrit (D) were determined by complete blood cell counts. The bar represents the mean, and asterisks indicate the trisomy 16 values are significantly different from the values of wild type littermates at the indicated time by Student's t-test ($p<0.05$). (E) Survival of recipients of trisomy 16 or wild type fetal liver cells after transfer. (F) Survival of secondary recipients of trisomy 16 or wild type bone marrow cells from primary recipients.

Figure 4

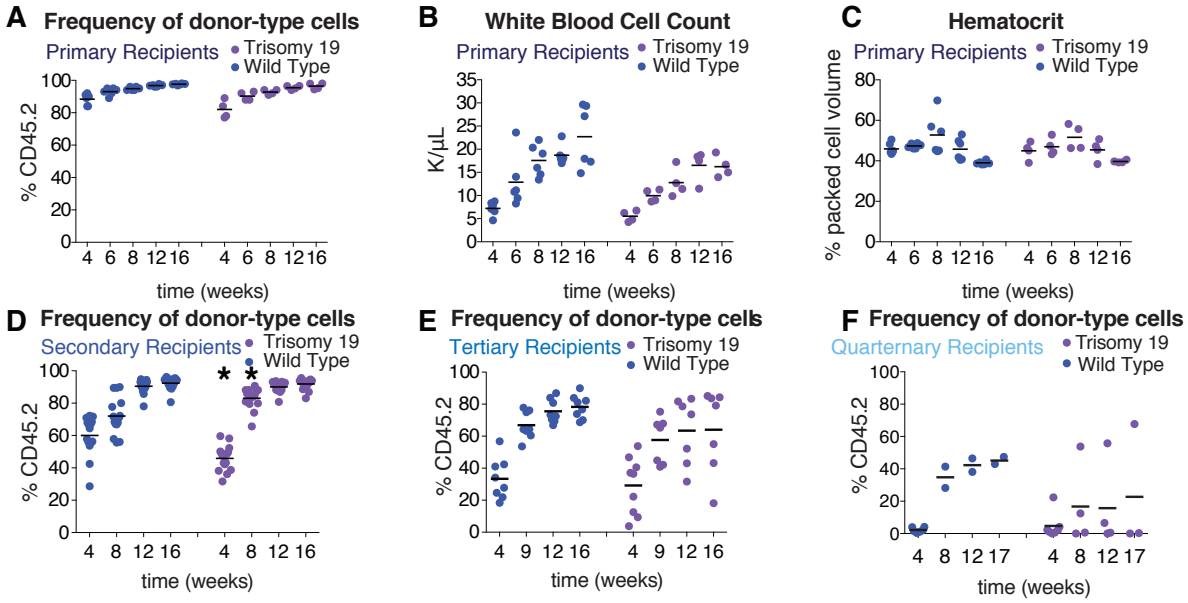


Figure 4. *Trisomy 19 HSC reconstitution potential is reduced upon serial reconstitution*
(A–E) For trisomy 19 or wild type primary recipients, the percentage of CD45.2 positive cells in the peripheral blood (A), white blood cell count (B) and hematocrit (C) was determined. The percentage of CD45.2 positive cells in the blood of trisomy 19 or wild type secondary recipients (D), tertiary recipients (E) and quaternary recipients (F) was also evaluated. The bar represents the mean value for each condition, and asterisks indicate the trisomy 19 values are significantly different from the values of wild type littermates at the indicated time by Student's t-test ($p<0.05$).

Figure 5

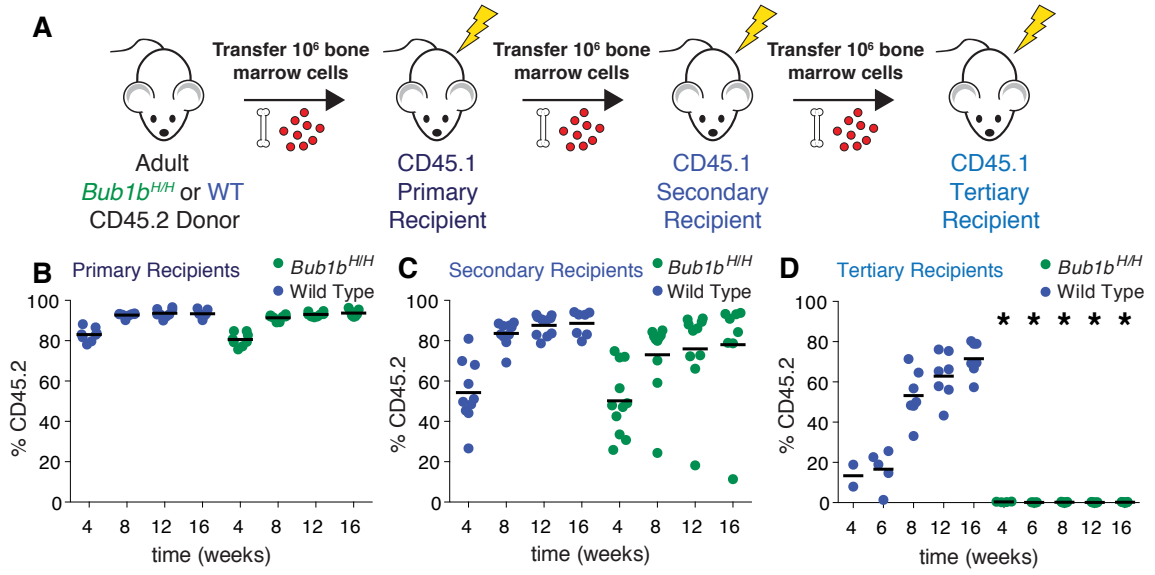


Figure 5. *Bub1b$^{H/H}$ BM-HSCs undergo stem cell exhaustion upon serial reconstitution*
(A) Schematic of serial reconstitution experiments. (B–D) CD45.2 adult *Bub1b$^{H/H}$* bone marrow cells were serially transplanted into lethally irradiated CD45.1 recipients. The percentage of the peripheral blood that is CD45.2 positive was determined in primary (B), secondary (C) and tertiary (D) recipients at the indicated times. The bar represents the mean value for each condition. Asterisks indicate the *Bub1b$^{H/H}$* values are significantly different from the values of wild type at the indicated time by Student's t-test ($p<0.05$).

Figure 6

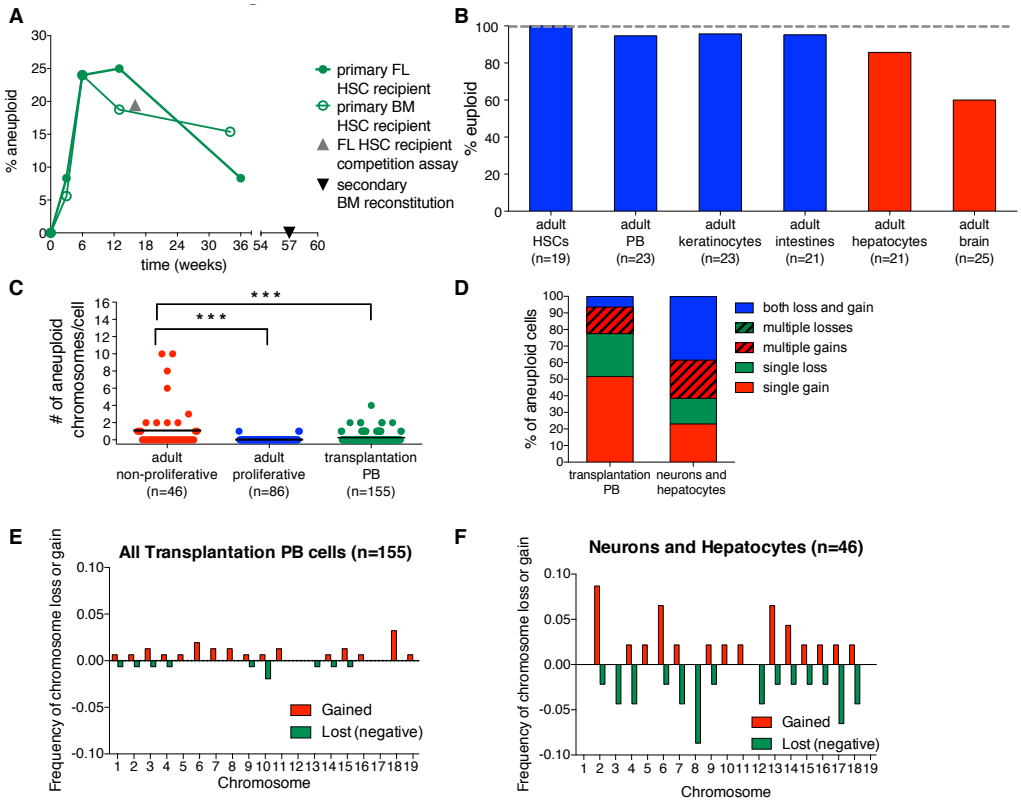


Figure 6. *Bub1b$^{H/H}$ adult regenerative tissues show evidence of selection against aneuploid cells* (A) Percent aneuploidy over time during hematopoietic reconstitution with *Bub1b$^{H/H}$* HSCs was determined by single cell sequencing of peripheral blood cells derived from primary recipient mice of *Bub1b$^{H/H}$* bone marrow (open circles) or *Bub1b$^{H/H}$* fetal liver (closed circles) cells at the indicated times after transfer. Peripheral blood cells from a mouse reconstituted with *Bub1b$^{H/H}$* fetal liver HSCs 16 weeks after competitive reconstitution (shown as percent of total peripheral blood; gray triangle; Fig. 1G), and from a *Bub1b$^{H/H}$* secondary bone marrow recipient mouse (black triangle; Fig. 5C) were also sequenced. Baseline aneuploidy was determined by single cell sequencing of FL- and BM-HSCs. Segmentation plots of all sequenced cells are shown in Supplemental Fig. S7 and S8. (B) Percentage of euploid cells found in different adult *Bub1b$^{H/H}$* cell types. BM-HSCs, peripheral blood cells (PB), keratinocytes and intestines (in blue) are from ~4-month old *Bub1b$^{H/H}$* mice. Data from hepatocytes and brain (in red) are from (Knouse et al. 2014) and from ~6-month old *Bub1b$^{H/H}$* mice. Segmentation plots of all newly sequenced cells are shown in Supplemental Fig. S9. (C) The number of aneuploid chromosomes per cell in all adult cells analyzed in (A-B). The bar represents the mean value for each population. t-tests were performed for significance. *** indicates $p<0.001$. (D) Summary of chromosome gain and loss events observed in each cell from transplantation peripheral blood cells and adult non-proliferative neurons and hepatocytes. Multiple gains or losses describe cells that have gained or lost 2 or more chromosomes. Both loss and gain describes cells that have gained at least one chromosome and lost at least one chromosome. (E-F) Frequency of chromosome gain (red) or chromosome loss (green) by chromosome observed in all peripheral blood cells after transplantation (E) and in all adult neurons and hepatocytes (F).

References

Adorno M, Sikandar S, Mitra SS, Kuo A, Di Robilant BN, Haro-Acosta V, Ouadah Y, Quarta M, Rodriguez J, Qian D, et al. 2013. Usp16 contributes to somatic stem-cell defects in Down's syndrome. *Nature* 1–17.

Babu J, Jeganathan K, Baker D, Wu X, Kang-Decker N, van Deursen J. 2003. Rae1 is an essential mitotic checkpoint regulator that cooperates with Bub3 to prevent chromosome missegregation. *J Cell Biol* **160**: 341–353.

Baker DJ, Dawlaty MM, Wijshake T, Jeganathan KB, Malureanu L, van Ree JH, Crespo-Diaz R, Reyes S, Seaburg L, Shapiro V, et al. 2012. Increased expression of BubR1 protects against aneuploidy and cancer and extends healthy lifespan. *Nature Cell Biology* **14**: 1–9.

Baker DJ, Jeganathan KB, Cameron JD, Thompson M, Juneja S, Kopecka A, Kumar R, Jenkins RB, de Groen PC, Roche P, et al. 2004. BubR1 insufficiency causes early onset of aging-associated phenotypes and infertility in mice. *Nature Genetics* **36**: 744–749.

Boveri T. 1914. *Zur Frage der Entstehung Maligner Tumoren*. Gustav Fischer, Jena.

Busch K, Klapproth K, Barile M, Flossdorf M, Holland-Letz T, Schlenner SM, Reth M, Höfer T, Rodewald H-R. 2015. Fundamental properties of unperturbed haematopoiesis from stem cells in vivo. *Nature* **518**: 542–546.

Campisi J, d'Adda di Fagagna F. 2007. Cellular senescence: when bad things happen to good cells. *Nat Rev Mol Cell Biol* **8**: 729–740.

Carson HL. 1968. The population flush and its genetic consequences. In *Population biology and evolution* (ed. R.C. Lewontin), pp. 123–137, Syracuse University Press, Syracuse, NY.

Cheng Y, Wang H, Wang H, Chen Z, Jin J. 2009. Trisomy 21 in patients with acute leukemia. *Am J Hematol* **84**: 193–194.

Choi JK. 2008. Hematopoietic disorders in down syndrome. *International journal of clinical and experimental pathology* **1**: 387–395.

Chou ST, Opalinska JB, Yao Y, Fernandes MA, Kalota A, Brooks JSJ, Choi JK, Gewirtz AM, Danet-Desnoyers G-A, Nemiroff RL, et al. 2008. Trisomy 21 enhances human fetal erythro-megakaryocytic development. *Blood* **112**: 4503–4506.

Crasta K, Ganem NJ, Dagher R, Lantermann AB, Ivanova EV, Pan Y, Nezi L, Protopopov A, Chowdhury D, Pellman D. 2012. DNA breaks and chromosome pulverization from errors in mitosis. *Nature* **482**: 53–58.

Davoli T, Xu AW, Mengwasser KE, Sack LM, Yoon JC, Park PJ, Elledge SJ. 2013. Cumulative Haploinsufficiency and Triplosensitivity Drive Aneuploidy Patterns and Shape the Cancer Genome. *Cell* 1–15.

Ema H, Nakauchi H. 2000. Expansion of hematopoietic stem cells in the developing liver of a mouse embryo. *Blood* **95**: 2284–2288.

Frasca D, Guidi F, Arbitrio M, Pioli C, Poccia F, Cicconi R, Doria G. 2000. Hematopoietic reconstitution after lethal irradiation and bone marrow transplantation: effects of different hematopoietic cytokines on the recovery of thymus, spleen and blood cells. *Bone Marrow Transplant* **25**: 427–433.

Gamis AS, Smith FO. 2012. Transient myeloproliferative disorder in children with Down syndrome: clarity to this enigmatic disorder. *Br J Haematol* **159**: 277–287.

Gilbert LA, Hemann MT. 2012. Context-specific roles for paracrine IL-6 in lymphomagenesis. *Genes Dev* **26**: 1758–1768.

Gogendeau D, Siudeja K, Gambarotto D, Pennetier C, Bardin AJ, Basto R. 2015. Aneuploidy causes premature differentiation of neural and intestinal stem cells. *Nature Communications* **6**: 1–15.

Gropp A, Winking H, Herbst EW, Claussen CP. 1983. Murine trisomy: developmental profiles of the embryo, and isolation of trisomic cellular systems. *Journal of Experimental Zoology* **228**: 253–269.

Harrison DE, Astle CM. 1982. Loss of stem cell repopulating ability upon transplantation. Effects of donor age, cell number, and

transplantation procedure. *J Exp Med* **156**: 1767–1779.

Harrison DE, Zhong RK, Jordan CT, Lemischka IR, Astle CM. 1997. Relative to adult marrow, fetal liver repopulates nearly five times more effectively long-term than short-term. *Exp Hematol* **25**: 293–297.

Henry E, Walker D, Wiedmeier SE, Christensen RD. 2007. Hematological abnormalities during the first week of life among neonates with Down syndrome: Data from a multihospital healthcare system. *Am J Med Genet* **143A**: 42–50.

Herbst EW, Winking H. 1991. Adoptive transfer of the hematopoietic system of trisomic mice with limited life span: stem cells from six different trisomies are capable of survival. *Dev Genet* **12**: 415–422.

Hitzler JK. 2003. GATA1 mutations in transient leukemia and acute megakaryoblastic leukemia of Down syndrome. *Blood* **101**: 4301–4304.

Jeganathan K, Malureanu L, Baker DJ, Abraham SC, van Deursen JM. 2007. Bub1 mediates cell death in response to chromosome missegregation and acts to suppress spontaneous tumorigenesis. *J Cell Biol* **179**: 255–267.

Kiel MJ, Yilmaz OH, Iwashita T, Yilmaz OH, Terhorst C, Morrison SJ. 2005. SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell* **121**: 1109–1121.

Kiladjian J-J. 2012. The spectrum of JAK2-positive myeloproliferative neoplasms. *Hematology* **2012**: 561–566.

Kim I, He S, Yilmaz OH, Kiel MJ, Morrison SJ. 2006. Enhanced purification of fetal liver hematopoietic stem cells using SLAM family receptors. *Blood* **108**: 737–744.

Knouse KA, Wu J, Whittaker CA, Amon A. 2014. Single cell sequencing reveals low levels of aneuploidy across mammalian tissues. *Proceedings of the National Academy of Sciences* **111**: 13409–13414.

Kondo M, Wagers AJ, Manz MG, Prohaska SS, Scherer DC, Beilhack GF, Shizuru JA, Weissman IL. 2003. Biology of hematopoietic stem cells and progenitors: implications for clinical application. *Annu Rev Immunol* **21**: 759–806.

Lane AA, Chapuy B, Lin CY, Tivey T, Li H, Townsend EC, van Bodegom D, Day TA, Wu S-C, Liu H, et al. 2014. Triplication of a 21q22 region contributes to B cell transformation through HMGN1 overexpression and loss of histone H3 Lys27 trimethylation. *Nature Genetics* **46**: 618–623.

Lapidot T. 2005. How do stem cells find their way home? *Blood* **106**: 1901–1910.

Li M, Fang X, Baker DJ, Guo L, Gao X, Wei Z, Han S, van Deursen JM, Zhang P. 2010. The ATM-p53 pathway suppresses aneuploidy-induced tumorigenesis. *Proc Natl Acad Sci USA* **107**: 14188–14193.

Mitelman F, Heim S, Mandahl N. 1990. Trisomy-21 in Neoplastic-Cells. *Am J Med Genet* **7**: 262–266.

Morrison SJ, Hemmati HD, Wandycz AM, Weissman IL. 1995. The purification and characterization of fetal liver hematopoietic stem cells. *Proc Natl Acad Sci USA* **92**: 10302–10306.

Mundschau G, Gurbuxani S, Gamis AS, Greene ME, Arceci RJ, Crispino JD. 2003. Mutagenesis of GATA1 is an initiating event in Down syndrome leukemogenesis. *Blood* **101**: 4298–4300.

Nicholson JM, Macedo JC, Mattingly AJ, Wangsa D, Camps J, Lima V, Gomes AM, Dória S, Ried T, Logarinho E, et al. 2015. Chromosome mis-segregation and cytokinesis failure in trisomic human cells. *eLife* **4**.

Orkin SH, Zon LI. 2008. Hematopoiesis: an evolving paradigm for stem cell biology. *Cell* **132**: 631–644.

Pfau SJ, Amon A. 2016. A System to Study Aneuploidy In Vivo. *Cold Spring Harbor Symposia on Quantitative Biology* **80**: 027193.

Pfau SJ, Amon AA. 2012. Chromosomal instability and aneuploidy in cancer: from yeast to man. *EMBO Rep* **13**: 515–527.

Purton LE, Scadden DT. 2007. Limiting Factors in Murine Hematopoietic Stem Cell Assays. *Cell Stem Cell* **1**: 8–8.

Rajagopalan H, Lengauer C. 2004. Aneuploidy and cancer. *Nature* **432**: 338–341.

Roy A, Cowan G, Mead AJ, Filippi S, Bohn G, Chaidos A, Tunstall O, Chan JKY, Choolani M, Bennett P, et al. 2012. Perturbation of fetal liver hematopoietic stem and progenitor cell development by trisomy 21. *Proceedings of the National Academy of Sciences* **109**: 17579–17584.

Rozhok AI, DeGregori J. 2015. Toward an evolutionary model of cancer: Considering the mechanisms that govern the fate of somatic mutations. *Proceedings of the National Academy of Sciences* **112**: 8914–8921.

Santaguida S, Amon A. 2015. Short- and long-term effects of chromosome mis-segregation and aneuploidy. *Nat Rev Mol Cell Biol* **16**: 473–485.

Satge D, Sommelet D, Geneix A, Nishi M, Malet P, Vekemans M. 1998. A tumor profile in Down syndrome. *Am J Med Genet* **78**: 207–216.

Segal DJ, McCoy EE. 1974. Studies on Down's syndrome in tissue culture. I. Growth rates protein contents of fibroblast cultures. *J Cell Physiol* **83**: 85–90.

Sheltzer JM, Blank HM, Pfau SJ, Tange Y, George BM, Humpton TJ, Brito IL, Hiraoka Y, Niwa O, Amon A. 2011. Aneuploidy drives genomic instability in yeast. *Science* **333**: 1026–1030.

Stingele S, Stoehr G, Peplowska K, Cox JUR, Mann M, Storchova Z. 2012. Global analysis of genome, transcriptome and proteome reveals the response to aneuploidy in human cells. *Mol Syst Biol* **8**: 1–12.

Sun J, Ramos A, Chapman B, Johnnidis JB, Le L, Ho Y-J, Klein A, Hofmann O, Camargo FD. 2015. Clonal dynamics of native haematopoiesis. *Nature* **514**: 322–327.

Thompson SL, Compton DA. 2010. Proliferation of aneuploid human cells is limited by a p53-dependent mechanism. *J Cell Biol* **188**: 369–381.

Till JE, McCulloch EA. 1961. A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. *Radiation Research* **14**: 213–222.

Tunstall-Pedoe O, Roy A, Karadimitris A, la Fuente de J, Fisk NM, Bennett P, Norton A, Vyas P, Roberts I. 2008. Abnormalities in the myeloid progenitor compartment in Down syndrome fetal liver precede acquisition of GATA1 mutations. *Blood* **112**: 4507–4511.

Waterstrat A, Liang Y, Swiderski CF, Shelton BJ, Van Zant G. 2010. Congenic interval of CD45/Ly-5 congenic mice contains multiple genes that may influence hematopoietic stem cell engraftment. *Blood* **115**: 408–417.

Weaver BAA, Silk AD, Montagna C, Verdier-Pinard P, Cleveland DW. 2007. Aneuploidy Acts Both Oncogenically and as a Tumor Suppressor. *Cancer Cell* **11**: 25–36.

Williams BR, Prabhu VR, Hunter KE, Glazier CM, Whittaker CA, Housman DE, Amon A. 2008. Aneuploidy Affects Proliferation and Spontaneous Immortalization in Mammalian Cells. *Science* **322**: 703–709.

Zhu J, Pavelka N, Bradford WD, Rancati G, Li R. 2012. Karyotypic Determinants of Chromosome Instability in Aneuploid Budding Yeast ed. B.A. Sullivan. *PLoS Genetics* **8**: e1002719.

Zimmermann A. 2004. Regulation of liver regeneration. *Nephrol Dial Transplant* **19 Suppl 4**: iv6–i10.

Supplemental Information

Supplemental Materials and Methods

Sample preparation and karyotyping

Fetal livers were removed from embryos by dissection. Samples were homogenized by pipetting and passage through a 70 μm cell strainer, frozen in FBS + 5% dimethylsulfoxide and stored in liquid nitrogen. For *Bub1b*$^{H/H}$ embryos and adult mice, tissue samples were sent to Transnetyx (Cordova, TN) for genotyping using the protocol described previously (Baker et al. 2004). For trisomic embryos, mouse embryonic fibroblasts (MEFs) were derived, and karyotype was determined by metaphase spreads of MEFs as described previously (Williams et al. 2008). Karytoype was confirmed by qPCR of genomic DNA derived from MEF cell pellets. Briefly, cells were digested 4-16 hours at 55ºC in 10 mM Tris pH 8.0, 100 mM NaCl, 10 mM EDTA pH 8.0, 0.5% SDS solution with 0.4 mg/ml proteinase K. Genomic DNA was precipitated with isopropanol, washed with 70% ethanol, resuspended in double deionized water, and incubated at 95ºC for 5 minutes. qPCR was performed on DNA samples by mixing Takara SYBR *Premix Ex Taq II* (Tli RNase H Plus, ClonTech, Mountain View, CA) master mix with the primers described in Supplemental Table 1, and amplified using a LightCycler 480 II (Roche). Samples were normalized to the copy number-invariant reference gene *Glucagon* (GCG) as described (Ballester et al. 2013). Genes amplified were selected based on low propensity for copy number variation (She et al. 2008) and chromosomal location, and primers were designed to amplify across intron-exon boundaries.

Homing assays

Recipient mice for homing assays were irradiated with 8.5 Gy administered as a single dose. After thawing, fetal liver cells were labeled with CM-DiI (Life Technologies) according to the manufacturer's instructions, and then evaluated for degree of labeling and viability by propidium

iodide exclusion with an LSR II flow cytometer (Becton Dickinson). $2x10^6$ live cells were injected intravenously, and bone marrow was harvested by flushing the long bones 24 hours after injection. Red blood cells were lysed in ACK Lysing Buffer, samples were washed in IMDM containing 2% FBS, then the proportion of labeled cells in the bone marrow was evaluated on an LSR II flow cytometer (Becton Dickinson). Three trisomy 19 fetal livers were pooled for homing assays due to low fetal liver cell recovery from these embryos.

Single cell sequencing

Peripheral blood was collected from CD45.1 recipient mice reconstituted with fetal liver cells isolated from E14.5 *Bub1b*$^{H/H}$ embryos and a common euploid CD45.1 donor, from mice reconstituted with E14.5 *Bub1b*$^{H/H}$ fetal liver cells alone or from mice reconstituted with bone marrow cells from 6- to 12-week old adult mice. Blood was processed as described in the Methods, and *Bub1b*$^{H/H}$ white blood cells were sorted using a MoFlo cell sorter (Beckman-Coulter) or Aria I cell sorter (Beckson Dickinson). *Bub1b*$^{H/H}$ cells were sorted using a CD45.2 antibody, and CD45.1 recipient-derived cells were excluded using a CD45.1 antibody (Biolegend).

To isolate BM-HSCs, bone marrow cells were isolated as described for bone marrow transfers. White blood cells were counted before the bone marrow was incubated at an appropriate concentration in antibodies detecting CD150, CD48, CD117, Sca-1 (from BioLegend) and a mouse lineage antibody cocktail labeled with biotin obtained from Miltenyi Biotech. Cells were subsequently incubated with anti-biotin microbeads, then differentiated cells were depleted from the cell suspension by retention of biotin-positive cells on a MACS column (Miltenyi Biotech). Lineage-depleted cell suspensions were then double sorted on a FACSAria II cell sorter (Becton Dickinson), first for lineage negative, CD48 negative cells, then for lineage negative, CD48 negative, CD150 positive, CD117 positive, Sca-1 positive cells. FL-HSCs were

isolated in the same manner, except using the markers and lineage cocktail described in the Methods for HSC quantification.

Keratinocytes were isolated from the backskin of a 4-month old *Bub1b*$^{H/H}$ mouse. Hair was plucked from the backskin, then the backskin was digested in 0.25% trypsin in PBS overnight at 4°C. The epidermis was removed from the underlying dermis using forceps, then transferred to DMEM containing 10% FBS and minced with a razor blade. The minced dermis was then both sedimented and passed through a 70 μm cell strainer twice. The pellet was again passed through a 70 μm cell strainer before isolation by microaspiration.

Intestinal crypt cells were isolated from a 4-month old *Bub1b*$^{H/H}$ mouse as described previously (Yilmaz et al. 2012). Briefly, the small intestines were dissected and the fat and mesentery was removed. The intestinal lumen was washed thoroughly with ice cold PBS using a 20G feeding needle. The intestines were then opened longitudinally, and the mucus was removed by gently rubbing the intestines in cold PBS. The intestines were then cut into small pieces and washed further in ice cold PBS with occasional shaking. The pieces were then incubated and washed three times in ice cold PBS/EDTA (10 mM) with occasional shaking before trituration with a pipette then passage through a 70 μm cell strainer. The resulting cell suspension was greatly enriched for crypt cells as assessed by visual inspection under a microscope. To liberate cells from the crypts, the cell suspension was concentrated by centrifugation, resuspended in TrypLE Express (Thermo Fisher) and incubated for 1 minute at 37°C, then placed on ice. S-MEM was added, then the cells was concentrated by centrifugation before dilution in S-MEM for isolation by microaspiration.

All single cells were isolated after appropriate preparation by microaspiration, and then amplified, sequenced and analyzed as described (Knouse et al. 2014).

Quantification of trisomic fetal liver area
Embryos were isolated at embryonic day E14.5 and fixed overnight in Bouin's solution. Embryos were sectioned into 5 μm slices and stained with hematoxylin and eosin. Slides were then scanned on a Leica Aperio slide scanner, and fetal liver area was measured using ImageJ.

Quantification of relative fetal liver cell recovery
Fetal livers were preserved and then thawed as described above. Cell number was counted in triplicate using a Cellometer Auto T4 cell counter (Nexelcom).

Supplemental References

Ballester M, Castelló A, Ramayo-Caldas Y, Folch JM. 2013. A quantitative real-time PCR method using an X-linked gene for sex typing in pigs. *Mol Biotechnol* **54**: 493–496.

She X, Cheng Z, Zöllner S, Church DM, Eichler EE. 2008. Mouse segmental duplication and copy number variation. *Nature Genetics* **40**: 909–914.

Yilmaz OH, Katajisto P, Lamming DW, Gültekin Y, Bauer-Rowe KE, Sengupta S, Birsoy K, Dursun A, Yilmaz VO, Selig M, et al. 2012. mTORC1 in the Paneth cell niche couples intestinal stem-cell function to calorie intake. *Nature* **486**: 490–495.

Supplemental Figure Legends

Supplemental Figure 1. *Evaluation of trisomic fetal livers*
Representative image of sagittal cross section of fetal livers from E14.5 WT (A) or trisomy 19 (B) embryos. (C) Quantification of trisomic fetal liver cross section area relative to cross section area of wild type littermate. (D) Relative recovery of trisomy 16, trisomy 19 and *Bub1b*$^{H/H}$ E14.5 fetal liver cells compared to wild type littermates after cryo-preservation and thawing. Data are shown as mean ± SD.

Supplemental Figure 2. *Peripheral blood analyses of trisomy 16 primary recipients*
Flow cytometry of peripheral blood from trisomy 16 and wild type primary recipients was performed to quantify the percentage of B220-positive B cells (A) and Thy1.2-positive T cells (B) in the peripheral blood. Complete blood cell counts of peripheral blood from trisomy 16 and wild type primary recipients was performed to determine red blood cell count (C), mean corpuscular volume (MCV) (D), mean corpuscular hemoglobin (MCH) (E), mean corpuscular hemoglobin concentration (MCHC) (F), hemoglobin content (G) and platelet counts (H). The composition of the leukocyte population was evaluated by automated differential for trisomy 16 and wild type primary recipients (I). In all graphs, the bars represent the mean value for each time point and condition, and asterisks indicate the trisomy 16 values are significantly different from the values of wild type littermates at the indicated time by t-test ($p<0.05$). (J) Plots of red blood cell number versus mean corpuscular volume (MCV) of all trisomy 16 primary recipients and wild type primary recipients at all times evaluated above. (K) Quantification of the percentage of $CD150^+$, $Sca\text{-}1^+$, $CD117^+$, $CD48^-$ cells in the bone marrow of trisomy 16 and wild type primary recipients at 25 weeks post-transplantation. Data are represented as mean ± SD.

Supplemental Figure 3. *Peripheral blood analyses of trisomy 19 primary and secondary recipients*
Flow cytometry of peripheral blood from trisomy 19 and wild type primary recipients was performed to quantify the percentage B220-positive B cells (A) and Thy1.2-positive T cells (B) in the peripheral blood. Complete blood cell counts of peripheral blood from trisomy 19 and wild type primary recipients was performed to determine red blood cell counts (C), mean corpuscular volume (MCV) (D), mean corpuscular hemoglobin (MCH) (E), mean corpuscular hemoglobin concentration (MCHC) (F), hemoglobin content (G) and platelet counts (H). The composition of the leukocyte population was evaluated by automated differential for trisomy 19 and wild type

primary recipients (I). In all graphs, the bars represent the mean value for each time point and condition. (J) Plots of red blood cell number versus mean corpuscular volume (MCV) of all trisomy 19 primary recipients and wild type primary recipients at all times evaluated above. Flow cytometry of peripheral blood from trisomy 19 and wild type secondary recipients was performed to quantify the percentage of B220-positive B cells (K) and Thy1.2-positive T cells (L) in the peripheral blood. Complete blood cell counts of peripheral blood from trisomy 19 and wild type secondary recipients was performed to determine white blood cell counts (M), hematocrit (N), red blood cell counts (O), mean corpuscular volume (MCV) (P), mean corpuscular hemoglobin (MCH) (Q), mean corpuscular hemoglobin concentration (MHCH) (R), hemoglobin content (S) and platelet counts (T). The composition of the leukocyte population was evaluated by automated differential for trisomy 19 and wild type secondary recipients (U). (V) Quantification of the percentage of $CD150^{+}$, $Sca\text{-}1^{+}$, $CD117^{+}$, $CD48^{-}$ cells in the bone marrow of trisomy 19 and wild type secondary recipients at 62 weeks post-transplantation. Data are represented as mean ± SD. In all graphs, the bars represent the mean value for each time point and condition, and asterisks indicate the trisomy 19 values are significantly different from the values of wild type littermates at the indicated time by Student's t-test ($p<0.05$).

Supplemental Figure 4. *Peripheral blood cell and HSC analyses of trisomy 19 tertiary recipients and quaternary recipient survival*
Flow cytometry of peripheral blood from trisomy 19 and wild type tertiary recipients was performed to quantify the percentage of B220-positive B cells (A) and Thy1.2-positive T cells (B) in the peripheral blood. Complete blood cell counts of peripheral blood from trisomy 19 and wild type tertiary recipients was performed to determine white blood cell counts (C), hematocrit (D), red blood cell counts (E), mean corpuscular volume (MCV) (F), mean corpuscular

hemoglobin (MCH) (G), mean corpuscular hemoglobin concentration (MHCH) (H), hemoglobin content (I) and platelet counts (J). The composition of the leukocyte population was evaluated by automated differential for trisomy 19 and wild type tertiary recipients (K). (L) Survival of trisomy 19 and wild type tertiary recipients. (M) Quantification of the percentage of $CD150^+$, $Sca\text{-}1^+$, $CD117^+$, $CD48^-$ cells in the bone marrow of trisomy 19 and wild type tertiary recipients at 44 weeks post-transplantation. Data are represented as mean ± SD. (N) Segmentation plots showing the karyotype of CD45.2 peripheral blood cells isolated from two trisomy 19 quaternary recipients (one with high trisomy 19 contribution, one with low trisomy 19 contribution). Segmentation plots show the copy number of all chromosomes from 1 to X relative to a euploid reference on a $\log_2$ scale. Trisomic chromosome is highlighted in red. (O) Survival of trisomy 19 and wild type quaternary recipients. In all graphs, the bars represent the mean value for each time point and condition, and asterisks indicate the trisomy 16 values are significantly different from the values of wild type littermates at the indicated time by t-test ($p<0.05$).

Supplemental Figure 5. *Peripheral blood of adult $Bub1b^{H/H}$ mice is normal and $Bub1b^{H/H}$ adult bone marrow HSCs show no fitness defects in competitive reconstitution assays*
Complete blood cells counts of 12-week old *$Bub1b^{H/H}$* mice and wild type littermate controls: (A) red blood cell count, (B) white blood cell count, (C) mean corpuscular volume (MCV), (D) mean corpuscular hemoglobin (MCH), (E) mean corpuscular hemoglobin concentration (MCHC), (F) hematocrit, (G) hemoglobin content and (H) platelet counts. (I) Quantification of the percentage of $CD150^+$, $Sca\text{-}1^+$, $CD117^+$, $CD48^-$ cells in the bone marrow of adult *$Bub1b^{H/H}$* mice and wild type littermates (n=3 for each). Data are represented as mean ± SD. (J-L) CD45.2 *$Bub1b^{H/H}$* (J) or wild type littermate (K) bone marrow cells were co-transferred with an equal number of bone marrow cells from a common CD45.1 donor mouse of the same age into a

lethally irradiated CD45.1 recipient. Peripheral blood was sampled at indicated times and the percentage of the white blood cell population contributed by each donor was quantified by flow cytometry with antibodies against CD45.1 and CD45.2. Data are represented as mean ± SD for each time point. (L) Ratios of the average percentage of the peripheral blood reconstituted by the *Bub1b*$^{H/H}$ bone marrow cells to the average percentage of the peripheral blood reconstituted by wild type littermate bone marrow cells at all indicated times.

Supplemental Figure 6. *Peripheral blood analyses and HSC counts from Bub1b*$^{H/H}$ *bone marrow primary and secondary recipients*

(A-I) Complete blood cell counts of peripheral blood from *Bub1b*$^{H/H}$ and wild type primary bone marrow recipients were performed to determine white blood cell counts (A). Flow cytometry was performed to quantify the percentage B220-positive B cells (B) and Thy1.2-positive T cells (C) in the peripheral blood of primary recipients. Complete blood cell counts further determined red blood cell counts (D), mean corpuscular volume (MCV) (E), hematocrit (F), hemoglobin content (G) and platelet counts (H). The composition of the leukocyte population was evaluated by automated differential for *Bub1b*$^{H/H}$ and wild type primary recipients (I). (J) Quantification of the percentage of CD150^{+}, Sca-1^{+}, CD117^{+}, CD48^{-} cells in the bone marrow of *Bub1b*$^{H/H}$ and *Bub1b*$^{+/H}$ primary recipients 63 weeks post-transplantation. Data are represented as mean ± SD. (K-S) Complete blood cell counts of peripheral blood from *Bub1b*$^{H/H}$ and wild type secondary bone marrow recipients were performed to determine white blood cell counts (K). Flow cytometry was performed to quantify the percentage B220-positive B cells (L) and Thy1.2-positive T cells (M) in the peripheral blood of secondary recipients. Complete blood cell counts further determined hematocrit (N), red blood cell counts (O), mean corpuscular volume (MCV) (P), hemoglobin content (Q) and platelet counts (R). The composition of the leukocyte

population was evaluated by automated differential for *Bub1b*$^{H/H}$ and wild type secondary recipients (S). (T) Quantification of the percentage of CD150^{+}, Sca-1^{+}, CD117^{+}, CD48^{-} cells in the bone marrow of *Bub1b*$^{H/H}$ and wild type secondary recipients at 47 weeks post-transplantation. Data are represented as mean ± SD. In all graphs, the bars represent the mean value for each time point and condition, and asterisks indicate the *Bub1b*$^{H/H}$ values that are significantly different from the values of wild type littermates at the indicated time by t-test ($p<0.05$).

Supplemental Figure 7. *Single cell sequencing analyses of Bub1b*$^{H/H}$ *fetal liver HSCs and peripheral blood cells from a Bub1b*$^{H/H}$ *fetal liver competition assay recipient and a Bub1b*$^{H/H}$ *secondary bone marrow recipient*
Segmentation plots of single cell karyotypes determined by single cell sequencing of CD45.2 positive peripheral blood cells derived from *Bub1b*$^{H/H}$ FL-HSCs from Fig. 1F (A) and from bone marrow HSCs from secondary recipients from Fig. 6A (B). Segmentation plots of single cell karyotypes determined by single cell sequencing of HSCs from *Bub1b*$^{H/H}$ fetal liver (C) or bone marrow (D) from Fig. 6A. Segmentation plots show the copy number of single cells from chromosome 1 to X relative to a euploid reference on a $\log_2$ scale. Segments above the threshold for whole chromosome gain are shown in red, and segments below the threshold for whole chromosome loss are shown in green. Cells that are classified as aneuploid are highlighted with and asterisk on the right.

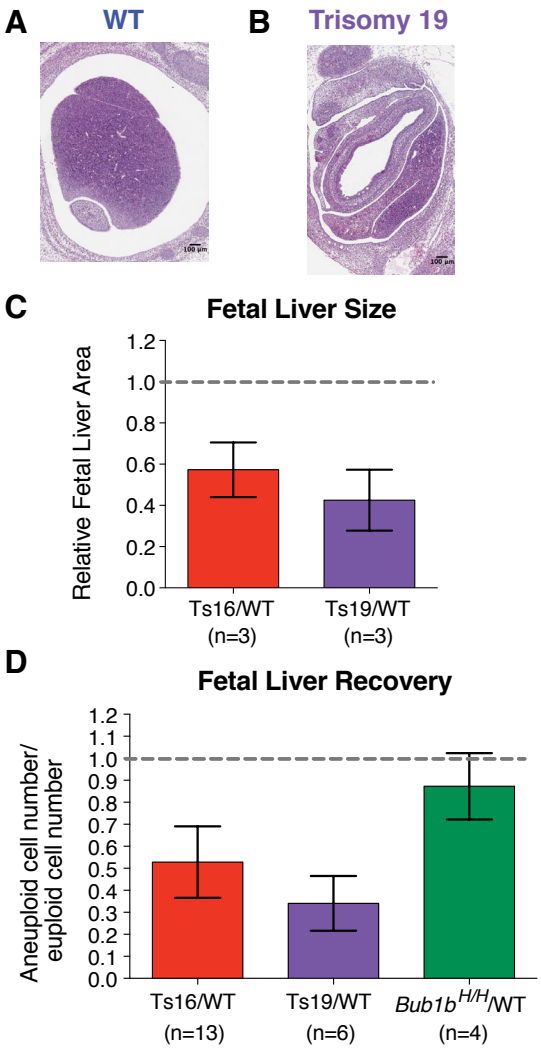
Supplemental Figure 8. *Single cell sequencing analyses of Bub1b*$^{H/H}$ *fetal liver and bone marrow primary recipients with time*
Segmentation plots of single cell karyotypes determined by single cell sequencing of CD45.2 positive peripheral blood cells derived from *Bub1b*$^{H/H}$ FL-HSCs 3 weeks (A), 6 weeks (B), 13 weeks (C) and 36 weeks (D) after transfer; and of CD45.2 positive peripheral blood cells derived

from *Bub1b*$^{H/H}$ BM-HSCs 3 weeks (E), 6 weeks (F), 13 weeks (G) and 34 weeks (H) after transfer. Segmentation plots show the copy number of single cells from chromosome 1 to X relative to a euploid reference on a $\log_2$ scale. Segments above the threshold for whole chromosome gain are shown in red, and segments below the threshold for whole chromosome loss are shown in green. Cells that are classified as aneuploid are highlighted with and asterisk on the right.

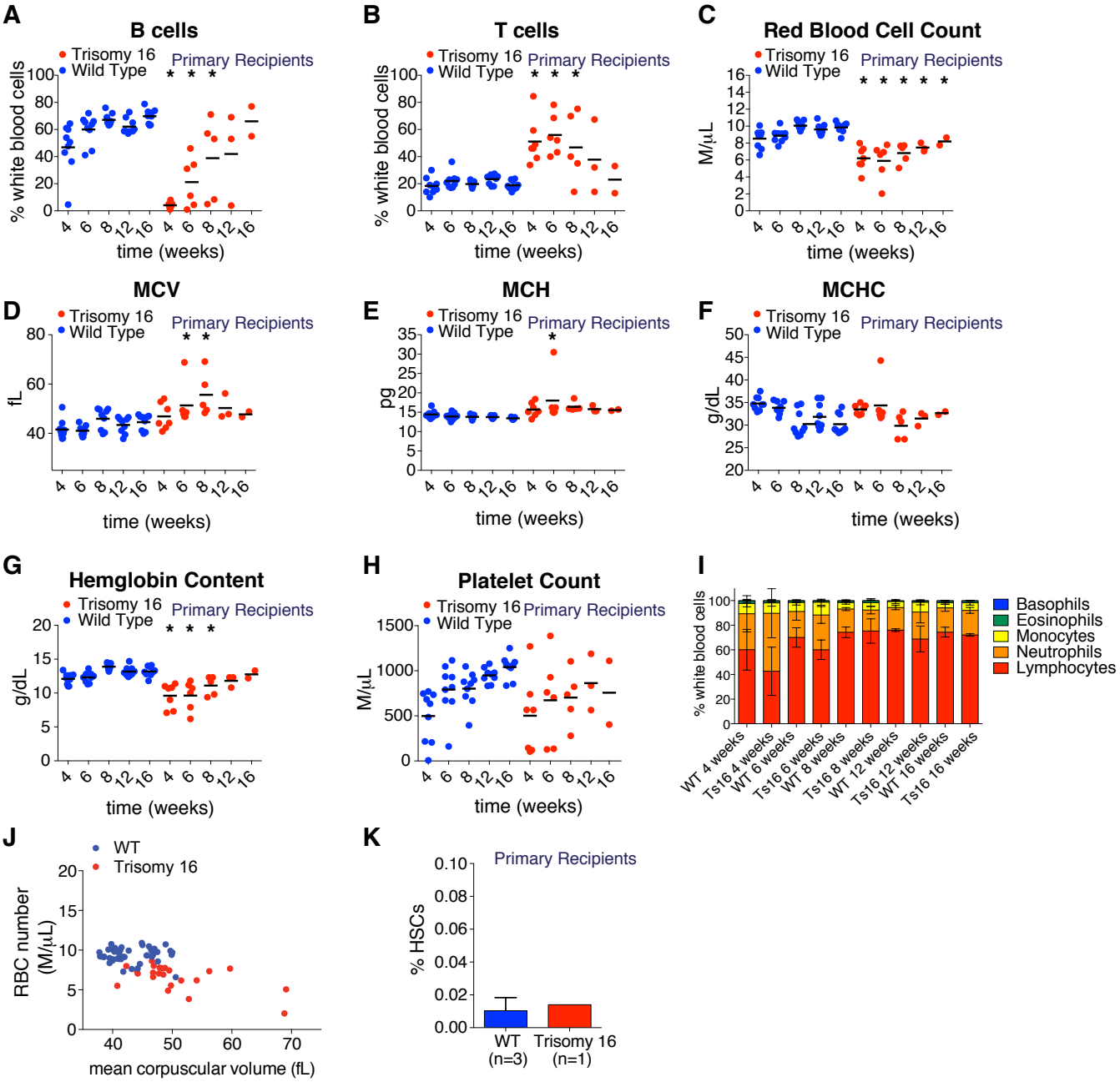
Supplemental Table 1. *Primers used for karyotyping qPCR*

Chromosome	Gene	Forward Primer	Reverse Primer	Source
2	Glucagon (*GCG*)	5'–AACATTGCCAAACGTCATGATG–3'	5'–GCCTTCCTCGGCCTTTCA–3'	Ballester et al., 2013
X	Glycoprotein M6B (*GM6B*)	5'–CTCTTCCACCAGCTGATCTACATG–3'	5'–TCCCGACTCTTAAACTTCAAAACC–3'	Ballester et al., 2013
16	Runt-related transcription factor 1 (*RUNX1*)	5'–CAGGTATACCTTGGATCAGTGC–3'	5'–CAACACAGCATCTTCTGATGGC–3'	This study
16	Eph receptor A3 (*EPHA3*)	5'–AGGAATCATCCCAGCAACACAC–3'	5'–GAGAGCAATCTAGTATTGTTCTGGG–3'	This study
16	Oxysterol binding protein-like 11 (*OSBPL11*)	5'–CCCAATTAAGTGCATACCCAGC–3'	5'–CAAGAGACAGTCAGCAAACACGG–3'	This study
16	Epithelial membrane protein 2 (*EMP2*)	5'–CTCTGTTCTCATGAATGAGCCTG–3'	5'–CAGAAAGAATCGAAGGGAGATTG–3'	This study
19	Bestrophin 1 (*BEST1*)	5'–CAGGGCAGAGGTCATGGTTC–3'	5'–CTGGTGCTCAAGGCAGACCT–3'	This study
19	Ankyrin repeat domain 1 (cardiac muscle) (*ANKRD1*)	5'-GTGCACATGGAAATGACTGG-3'	5'-TGGGCCACAACTCAATGTTA-3'	This study
19	Oligonucleotide/ oligosaccharide-binding fold containing 1 (*OBFC1*)	5'-CTGCACGAAACCTTGCATGA-3'	5'-GCCCCGGCTGATCTTAATCT-3'	This study
19	Caspase 7 (CASP7)	5'–CAATCTGCCACTCTGCAACC–3'	5'–CAGCAACATTGAACAGGCT– 3'	This study

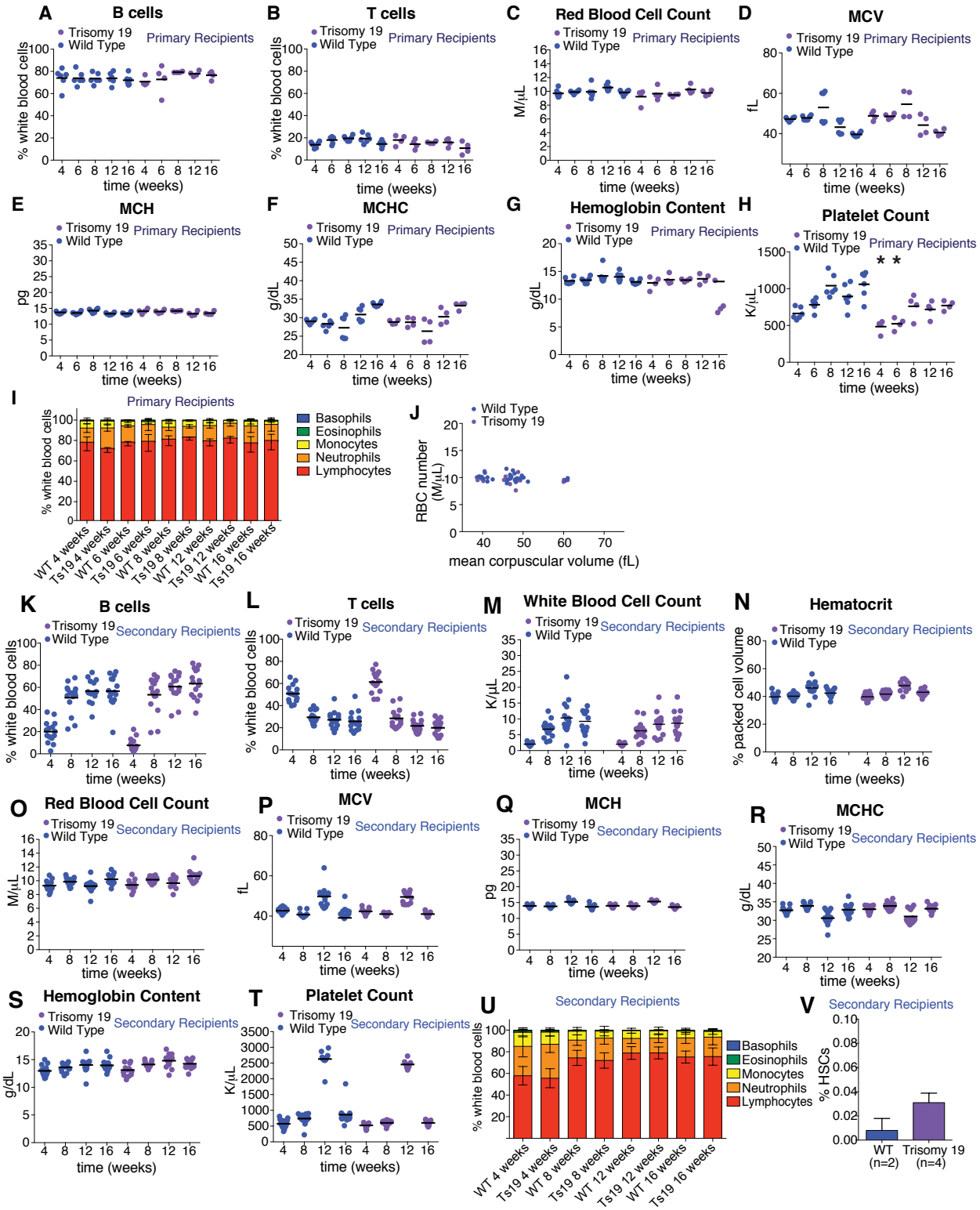
Supplemental Figure S1



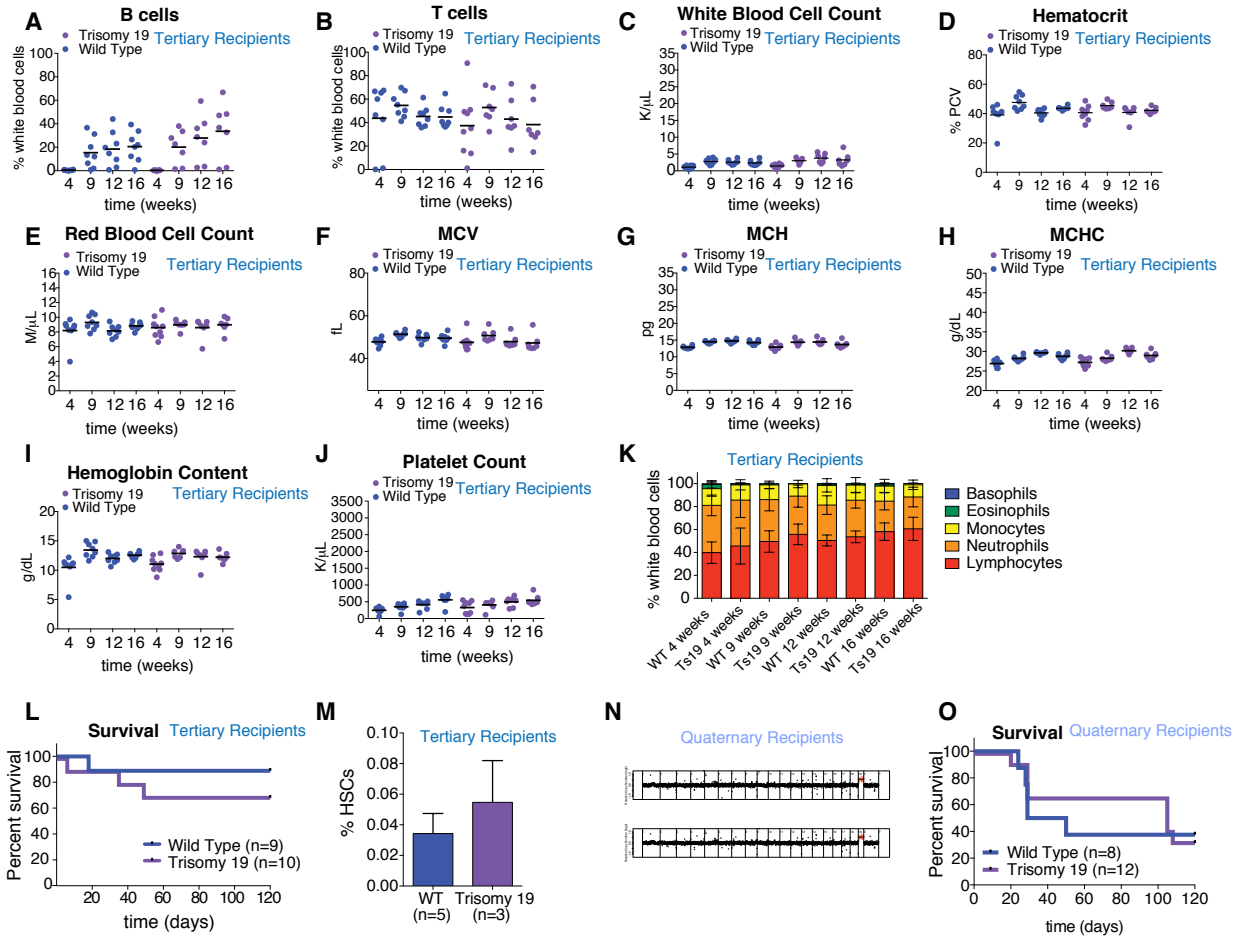
Supplemental Figure S2



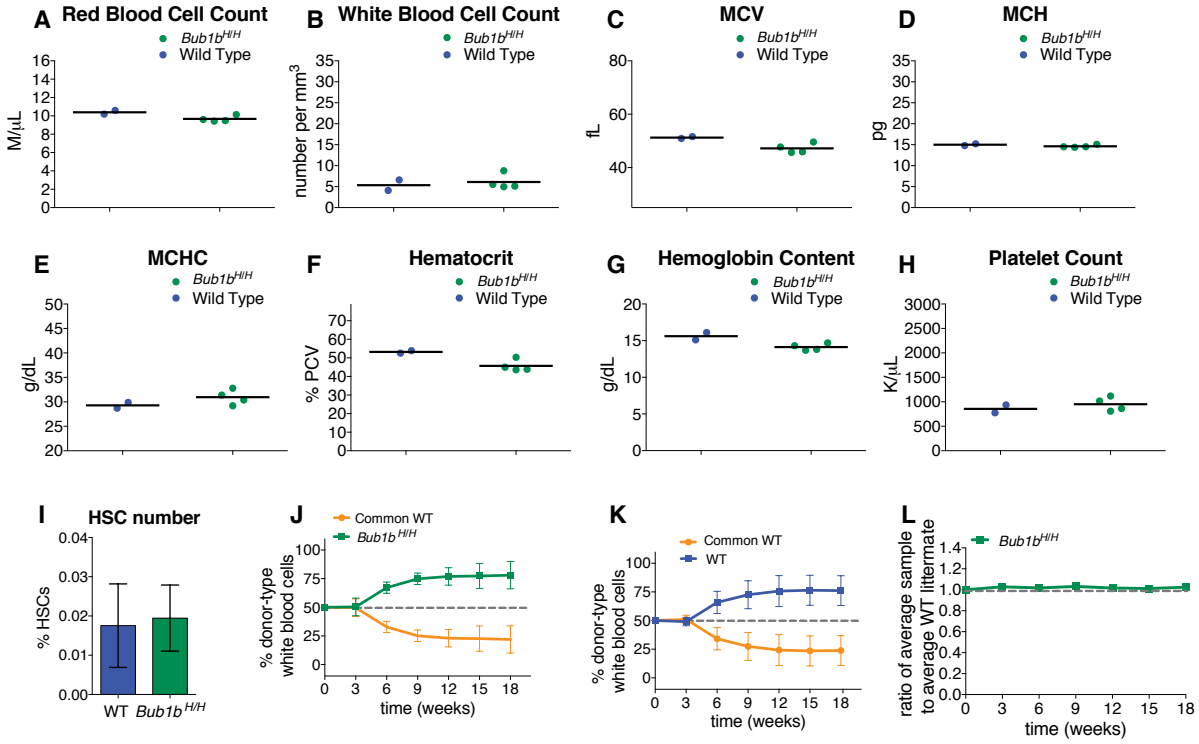
Supplemental Figure S3



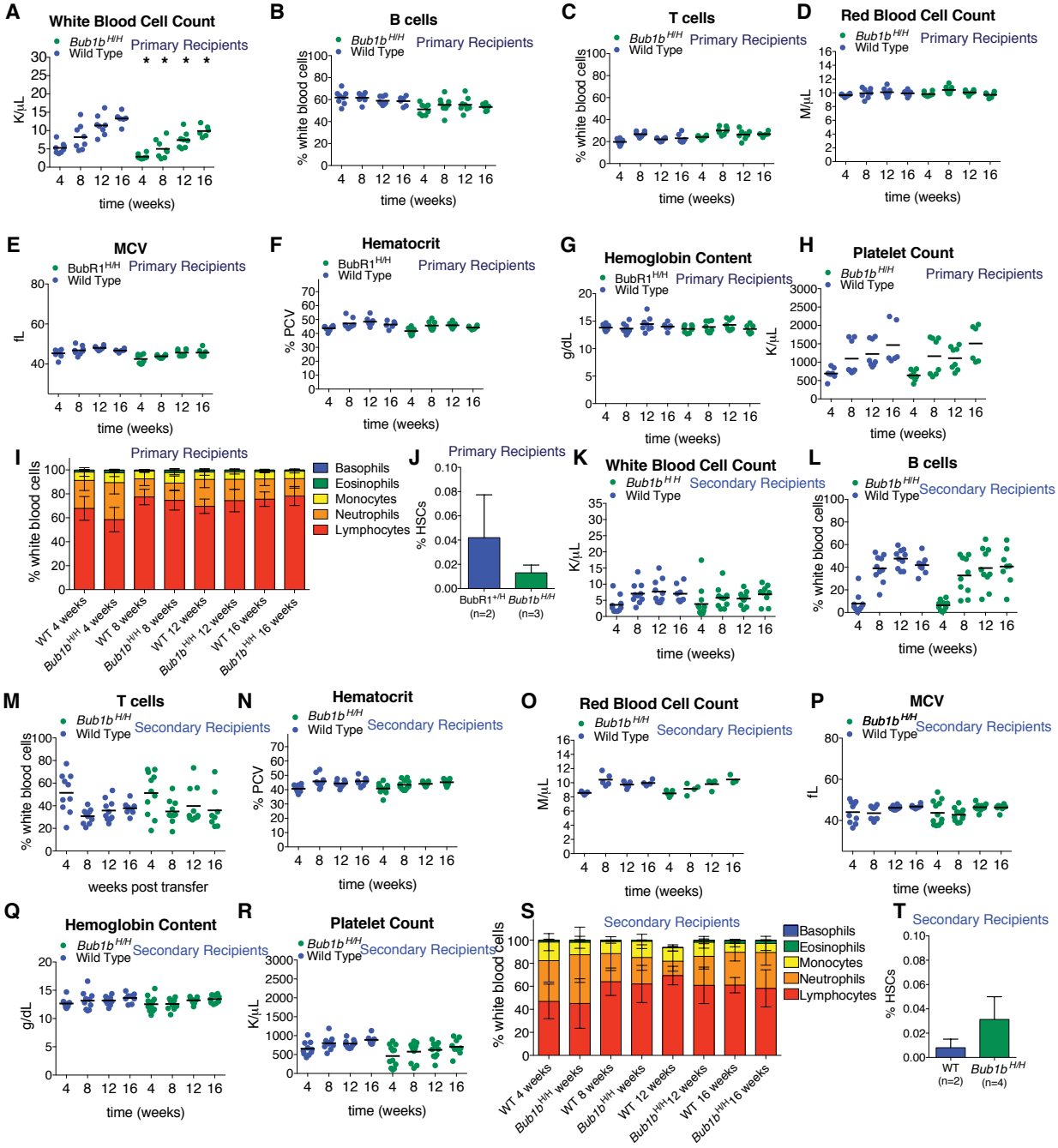
Supplemental Figure S4



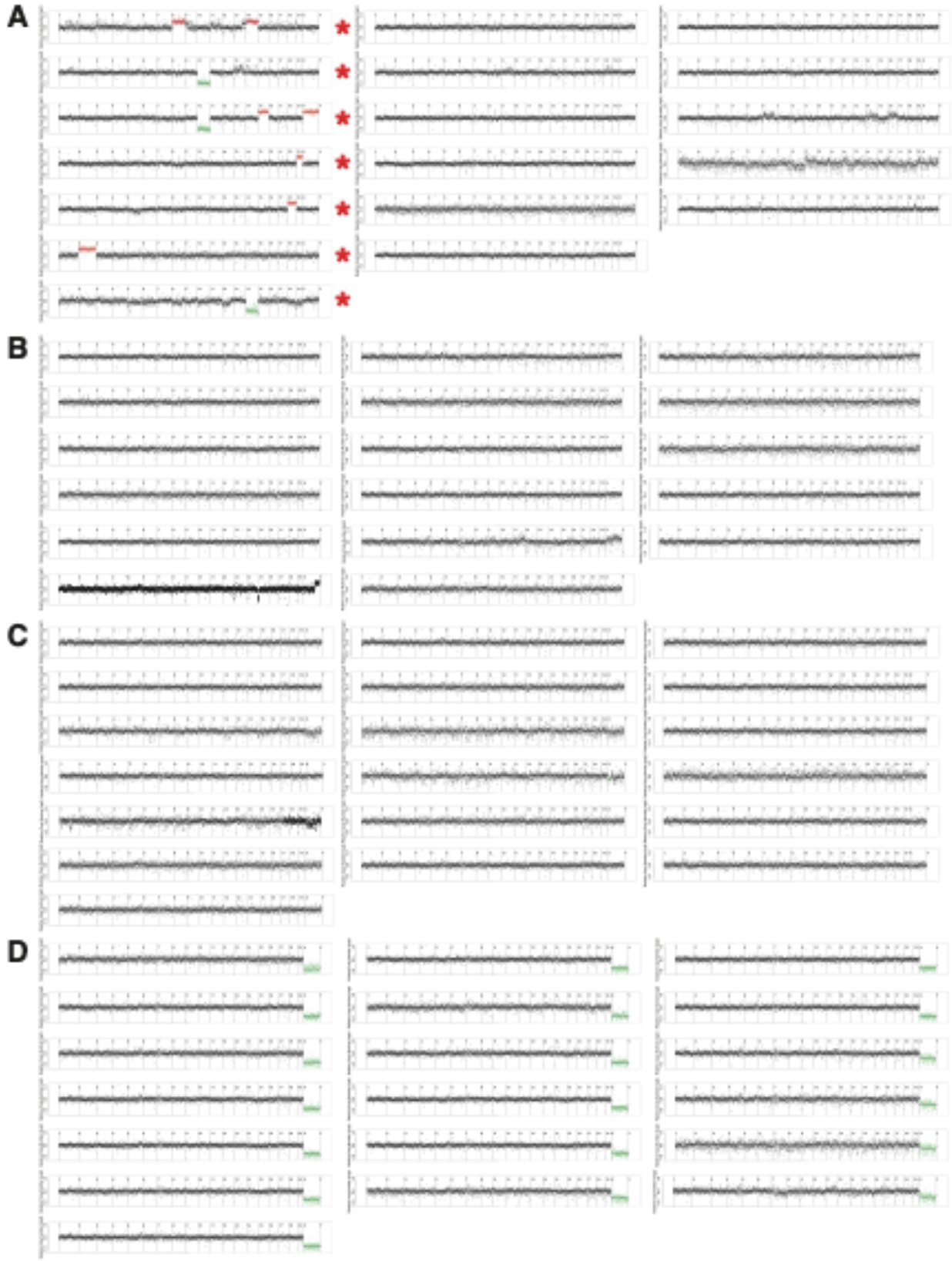
Supplemental Figure S5



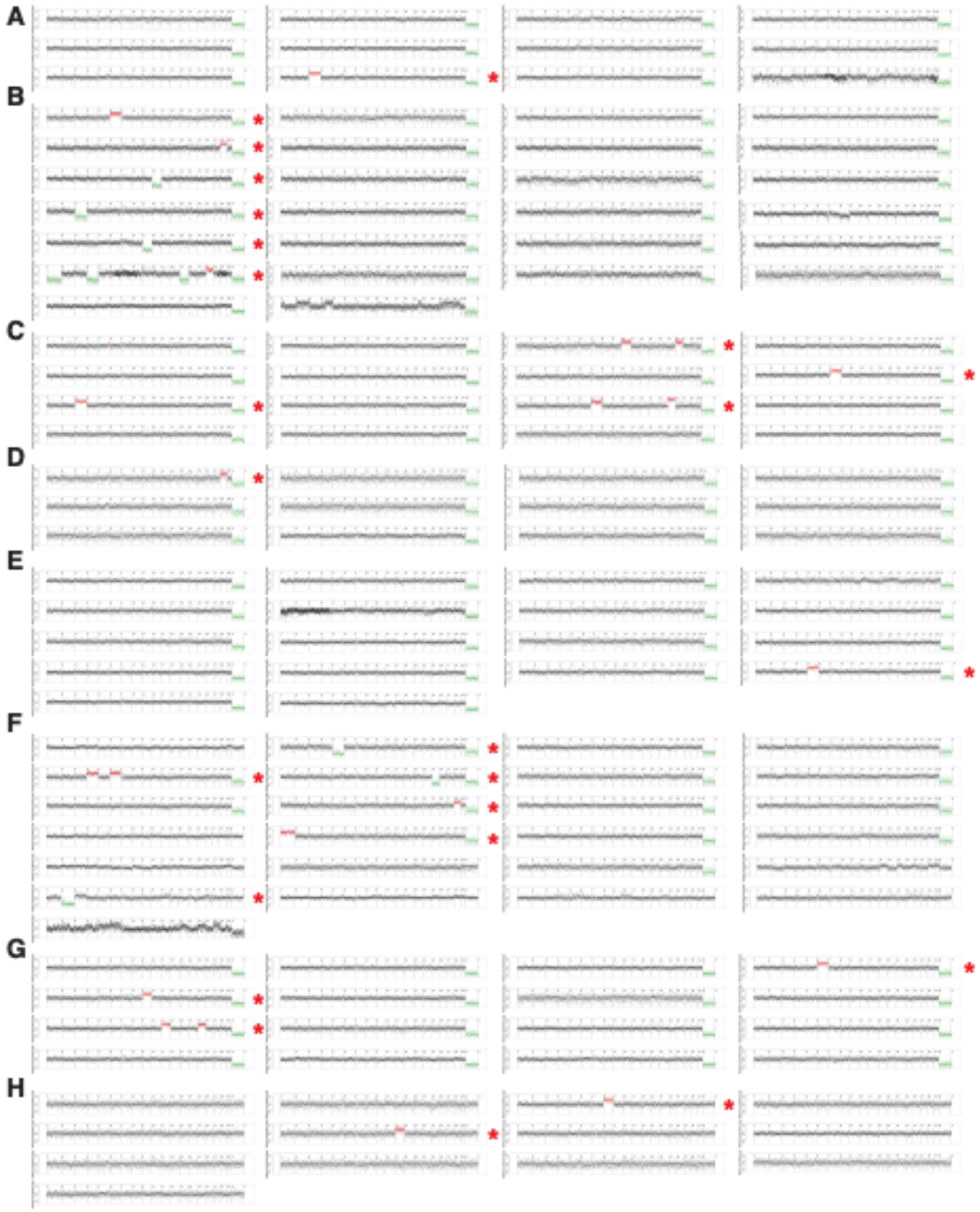
Supplemental Figure S6



Supplemental Figure S7



Supplemental Figure S8



Chapter Three: A somatic evolutionary model of the dynamics of aneuploid cells during hematopoietic reconstitution

Rozhok AI, Silberman RE, Higa KC, Liggett LA, Amon A, DeGregori J. 2020. A somatic evolutionary model of the dynamics of aneuploid cells during hematopoietic reconstitution. *Science Reports* **10**:1-10.

Abstract

Aneuploidy is a feature of many cancers. Recent studies demonstrate that in the hematopoietic stem and progenitor cell (HSPC) compartment aneuploid cells have reduced fitness and are efficiently purged from the bone marrow. However, early phases of hematopoietic reconstitution following bone marrow transplantation provide a window of opportunity whereby aneuploid cells rise in frequency, only to decline to basal levels thereafter. Here we demonstrate by Monte Carlo modeling that two mechanisms could underlie this aneuploidy peak: rapid expansion of the engrafted HSPC population and bone marrow microenvironment degradation caused by pre-transplantation radiation treatment. Both mechanisms reduce the strength of purifying selection acting in early post-transplantation bone marrow. We explore the contribution of other factors such as alterations in cell division rates that affect the strength of purifying selection, the balance of drift and selection imposed by the HSPC population size, and the mutation selection balance dependent on the rate of aneuploidy generation per cell division. We propose a somatic evolutionary model for the dynamics of cells with aneuploidy or other fitness-reducing mutations during hematopoietic reconstitution following bone marrow transplantation. Similar alterations in the strength of purifying selection during cancer development could help explain the paradox of aneuploidy abundance in tumors despite somatic fitness costs.

Introduction

Aneuploidy, or deviation of the chromosome number from the normal karyotype (resulting from chromosome mis-segregation), is associated with many cancers, being prevalent in both solid cancers and leukemia [1–4]. The effects of aneuploidy vary, with some cellular phenotypes dependent on what specific chromosome(s) is lost or gained, and other phenotypes arising from a general stress response to aneuploidy [5]. Consequently, associations of aneuploidy with cancers

range widely, from a few percent, such as the loss of chromosome 1 or gain of chromosome 5 in kidney adenocarcinoma, to 50%, such as the loss of chromosome 3 in melanoma, and even 70% for the loss of chromosome 22 in meningiomas [2]. In total, almost 90% of cancers exhibit gains or losses of at least one chromosome arm, with patterns specific to particular tumor types [4]. For example, squamous cell cancers originating in multiple organs exhibit a common pattern of chromosome arm 3p loss and 3q gain.

Aneuploid cells have been shown to drive adaptation in yeast [6–8]. This evidence has led to speculations that by conferring greater adaptability, variability of chromosomal ploidy in a cell population might lead to the expansion of aneuploid clones in human tissues, fueling further accumulation of oncogenic alterations in cells and progression to cancer [9–11]. However, aneuploidy has been shown to more commonly reduce the fitness of animal somatic cells [5], just as most chromosome gains in yeast reduce their fitness [12,13]. Similarly, mice engineered to model the human trisomy of chromosome 21, the cause of Down syndrome, demonstrate lower proliferative potential of hematopoietic stem cells (HSC), mammary epithelial cells, neural progenitors and fibroblasts [14]. Multiple mouse models of spindle assembly checkpoint mutants are lethal indicating that high level chromosome mis-segregation is highly detrimental [15]. Aneuploidy has also been shown to promote premature differentiation and depletion of neural and intestinal stem cells in *Drosophila melanogaster* [16]. We recently examined how aneuploidy impacts hematopoietic stem and progenitor cell (HSPC) fitness using transplantation of bone marrow from aneuploid or aneuploid-prone mouse models [15]. This study demonstrated that increased aneuploidy is associated with reduced somatic stem cell fitness *in vivo,* as such aneuploid cells are efficiently purged from the hematopoietic compartment.

These experiments raise the question of how aneuploidy can be so tightly associated with a

vast array of cancers, given that cancer development requires a series of expansions and fitness gains by more proliferative cell clones. One answer would be that only specific types of aneuploidy are involved in cancer. However, evidence shows that aneuploidy has various degrees of association with cancers across the board, including a gain or loss of almost any human chromosome [4]. We performed computational modeling that indicates that rapid expansion of the engrafted HSC population together with reduced support of HSC stemness from damaged bone marrow microenvironments are plausibly the two primary mechanisms weakening purifying selection in early post transplant bone marrow, providing a window of opportunity for the expansion of aneuploid HSCs. These results have implications for the generation of aneuploid cells in other contexts, including during cancer development.

Results

In the context of bone marrow transplantation in mice, we previously showed that the peripheral blood descendants of aneuploidy-prone HSPCs demonstrate an immediate and substantial rise in the frequency of aneuploidy after bone marrow transplantation, despite a clear fitness disadvantage relative to euploid cells [15]. For these experiments, aneuploid cells were generated at an increased rate due to a hypomorphic mutation in the mitotic spindle assembly checkpoint protein gene BUB1-related 1 (*BUBR1*). As shown in Fig. 1A, following the rise in the fraction of aneuploid cells in peripheral blood post transplantation, the frequency of aneuploid cells subsequently declines to the low baseline levels typical of unperturbed blood cells. Given that peripheral blood is regularly generated from HSC and downstream progenitors in the bone marrow, this pattern suggests that early reconstitution provides a window of opportunity for the enrichment of aneuploid self-renewing cell types, such as HSC and HSPC, despite their lower fitness. We currently do not have data to discriminate whether the observed

aneuploidy originates from HSC or later more committed cell lineages of hematopoiesis. However, as aneuploidy declines later during BM reconstitution, we focus this research on finding the answer to the observation that tolerance of aneuploidy demonstrates a temporal change, being higher early in the process. Thus, while our model starts with parameters determined for HSC, the general principles to be explored here should be relevant for different progenitor stages and cell lineages.

The early phase of bone marrow reconstitution after transplantation differs from steady-state hematopoiesis in several respects. First, HSCs and HSPCs are known to divide much faster immediately after transplantation and return to their normal cell cycle rate later [17]. Early post-transplantation bone marrow also has free niche space after irradiation kills recipient HSCs, such that the transplanted population is not at an equilibrium but expands until the entire bone marrow niche space is reclaimed by the engrafted HSCs, in order to restore normal hematopoiesis. We expect that other progenitor compartments will behave similarly, as their numbers are reduced post-irradiation followed by recovery. Also, radiation exposure causes substantial damage to the bone marrow microenvironment, including via genomic damage, oxidative stress causing profound inflammation [18,19]. Such perturbation likely reduces the functionality of HSC niches for stem cell maintenance, similarly to the effect shown for mesenchymal stem cells [18]. Niche perturbation should thus reduce the strength of purifying selection by impairing the support of HSPC relatively independently of HSPC phenotype.

Model Architecture

Using Matlab, we simulated bone marrow reconstitution by creating a virtual niche space as a matrix of 10,000 single-cell niches, based on estimated numbers of HSC in mice [20,21]. The initial number of HSCs was 100, reflecting the approximate number of HSC in a million transplanted mouse bone marrow cells [15]. We set a rate of generation to aneuploidy and the

average fitness cost relative to normal HSC and explored a range of both parameters in simulations. Cell division rate started from fast division (once per ~3 days) and returned to once per ~40 days when bone marrow HSC population size returned to the physiological ~10,000 cells. Total reconstitution was achieved in ~8 simulated weeks. The simulation lasted for 600 simulated days with daily updates. At each update, HSCs divided stochastically based on the current division rate. Excess cells were removed as a result of a binomial trial with probabilities of staying an HSC based on the number of HSCs after division, the assumed current population size based on the growth curve, and their relative fitness.

HSC niches had an additional property – the ability to maintain HSC stemness dependent on niche health after irradiation. This ability was realized by implementing an additional binomial "survival" trial with a certain probability for an HSC to leave the pool regardless of the cell's relative fitness. Percentages of aneuploid cells were then tracked over the entire simulation time (600 days) under various values of model parameters. Whenever not directly manipulated as shown in Fig. 1B,C and Fig. 2D,E, aneuploidy generation rate per cell division was kept at 1%, and aneuploid cell fitness effects at -1% as standard parameters for all experiments. Higher generation rates and fitness effects are also explored. Cell division rate profile was as described in Methods and shown in Fig. S1 whenever such a parameter was not varied on purpose or unless indicated otherwise. More details of the model are explained in Methods and the Matlab code is provided in Supplements, section "Model code". The niche health model and rationale will be explained later.

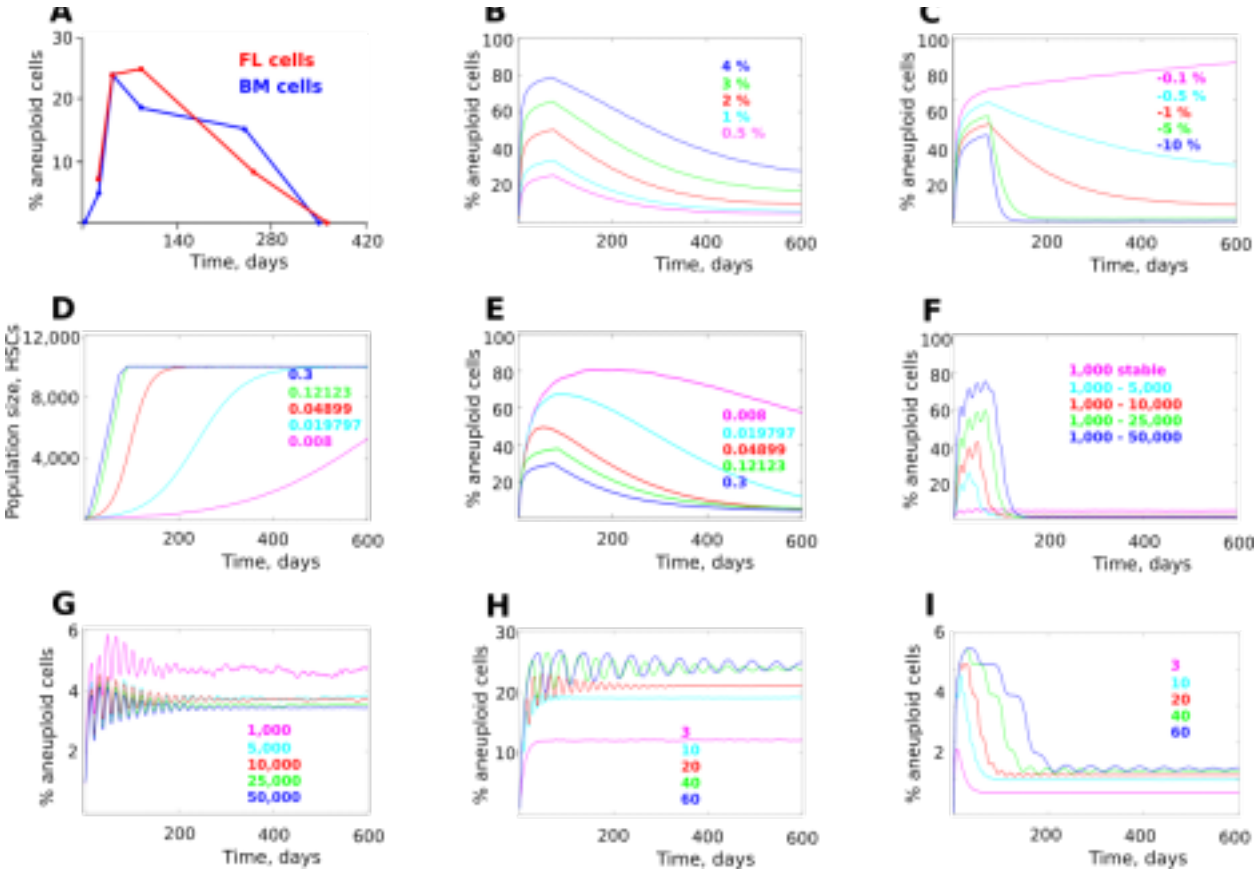


Fig. 1. Frequency of aneuploid HSC cells in the simulated post-transplantation bone marrow. (A) Frequency of aneuploid cells observed in peripheral blood of recipient mice after receiving transplanted bone marrow from BUBR1$^{H/H}$ (hypomorphic) mice; data are from Pfau et al. [15]; data at days 350 and 364 says was collected following the protocol in Pfau et al. [15]; FL – transplanted fetal liver cells, BM – transplanted bone marrow cells; see Supplements section Aneuploidy counts for a summary of data. (B) Simulated aneuploidy dynamics with varying aneuploidy generation rate per cell division (numbers color matched to respective data lines; statistics in Fig.S2). (C) Simulated aneuploidy dynamics with a range of cell fitness cost induced by aneuploidy (statistics in Fig.S3). (D) Dynamics of HSC population increase post transplantation over time (color-matched numbers represent growth coefficients which determined the shape of the population size growth). (E) Simulated aneuploidy dynamics under various cell population expansion regimens (numbers color-matched as in (D); statistics in Fig.S4). (F) Simulated aneuploidy frequency at stable cell division rate of 1 in 20 days and various extent of cell population size expansion (color-matched numbers indicate initial and final population size in # of cells; statistics in Fig.S5). (G) Simulated aneuploidy frequency at a stable cell division rate of 1 in 20 days and different stable cell population sizes (color-matched numbers indicate population size in # of cells; statistics in Fig.S6). (H) Simulated aneuploidy frequency at a stable population size of 10,000 cells and varying stable cell division rates (color-matched numbers indicate the average interval in days between successive cell divisions; statistics in Fig.S7). (I) Simulated aneuploidy frequency under population expansion from 1,000 to 10,000 cells and varying stable cell division rates (color-matched numbers as in (H); statistics in Fig.S8).

Niche-independent effects of bone marrow transplantation on aneuploid cell frequency

We set out to explore if altered purifying selection underlies the observed aneuploidy peak following bone marrow transplantation (see Fig. 1A). As shown in Fig. 1B, with a fitness cost applied to aneuploid cells relative to normal cells, the model replicated the general pattern observed experimentally, with aneuploid cell frequency peaking early and subsiding afterwards. Varying aneuploidy generation rate per cell division with a fixed aneuploidy fitness cost of -1% significantly affected the height of the frequency peak, while the pattern of the subsequent frequency decline remained similar (Fig. 1B). This result indicates that the mutation-selection balance that depends on the aneuploidy generation frequency per cell division is one of the evolutionary forces that should impact the aneuploidy frequency peak, particularly its height. We further set a fixed aneuploidy generation rate at 1% and applied the following range of aneuploidy fitness costs [-10%, -5%, -1%, -0.5%, -0.1%]. While increasing fitness cost resulted in a lowering of the peak of aneuploidy (Fig. 1C), the effect was relatively modest over the 100-fold range of fitness costs, indicating that purifying selection is weak during the expansion phase. In contrast, fitness cost demonstrates a significantly stronger effect on the purging of aneuploid cells that follows this peak, consistent with a strengthening of purifying selection during the post-reconstitution period. Fig. 1C also shows that if the aneuploidy fitness cost is low enough, aneuploidy frequency does not decrease. These results demonstrate that one explanation for the post-transplantation frequency of aneuploidy that we have observed earlier [15] could be the shifting balance between mutation (aneuploidy generation) rate and the strength of purifying selection.

Pre-transplantation radiation treatment eliminates resident HSCs and thus vacates HSC niche space for transplants, leaving room for the transplanted HSC population to expand.

Population expansion should increase the presence of random drift and reduce the strength of purifying selection [22–24]. As shown in Fig. 1D, we applied a range of HSC population expansion rates following transplantation with the same maximum population size. The rate of population expansion demonstrates an inverse relationship with the peak frequency of aneuploidy (Fig. 1E), with faster expansion lowering the aneuploidy peak. A number of confounding factors could potentially interfere with the effects of population expansion. First, a more rapidly increasing population size can counteract the effect of population expansion, whereby the expansion itself relaxes selection while the resulting increased population size intensifies it. Also, the rapidly dividing HSCs during the early post-transplantation phase could also intensify the strength of selection by increasing the number of cell generations per time unit, as argued previously [25].

To isolate the effect of population expansion per se and explore the effect of these additional factors, we first fixed the simulated cell division rate to an intermediate average of 1 division per 20 days throughout the simulation run. We further explored a range of population expansion rates from 1,000 cells to [1,000; 5,000; 10,000, 25,000; 50,000]. As shown in Fig. 1F, without the contribution of the changing cell division rate, greater population expansion caused higher aneuploidy frequency peaks. The effect of the selection-intensifying cell population size was still present, but obviously overcome by the selection-reducing effect of population expansion. Interestingly, with no expansion the aneuploidy peak was absent, validating a role of population expansion in generating the peak (Fig. 1F). In order to test the effect of population size we tested the model with a range of stable population sizes [1,000; 5,000; 10,000; 25,000; 50,000] and with the same stable cell division rate as in Fig. 1F. In the absence of the effects of population expansion and the changing cell division rate, we see that population size does have the predicted selection-suppressing effect, producing differences in peak aneuploidy frequency

early and influencing the frequency during later phases of reconstitution. However, the difference was tangible only between the smallest population size (1000 HSC) and the rest, revealing non-linear effects and indicating that at the size of 5,000 and above drift is perhaps a minor factor (Fig. 1G). However, in the absence of the early-phase population expansion, HSCs do not demonstrate an early frequency peak.

We further fixed population size at 10,000 cells and applied a range of stable cell division rates at once per [3, 10, 20, 40, 60] days to isolate the effect of cell division rates. As predicted, Fig. 1H demonstrates that faster cell division does have a suppressive effect on aneuploidy frequency, presumably by intensifying the strength of selection (by increasing the number of cell generations per unit time). Just as in Fig. 1G, an aneuploidy peak is not produced in the absence of population expansion. In order to corroborate the role of expansion per se, we further applied population expansion from 1,000 to 10,000 cells but across the same range of fixed cell division rates at once per [3, 10, 20, 40, 60] days. Fig. 1I demonstrates that in the presence of population expansion the model generates an early peak of aneuploidy frequency. Interestingly, the height of the peak and the rate of the subsequent aneuploidy elimination is dependent on cell division rate, with faster rates being more aneuploidy-suppressive, consistent with the results in Fig. 1H. Based on these results, we can conclude that HSC population expansion during early post transplantation bone marrow reconstitution is likely to have a profound effect on the intensity of purifying selection and is another likely mechanism underlying the pattern observed *in vivo* [15].

Altogether, the results shown thus far indicate that the early post-transplantation peak of aneuploidy could be explained by shifts in the character of purifying selection that occur during hematopoietic reconstitution.

Effects of bone marrow niche/microenvironment on aneuploid cell frequency

Another process characteristic of bone marrow transplantation following ablation of the resident bone marrow cells with radiation is damage to and recovery of the bone marrow microenvironment. The post-radiation bone marrow microenvironment appears to be highly perturbed by direct radiation-induced damage and the ensuing reactive oxygen species generation and inflammation [18,19]. Similar to such effects shown for mesenchymal stem cells [18], these changes should reduce the ability of the bone marrow niche to support stemness in HSC, which require proper microenvironmental signaling to maintain homeostatic differentiation rates [26]. We therefore further reasoned that the perturbed microenvironment should decrease the ability of HSC to maintain stemness, and this influence should be less discriminating between aneuploid and normal HSCs, being poorly supportive for all HSC (Fig. 2A). Such an effect might reduce the strength of selection based on intrinsic cell fitness differences.

In order to model the damaged niche shown in Fig. 2A, we added a stochastic effect exerted by the bone marrow niche and affecting all cells by adding an additional small probability within each cell's binomial trial that the cell will leave the pool, as shown in the increased "cell efflux" model in Fig. 2A. This effect was then reduced over time following a certain function and reflecting the bone marrow healing process. A range of "healing" functions was used, with a healing coefficient 0.005 (Fig. 2B) used as standard in simulations where bone marrow recovery function was fixed. The function's initial value (Y-axis in Fig. 2B at X=1) reflects the initial probability of a cell to maintain stemness per trial (lower probability indicates a more degraded niche). In other words, simulations started with the presence of an elevated HSC efflux from the niche (independent of cell phenotype) caused by bone marrow damage, and this effect gradually subsided over time as bone marrow "healed".

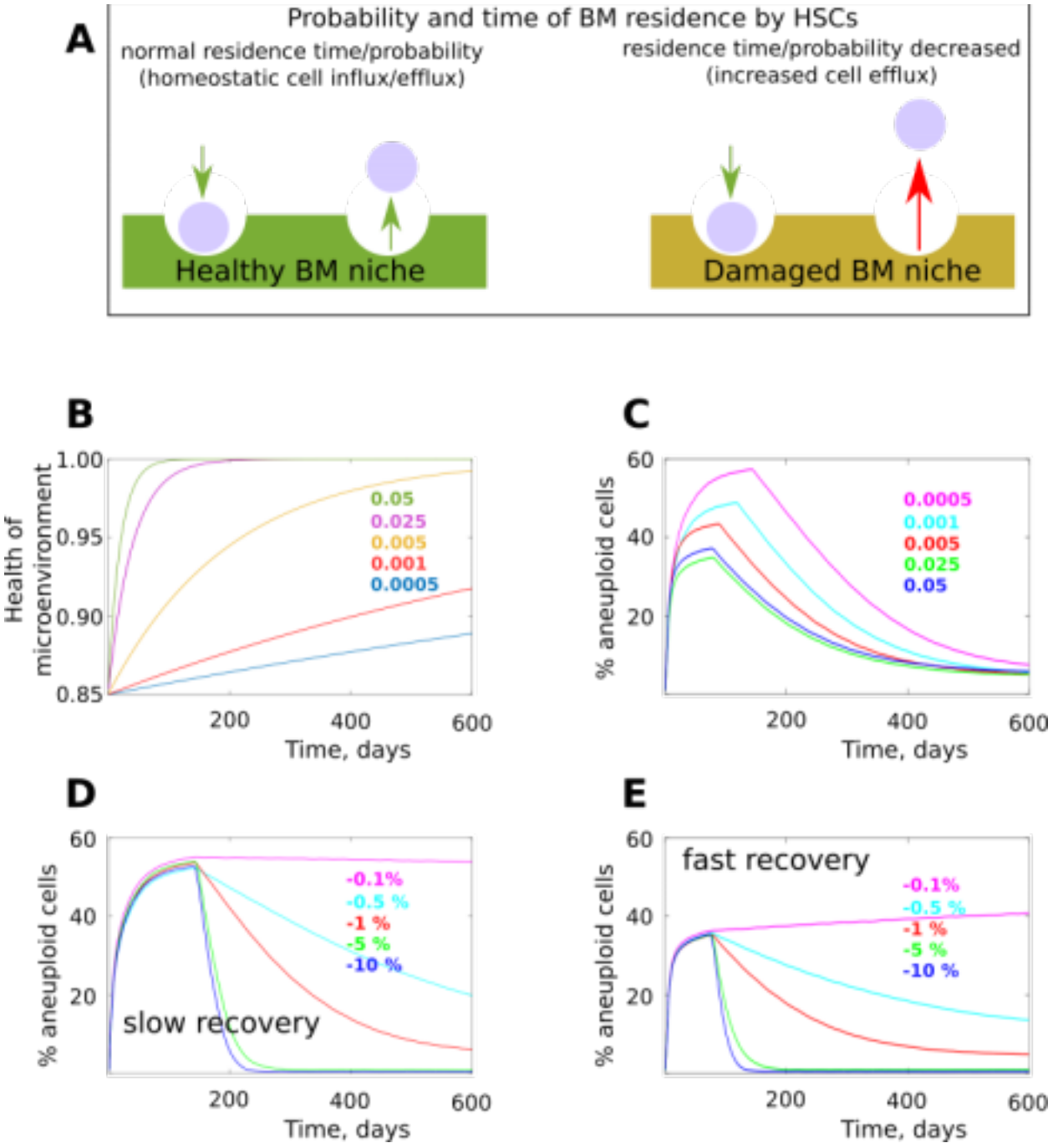


Fig. 2. The model of degraded bone marrow niches and their effect on the aneuploidy dynamics. (A) Damaged bone marrow niche model (BM stands for bone marrow). (B) Temporal profiles of post-radiation bone marrow niche healing; Y at X=0 represents initial bone marrow niche health as a fraction of the maximum health equal to 1; color-matched numbers represent coefficients of healing speed in the function of bone marrow niche health of time; the curve with the coefficient 0.005 was used as standard in simulations were this parameter was not investigated. (C) Simulated aneuploidy dynamics under various profiles of post-radiation bone marrow healing; numbers (healing temporal profile coefficients) indicate as in (B) but color-matched separately (statistics in Fig.S9). (D) Simulated aneuploidy

dynamics with a range of aneuploidy cell fitness cost (color matched numbers) and a slow bone marrow healing profile (coeff=0.0005, according to (B); statistics in Fig.S10). (E) Simulated aneuploidy as in (D) but under a rapid bone marrow healing profile (coeff=0.05, according to (B); statistics in Fig.S11).

In simulations with the niche effect added, we found that the default aneuploidy fitness cost of -10% was more appropriate to use, because the added aneuploidy promoting effect from niche degradation drove aneuploidy peaks to higher frequencies compared to niche-independent modeling, often up to the point of fixation. Notably, the real fitness cost of aneuploidy in HSC is not known and is likely distributed depending on various types of aneuploidy. This fact, however, does not confound the investigation of the general principles underlying the observed aneuploidy peak. Fig. 2C shows that the dynamics of niche healing has a significant effect on the aneuploid cell frequency peak and it also affects the time when the peak is reached. Slow bone marrow healing promotes an increased frequency of aneuploidy cells, showing that damaged bone marrow could counteract purifying selection relative to healthy bone marrow. To further investigate the ability of bone marrow health to impact purifying selection acting on aneuploid cells, we applied a range of aneuploidy fitness costs as indicated previously [- 10%, -5%, -1%, -0.5%, -0.1%], and tested the system under two regimens of niche recovery: slow (coefficient of healing 0.0005 in Fig. 2B) and fast (coefficient of healing 0.05). Comparison of the resulting dynamics demonstrates that the early increase and the maximal aneuploidy frequency reached are essentially unaffected by the aneuploidy fitness cost in the presence of the damaged bone marrow microenvironment, consistent with very weak purifying selection. Notably, in the absence of such bone marrow damage effects, altering the fitness cost of aneuploidy does affect the height of the aneuploidy peak, albeit not proportionally to change in fitness (Fig. 1C).

In contrast, during the later phase of aneuploidy frequency reduction, where purifying

selection is supposed to intensify when the HSC population has stabilized and the bone marrow microenvironment is healthier, we observe a substantial effect of fitness (Fig. 2D,E). After peaking, the frequency of aneuploid cells decreases dependent on the fitness cost of aneuploidy, and the effect of varying microenvironment health is minimal. Comparison of Fig. 2D and Fig. 2E further shows that the rate of bone marrow microenvironmental recovery has a significant effect on the height and timing of the peak frequency of aneuploidy, whereby slow niche recovery promotes aneuploidy and delays the onset of the second (aneuploidy reduction) phase.

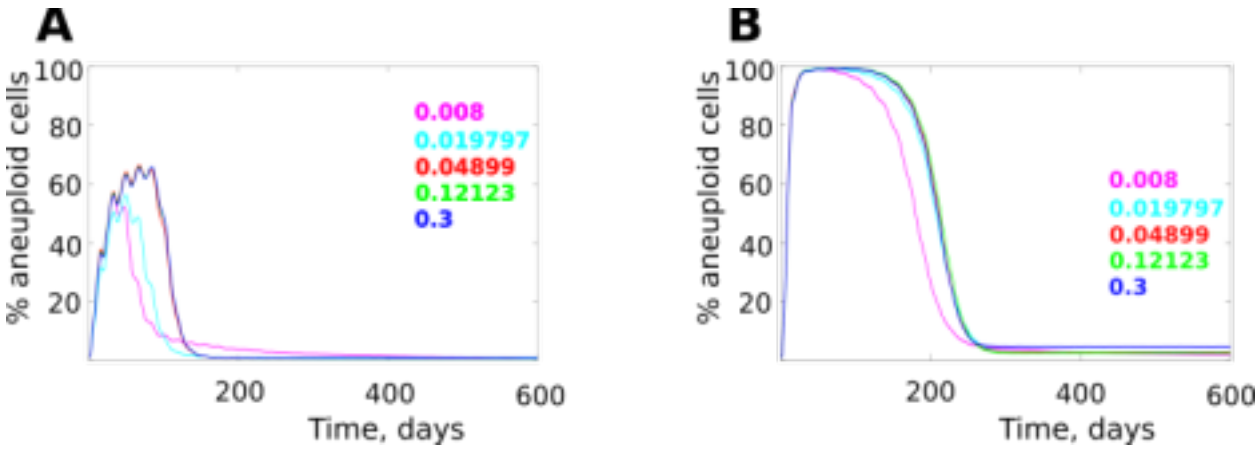


Fig. 3. The effect of bone marrow health and population expansion on simulated aneuploidy dynamics. (A) Less degraded bone marrow (initial niche health 90%; healing profile coeff=0.005, according to Fig. 2B; statistics in **Fig.S12**). **(B)** More degraded bone marrow (initial niche health 70%; healing profile as in (A); statistics in **Fig.S13**). Growth coefficients determining population expansion are varied from 0.008 to 0.3 in according to the scheme in **Fig.1D**.

We further explored the interaction of population expansion (according to the scheme in Fig. 1D) and bone marrow microenvironmental health by simulating the initial bone marrow health of 90% (relatively good) and 70% (more degraded), both under the healing coefficient of 0.005 (see Fig. 2B). Comparison of Fig. 3A and Fig. 3B demonstrates that in a relatively healthy bone marrow microenvironment (Fig. 3A), the population expansion effect is present, but peak

aneuploidy cell frequency is significantly lower than in a degraded bone marrow microenvironment (Fig. 3B). This result demonstrates a profound effect of bone marrow microenvironment degradation in suppressing purifying selection and promoting aneuploidy. More profound damage (Fig. 3B) to bone marrow also reduces the population expansion effect.

Discussion

Our results demonstrate that the pattern of aneuploidy frequency during post transplantation hematopoietic reconstitution observed by Pfau et al. [15] is likely the result of relaxed purifying selection during the early post-transplantation period. We demonstrate that at least two mechanisms could promote the observed aneuploidy frequency peak. First, a rapid expansion of the engrafted HSC population results in a reduction in the strength of somatic purifying selection by introducing an increased presence of drift, resembling the pattern shown in general population biology studies [22–24]. Additionally, pre-transplantation ablative radiation treatment, by damaging and perturbing the bone marrow microenvironment, should degrade the capability of the niche to maintain HSC stemness, which we propose should affect all HSC with reduced or minimal discrimination of cell phenotypes. Our results demonstrate that such niche damage should promote an increase in the frequency of aneuploid cells and that this effect can be strong enough to overcome the contribution of population expansion or the fitness cost of aneuploidy. Recovery of the bone marrow microenvironment over time restores the power of purifying selection, leading to elimination of aneuploid cells from the pool. Our results also demonstrate, consistent with general population biology, that increased cell population numbers act to suppress aneuploidy by reducing the amount of drift and elevating thus the strength of purifying selection. Rapid cell division rates early post transplantation confer the same effect, intensifying

selection by increasing the number of cell generations per time unit.

Previous studies have employed mathematical modeling to explore the parameters that underlie clonal expansions post-transplantation[27–31]. Here, we have explored potential mechanisms that can explain clonal *elimination* post-transplantation. We propose a model to explain changes in aneuploid HSC frequency post-transplantation (Fig. 4). The proposed scenario is that early hematopoietic reconstitution provides a window of opportunity for mutant cell clones of lower fitness, such as the aneuploid cells modeled here. This window is created by exposing transplanted HSCs to conditions of reduced strength of purifying selection. Later, purifying selection regains strength as the bone marrow microenvironment heals and the engrafted HSC population stabilizes at its homeostatic size, leading to elimination of mutant cells from the pool.

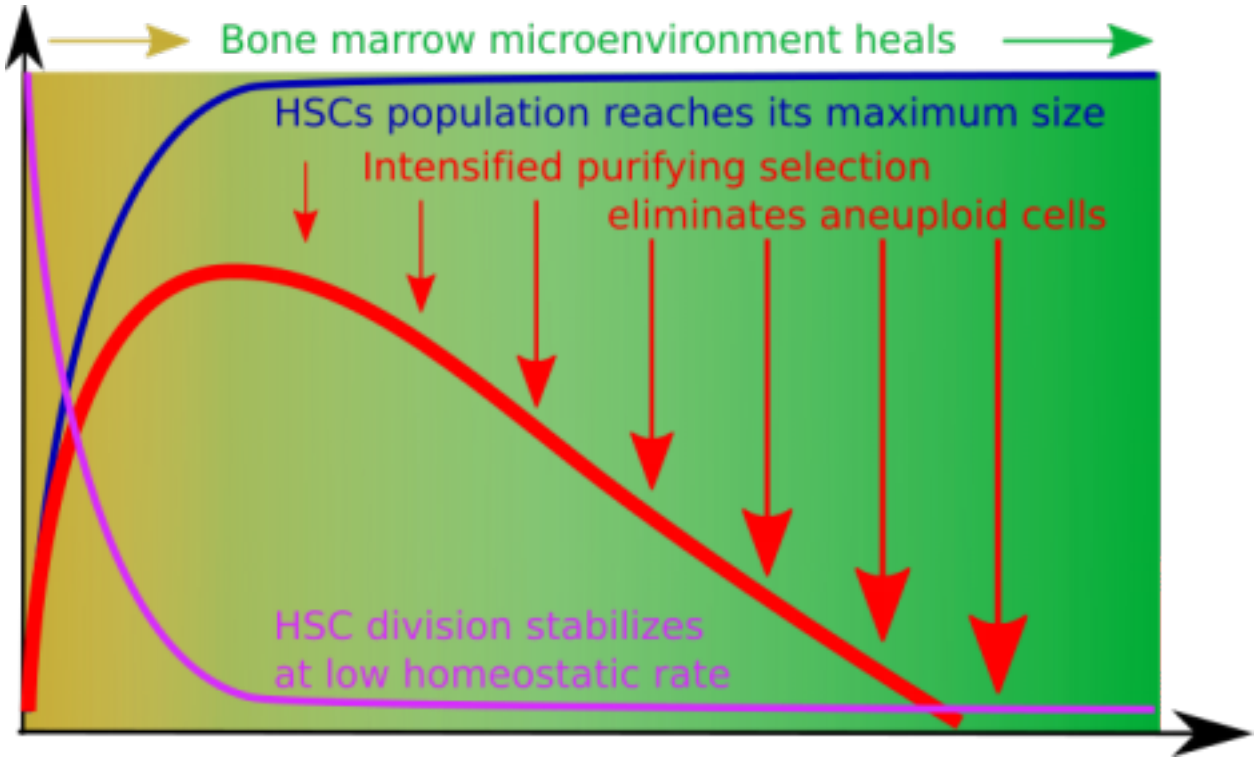


Fig. 4. A model of factors influencing the strength of purifying selection in the post

radiation post-transplantation bone marrow. Early post-transplantation bone marrow is characterized by a highly perturbed microenvironment, as well as rapid HSC population expansion. This combination of factors reduces the strength of purifying selection and promotes drift. Later, HSC population numbers reach their maximum and the bone marrow microenvironment partially restores from radiation damage, processes that intensify the strength of purifying selection and lead to the elimination of aneuploid HSCs from the pool. The rapid cell division rates and the increased HSC population size (per se, excluding the effect of expansion) should act to suppress aneuploidy frequency.

The observations made by Pfau et al. [15] and the mechanisms we propose to explain them may have important implications for cancer research. First of all, bone marrow transplantation is known to be associated with a higher post-transplantation risk of leukemia [32]. This increased risk has been attributed to the compromised immune system [33] and radiation-induced mutations [34–36]. Another mechanism that could contribute, in line with our results, would be conditions of reduced purifying selection that are conducive to greater proliferation of pre-malignant mutant cells (whether from the engrafted bone marrow or residual host cells) due to the perturbed bone marrow microenvironment and cell population expansion. Radiotherapy or chemotherapy, even when not associated with a stem cell transplantation, could similarly reduce purifying selection acting on aneuploid cells (or other cells with mutations reducing somatic fitness), both by depleting stem cell pools and by damaging their tissue niches. These mechanisms should be relevant both for normal and malignant stem cells, contributing to the genetic heterogeneity that can contribute to secondary malignancies and to cancer relapse, respectively.

We have argued earlier that a similar decreased purifying selection might be involved in childhood leukemia [37,38]. Although fetal development likely proceeds in a homeostatic environment, HSC population expands during development, which should promote a greater role for random drift, providing a window of opportunity for the accumulation and expansion of pre-malignant mutant clones that later in life could be purged by purifying selection. Interestingly,

the increased aneuploidy observed during fetal development, which is subsequently purged [39,40], could potentially be similarly explained by relaxing purifying selection associated with rapid tissue growth (although typically with healthy tissue niches). That damaged microenvironments could weaken purifying selection could also be relevant for tissue in the elderly. Increases in expanded hematopoietic clones are relatively common in the elderly [41–43]. While we and others have proposed that these expansions result from increased positive selection for mutations that confer adaptation to the aged bone marrow microenvironment [44,45], results presented here further indicate that relaxed purifying selection in the degraded bone marrow microenvironment could contribute to such somatic evolution by being more permissive to phenotypic variability. Indeed, previous studies reveal that increasing the severity of pre-conditioning prior to HSC transplantation can dramatically promote the expansion of a small number of HSPC clones [46,47]. Notably, aging-related changes likely do not provide a level of acute damage to bone marrow comparable to radiation treatment, but instead likely promote gradual alterations of the bone marrow microenvironment.

In each case, a temporary tolerance of aneuploidy (due to refilling of niches or due to degradation of niches) should increase the chances for a cell with a chromosomal aneuploidy event that would normally reduce fitness to undergo additional chromosomal rearrangements or other mutations that could decrease the cost of the initial aneuploidy. The more expanded the original aneuploidy cells are, the greater the odds of subsequent compensating events, the selective pressure for which should increase as the strength of selection returns following recovery of the stem cell pool and the niche. These chromosomal reassortments could generate more diversity in the population for potential oncogenic selection.

In a growing tumor, the disruption of normal niches could similarly contribute to the

increased presence of aneuploidy, by reducing the cost. From this perspective, the presence of aneuploidy in a cancer can be interpreted as not simply an increased rate of chromosomal gains and losses ("genomic instability"), but as changes in the strength of selection that alter the frequency of aneuploidy cells *independent* of their generation rate. Relaxed purifying selection in cancers, as previously shown [48–51], could even create permissive conditions for cells with large-scale genomic perturbations, such as chromothripsis [52].

The modeling results presented here should stimulate further experimental efforts to decipher the somatic evolutionary processes and mechanisms that govern stem cell dynamics under conditions of rapid population expansion and/or damaged microenvironments.

Methods

Simulations were performed in the Matlab programming environment (MathWorks Inc., MA). The model incorporates a matrix of simulated cells that divide with an average specified frequency determined by the curve of cell division shown in Fig. S1. Cells started from the average cell division rate of once per 3 days and reached the finals rate of once per 40 days [20,21], with the dynamics of change corresponding to the dynamics of population size (the closer to the final size the slower the cells became). Cell division is stochastic and determined by a normal distribution with the mean frequency determined by the curve in Fig. S1 and standard deviation equal to mean/8, as used in [53]. The cell population matrix had a time-specific maximum capacity (cell number) which is determined by the population growth curves as shown in Fig. 1D. The simulation continued for 600 time units (days) and was updated at each time unit.

At each cell division, aneuploid cells were generated with the probability of aneuploidy per

cell division as shown in Fig. 1B (1% in most simulations, unless indicated otherwise). Euploid cells were assigned fitness equal to 1. Aneuploid cells were assigned fitness as shown in Fig. 1C (-1%, or 0.99, in most simulations unless indicated otherwise). Based on cell fitness, the current matrix capacity and the total number of cells after cell divisions, at each simulation update each cell had a probability of leaving the pool (simulating death or differentiation) weighed by their fitness so that the remaining number of HSCs approximately corresponded to the current matrix capacity (fitness based competition for limited space). This probability was realized in corresponding binomial trials for each cell at each simulation update. The effect of niche degradation was realized by adding additional probability (in a binomial trial) for each cell to leave the pool; this probability did not depend on a cell's fitness, replicating a hypothesized effect of phenotype-indiscriminate lower capacity of bone marrow niche to support stemness. The probability was proportional to niche health (the Y-axis value at X=1 in a chart exemplified by Fig. 2B). For example, if the initial bone marrow niche was considered 15% degraded, each cell had an additional probability of 0.15 of leaving the pool at each simulation update. This effect was reduced over time, following the healing curves shown in Fig. 2B, with a corresponding decrease in this probability as the function approached 1 over time (perfect niche, all competition is based exclusively on intrinsic HSC fitness), reflecting bone marrow niche health recovery with time past transplantation.

Simulations started with the initial population size of 100 HSCs, according to data from [15]. All resulting aneuploidy curves are averages of 100 simulation repeats for each condition. The Matlab code for the model is presented in Supplements, section "Model code". We recognized that our results are limited by the accuracy of assumptions made by the model, such as parameter values. In order to factor out the effect of parameter values on the results of the modeling, we

used parameter ranges instead of single values and analyzed the effects of changing a parameter value instead of the effect of a particular value. Our results and conclusions are also contingent on the accuracy of our model for the HSC niche, which is based on the assumption that niche degradation leads in increased propensity of HSC to leave the niche, leading often to their commitment to differentiation.

Statistical comparisons of the simulated clonal dynamics were performed using the Matlab Statistics toolbox. Each simulated condition was run in 100 repeats. In order to elucidate as much statistical information about the relative behavior of clones as possible, we applied the following statistical procedure. At each time point (out of the 600 total simulation time points), we compared different conditions each represented by a sample of 100 runs by the Kruskal-Wallis method, which is a non-parametric version of ANOVA.

The obtained p-values were plotted along the X-axis (simulation time points), with the Y axis representing p-values (see Fig. S2-S13). This procedure allows visualizing the temporal dynamics of the differences in clonal behavior. The general magnitude of the difference in clonal behavior over time in this way can be visualized by the total sum of p values (area under the p-value curve). We calculated this area and divided it by the total area of the chart, the latter being 1x600. The total area represents a hypothetical scenario whereby p-values are equal to 1 during an entire simulation, meaning that the compared behavior of clones was identical throughout the simulation. Respectively, if the area under the p-value curve equals zero, it would mean that such clonal behaviors are totally distinct throughout the simulation time. Realistically, however, p-values always are within that range and never reach such extremes. Therefore, the above-mentioned ratio shown in the top right corner of the chart in Fig. S2-S13, reflects the overall relative magnitude of the difference in clonal behavior throughout the compared simulations.

The smaller the ratio, the greater the overall difference in clonal behavior. Following this statistical procedure, thus, we can demonstrate both the significance of the difference at each time point (p-value curve) and the overall magnitude of the difference throughout the simulation time.

Hematopoietic reconstitutions were performed following the protocol in Pfau *et al.* (2016). Briefly, B6.SJL-Ptprc^{a}Pepc^{b}/BoyJ (CD45.1) female mice were purchased from Jackson Laboratory and served as recipients for all reconstitutions. At 6-8 weeks old, recipients were treated with 9.5 Gy in a single dose, administered via 137^{Cs} irradiator (γ cell 40) at a dose rate of ~ 100 cGy/min, and were intravenously injected with 10^{6} donor (CD45.2) cells. In bone marrow reconstitutions, donor bone marrow cells were isolated from a BubR1$^{H/H}$ donor at 5-7 weeks old. Red blood cells were lysed in ACK lysing buffer. The remaining white blood cells were counted on a Cellometer Auto T4 automated hemacytometer (Nexelcom) and injected in Hank's balanced salt solution (HBSS). Fetal livers were harvested from BubR1$^{H/H}$ E14.5 embryos, homogenized by pipetting, passed through a 70 μM cell strainer, and frozen in FBS + 5% dimethylsulfoxide in liquid nitrogen. On the day of reconstitution, fetal liver cells were thawed in Iscove's modified Dulbecco's medium (IMDM) supplemented with 2% FBS. Viability was assessed by propidium iodide staining with a FACSCalibur flow cytometer (Becton Dickinson). Live cells were then counted and injected in HBSS. All BubR1$^{H/H}$ animals were genotyped by Transnetyx, following protocols described in Baker *et al.* (2004).

Peripheral blood was collected with heparinized capillary tubes into sodium heparin diluted in PBS. Following the lysing of red blood cells in ACK lysing buffer, cells were incubated with an anti-CD45.1 antibody, obtained from Biolegend (A20), per the manufacturer's specifications. CD45.1-negative cells were isolated via Aria I cell sorter (Beckerson Dickinson). These while

blood cells were analyzed by single-cell sequencing, following the protocol in Knouse *et al.* (2014).

Approval of Animal Experimentation

All animal experimentation was carried out in accordance with relevant guidelines and regulations, and the Institutional Animal Care and Use Committee at the Massachusetts Institute of Technology approved all mouse experiments.

References

1. Knouse KA, Davoli T, Elledge SJ, Amon A. Aneuploidy in Cancer: Seq-ing Answers to Old Questions. *Annu Rev Cancer Biol*. 2017;1(1):335-354. doi:10.1146/annurev-cancerbio-042616-072231

2. Gordon DJ, Resio B, Pellman D. Causes and consequences of aneuploidy in cancer. *Nat Rev Genet*. 2012;13(3):189-203. doi:10.1038/nrg3123

3. Rajagopalan H, Lengauer C. Aneuploidy and cancer. *Nature*. 2004;432(7015):338-341. doi:10.1038/nature03099

4. Taylor AM, Shih J, Ha G, et al. Genomic and Functional Approaches to Understanding Cancer Aneuploidy. *Cancer Cell*. 2018;33(4):676-689.e3. doi:10.1016/j.ccell.2018.03.007

5. Santaguida S, Amon A. Short- and long-term effects of chromosome mis segregation and aneuploidy. *Nat Rev Mol Cell Biol*. 2015;16(8):473-485. doi:10.1038/nrm4025

6. Chen G, Mulla WA, Kucharavy A, et al. Targeting the adaptability of heterogeneous aneuploids. *Cell*. 2015;160(4):771-784. doi:10.1016/j.cell.2015.01.026

7. Pavelka N, Rancati G, Zhu J, et al. Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. *Nature*. 2010;468(7321):321-325. doi:10.1038/nature09529

8. Rancati G, Pavelka N, Fleharty B, et al. Aneuploidy Underlies Rapid Adaptive Evolution of Yeast Cells Deprived of a Conserved Cytokinesis Motor. *Cell*. 2008;135(5):879-893. doi:10.1016/j.cell.2008.09.039

9. Sansregret L, Swanton C. The role of aneuploidy in cancer evolution. *Cold Spring Harb Perspect Med*. 2017;7(1). doi:10.1101/cshperspect.a028373

10. Davoli T, Xu AW, Mengwasser KE, et al. Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome. *Cell*. 2013;155(4):948-962. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=24183448.

11. Sotillo R, Schvartzman J-M, Socci ND, Benezra R. Mad2-induced chromosome instability leads to lung tumour relapse after oncogene withdrawal. *Nature*. 2010;464(7287):436-440. doi:10.1038/nature08803

12. Torres EM, Sokolsky T, Tucker CM, et al. Effects of aneuploidy on cellular physiology and cell division in haploid yeast. *Science*. 2007;317(5840):916- 924. doi:10.1126/science.1142210

13. Oromendia AB, Dodgson SE, Amon A. Aneuploidy causes proteotoxic stress in yeast. *Genes Dev*. 2012;26(24):2696-2708. doi:10.1101/gad.207407.112

14. Adorno M, Sikandar S, Mitra SS, et al. Usp16 contributes to somatic stem-cell defects in Down's syndrome. *Nature*. 2013;501(7467):380-384. doi:10.1038/nature12530

15. Pfau SJ, Silberman RE, Knouse KA, Amon A. Aneuploidy impairs hematopoietic stem cell fitness and is selected against in regenerating tissues in vivo. *Genes Dev*. 2016;30(12):1395-1408. doi:10.1101/gad.278820.116

16. Gogendeau D, Siudeja K, Gambarotto D, Pennetier C, Bardin AJ, Basto R. Aneuploidy causes premature differentiation of neural and intestinal stem cells. *Nat Commun*. 2015;6. doi:10.1038/ncomms9894

17. Chua HL, Anand V, Sampson C, et al. Kinetics of Self-Renewal and Differentiation Divisions of Transplanted Hematopoietic Stem and Progenitor Cells Early After Transplantation Into Lethally-Irradiated Recipients. *Blood*. 2009;114(22). http://www.bloodjournal.org/content/114/22/2443?sso-checked=true. Accessed April 5, 2018.

18. Cao X, Wu X, Frassica D, et al. Irradiation induces bone injury by damaging bone marrow microenvironment for stem cells. *Proc Natl Acad Sci*. 2011;108(4):1609-1614. doi:10.1073/pnas.1015350108

19. Sasi SP, Park D, Muralidharan S, et al. Particle Radiation-Induced Nontargeted Effects in Bone-Marrow-

Derived Endothelial Progenitor Cells. *Stem Cells Int*. 2015;2015. doi:10.1155/2015/496512

20. Abkowitz JL, Catlin SN, McCallie MT, Guttorp P. Evidence that the number of hematopoietic stem cells per animal is conserved in mammals. *Blood*. 2002;100(7):2665-2667. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12239184.

21. Abkowitz JL, Golinelli D, Harrison DE, Guttorp P. In vivo kinetics of murine hemopoietic stem cells. *Blood*. 2000;96(10):3399-3405. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11071634.

22. CARSON, L. H. The population flush and its genetic consequences. In: Lewontin R, ed. *Population Biology and Evolution*. Syracuse University Press; 1968:123-137. https://ci.nii.ac.jp/naid/10004957303/. Accessed April 5, 2018.

23. Jones JS. Models of speciation - The evidence from Drosophila. *Nature*. 1981;289(5800):743-744. doi:10.1038/289743a0

24. Lints F, Bourgois M. Population crash, population flush and genetic variability in cage populations of Drosophila melanogaster. *Genet Sel Evol*. 1984;16(1):45. doi:10.1186/1297-9686-16-1-45

25. Rozhok AIAI, DeGregori J. The Evolution of Lifespan and Age-Dependent Cancer Risk. *Trends in Cancer*. 2016;2(10):552-560. doi:10.1016/j.trecan.2016.09.004

26. Morrison SJ, Scadden DT. The bone marrow niche for haematopoietic stem cells. *Nature*. 2014;505(7483):327-334. doi:10.1038/nature12984

27. Becker NB, Günther M, Li C, Jolly A, Höfer T. Stem cell homeostasis by integral feedback through the niche. *J Theor Biol*. 2019;481:100-109. doi:10.1016/j.jtbi.2018.12.029

28. Ashcroft P, Manz MG, Bonhoeffer S. Clonal dominance and transplantation dynamics in hematopoietic stem cell compartments. *PLoS Comput Biol*. 2017;13(10). doi:10.1371/journal.pcbi.1005803

29. Park DS, Akuffo AA, Muench DE, et al. Clonal hematopoiesis of indeterminate potential and its impact on patient trajectories after stem cell transplantation. Komarova NL, ed. *PLOS Comput Biol*. 2019;15(4):e1006913. doi:10.1371/journal.pcbi.1006913

30. Stiehl T, Baran N, Ho AD, Marciniak-Czochra A. Clonal selection and therapy resistance in acute leukaemias: Mathematical modelling explains different proliferation patterns at diagnosis and relapse. *J R Soc Interface*. 2014;11(94). doi:10.1098/rsif.2014.0079

31. Manesso E, Teles J, Bryder D, Peterson C. Dynamical modelling of haematopoiesis: An integrated view over the system in homeostasis and under perturbation. *J R Soc Interface*. 2013;10(80). doi:10.1098/rsif.2012.0817

32. Bhatia S, Ramsay NK, Steinbuch M, et al. Malignant neoplasms following bone marrow transplantation. *Blood*. 1996;87(9):3633-3639. http://www.ncbi.nlm.nih.gov/pubmed/8611687.

33. Curtis RE, Rowlings PA, Deeg HJ, et al. Solid cancers after bone marrow transplantation. *N Engl J Med*. 1997;336(13):897-904. doi:10.1056/NEJM199703273361301

34. Kusunoki Y, Hayashi T. Long-lasting alterations of the immune system by ionizing radiation exposure: Implications for disease development among atomic bomb survivors. *Int J Radiat Biol*. 2008;84(1):1-14. doi:10.1080/09553000701616106

35. Little JB. Radiation carcinogenesis. *Carcinogenesis*. 2000;21(3):397-404. doi:10.1093/carcin/21.3.397

36. Yahata T, Takanashi T, Muguruma Y, et al. Accumulation of oxidative DNA damage restricts the self-renewal capacity of human hematopoietic stem cells. *Blood*. 2011;118(11):2941-2950. doi:10.1182/blood-2011-01-330050

37. Rozhok AIAI, Salstrom JLJL, DeGregori J. Stochastic modeling reveals an evolutionary mechanism

underlying elevated rates of childhood leukemia. *Proc Natl Acad Sci*. 2016;113(4):1050-1055. doi:10.1073/pnas.1509333113

38. Rozhok A, DeGregori J. A generalized theory of age-dependent carcinogenesis. *Elife*. 2019;8. doi:10.7554/eLife.39950

39. De S. Somatic mosaicism in healthy human tissues. *Trends Genet*. 2011;27(6):217-223. doi:10.1016/j.tig.2011.03.002

40. Bolton H, Graham SJL, Van der Aa N, et al. Mouse model of chromosome mosaicism reveals lineage-specific depletion of aneuploid cells and normal developmental potential. *Nat Commun*. 2016;7(1):11165. doi:10.1038/ncomms11165

41. Genovese G, Kahler AK, Handsaker RE, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med*. 2014;371(26):2477-2487. doi:10.1056/NEJMoa1409405

42. Jaiswal S, Fontanillas P, Flannick J, et al. Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med*. 2014;371(26):2488-2498. doi:10.1056/NEJMoa1408617

43. McKerrell T, Park N, Moreno T, et al. Leukemia-associated somatic mutations drive distinct patterns of age-related clonal hemopoiesis. *Cell Rep*. 2015;10(8):1239-1245. doi:10.1016/j.celrep.2015.02.005

44. Rozhok AII, DeGregori J. Toward an evolutionary model of cancer: Considering the mechanisms that govern the fate of somatic mutations. *Proc Natl Acad Sci U S A*. 2015;112(29):8914-8921. doi:10.1073/pnas.1501713112

45. McKerrell T, Vassiliou GS. Aging as a driver of leukemogenesis. *Sci Transl Med*. 2015;7(306). doi:10.1126/scitranslmed.aac4428

46. Lu R, Czechowicz A, Seita J, Jiang D, Weissman IL. Clonal-level lineage commitment pathways of hematopoietic stem cells in vivo. *Proc Natl Acad Sci*. 2019;116(4):1447-1456. doi:10.1073/pnas.1801480116

47. Sun J, Ramos A, Chapman B, et al. Clonal dynamics of native haematopoiesis. *Nature*. 2014;514(7522):322-327. doi:10.1038/nature13824

48. Yadav VK, Degregori J, De S. The landscape of somatic mutations in protein coding genes in apparently benign human tissues carries signatures of relaxed purifying selection. *Nucleic Acids Res*. 2016;44(5):2075-2084. doi:10.1093/nar/gkw086

49. Ling S, Hu Z, Yang Z, et al. Extremely high genetic diversity in a single tumor points to prevalence of non-Darwinian cell evolution. *Proc Natl Acad Sci U S A*. 2015;112(47):E6496-505. doi:10.1073/pnas.1519556112

50. Martincorena I, Raine KM, Gerstung M, et al. Universal Patterns of Selection in Cancer and Somatic Tissues. *Cell*. 2017;171(5):1029-1041.e21. doi:10.1016/j.cell.2017.09.042

51. Lipinski KA, Barber LJ, Davies MN, Ashenden M, Sottoriva A, Gerlinger M. Cancer Evolution and the Limits of Predictability in Precision Cancer Medicine. *Trends in cancer*. 2016;2(1):49-63. doi:10.1016/j.trecan.2015.11.003

52. Forment J V., Kaidi A, Jackson SP. Chromothripsis and cancer: Causes and consequences of chromosome shattering. *Nat Rev Cancer*. 2012;12(10):663- 670. doi:10.1038/nrc3352

53. Rozhok AI, Salstrom JL, DeGregori J. Stochastic modeling indicates that aging and somatic evolution in the hematopoietic system are driven by non-cell autonomous processes. *Aging (Albany NY)*. 2014;6(12):e1-16.

Supplemental Materials

Model code

```
hold off

% 1. Change 2nd var in clearvars to the titrated
% 2. Change a)nConds + b)the 1st var in the c-loop to
the titrated % 3. Check all the vars just below

clearvars legs

cellDivRateRange = [3, 10, 20, 40, 60];
initPopSizeRange = [1000 5000 10000 25000 50000];
maxPopSizeRange = [1000 5000 10000 25000 50000];
aneuFitRange = [-0.001 -0.005 -0.01 -0.05 -0.1];
aneuRateRange = [0.005 0.01 0.02 0.03 0.04];
expRange = [0.008 0.019796928 0.048989795
0.12123093 0.3]; nicheEffectRange = [0.7 0.75
0.8 0.85 0.9];
nicheEffectIncRange = [0.0005 0.001 0.005 0.025 0.05];
%>>>>>>>>>>>>>>>>>>>
nConds = 5;
%>>>>>>>>>>>>>>>>>>>

colorstring = 'mcrgbk';
legs = [];
legsLength = 145;

posTracker = 0;
% >>>>>>GLOBAL STORAGE>>>>>>>>>>>
globalAneu = [];
globalHSC = [];
globalDivStore = [];
% <<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<

for c = 1 : nConds

    cellDivRate = cellDivRateRange(3);
    initPopSize = initPopSizeRange(1);
    maxPopSize = maxPopSizeRange(3);
    aneuFit = aneuFitRange(5);
    aneuRate = aneuRateRange(2);
    expCoeff = expRange(c);
    nicheEffect = nicheEffectRange(1);
    nicheEffectInc = nicheEffectIncRange(3);
    %$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$4

    clearvars HSCpool
    timer = tic;

    % initial parameters:
    runs = 100; % number of simulation repeats
    maxPool = maxPopSize; % maximum capacity of the pool (max
    number of HSC) maxTime = 600; % maximum age (99 years)
    initHSC = initPopSize;
    %growthRate = 57;
    cellDivInit = 3;
    cellDivFin = 40;

    %===BEST==
    % aneuRate = 0.005;
    % aneuFit = -0.01;
    % nicheEffect = 0.83;
    % nicheEffectInc = 0.005;
    %=========

    % STORAGE MATRICES ---------------------------------
    HSCs = [];
    aneuStore = [];
    cellDivStore = [];
    % -----------------------------------------------------
```

```
    %POOL CAPACITY WITH AGE (%-barred code is an alternative for HSC pool size changing over
    lifetime) % growthCurve = [initHSC];
    % for i = 2 : maxTime
    % %range = 0.3 0.12123093 0.048989795 0.019796928 0.008
    % growthCurve(i) = growthCurve(i-1) + 0.5*growthRate*(1 - (growthCurve(i-
    1)/maxPool)); % end

    oldCapacity = initHSC;
    growthCurve = [initHSC];
    for i = 1 : maxTime
        %range = 0.3 0.12123093 0.048989795 0.019796928 0.008
        newCapacity = oldCapacity + (round(expCoeff * oldCapacity * ((maxPool -
oldCapacity) / maxPool)));
        growthCurve = [growthCurve, newCapacity];
        oldCapacity = newCapacity;
    end

    for run = 1 : runs % this loop contains the whole run of the program for each
individual
        disp(run)%------------------------------------------ for visualization only

        %//// INITIAL CELL POOL
        ////////////////////////////////////////////////////////////// %-- 1. cell
        IDs
        %-- 2. cell ages
        %-- 3. fitness
        HSCpool = [];
        initAneu(1, 1:initHSC) = binornd(1, aneuRate, 1, initHSC); % aneuploidy trial
        %* plugging in initial aneuploid cells by IDS
        HSCpool(1, initAneu(1, 1:initHSC) == 0) = 1; % initial normal cells
        HSCpool(1, initAneu(1, 1:initHSC) == 1) = 2; % initial aneuploid cells
        %* cell ages
        HSCpool(2, 1:initHSC) = randi([2, cellDivInit], 1, initHSC); % initial
        cell ages %* normal and aneuploid cell fitness
        HSCpool(3, initAneu(1, 1:initHSC) == 0) = 1; % fitness of normal cells
        HSCpool(3, initAneu(1, 1:initHSC) == 1) = 1 + aneuFit; % fitness of
        aneuploid cells
        %\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\

        nicheEffectLocal = nicheEffect;
        nicheEffectLocalInc = nicheEffectInc;

        for currentTime = 1 : maxTime

            %disp('current time')
            %disp(currentTime)

            %cellDiv = cellDivInit + ((cellDivFin-cellDivInit)*(abs(size(HSCpool, 2)-
initHSC)/abs(maxPool-initHSC))); %proportional to expansion
            cellDiv = cellDivRate;

            % data collection
            HSCs(run, currentTime) = size(HSCpool, 2);
            aneuStore(run, currentTime) = numel(HSCpool(1, HSCpool(1,
            :)==2))/size(HSCpool, 2)*100; cellDivStore(run, currentTime) = cellDiv;
            % ---------------



            % CELL DIVISIONS

            divMatrix = normrnd(cellDiv, cellDiv/8, 1, size(HSCpool, 2));

            HSCpool(2, HSCpool(2, :) >= divMatrix(1, :)) = 1; % resets the time past division
for cells that divide
            newCells = HSCpool(:, HSCpool(2, :) == 1); %new cells
            HSCpool(:, HSCpool(2, :) == 1) = 0;
            HSCpool = (HSCpool(:, HSCpool(1, :) > 0));

            %== Aneuploidy
            currAneu(1, 1: size(newCells, 2)) = binornd(1, aneuRate, 1,
            size(newCells, 2)); newCells(1, currAneu(1, :) == 1) = 2;
            newCells(3, currAneu(1, :) == 1) = 1 + aneuFit;
            %==
```

```
            %NEW CELLS ADDED TO THE POOL
            HSCpool = [HSCpool, newCells, newCells];

            %CELLS COMPETE FOR SPACE
            % overkill = size(HSCpool, 2)/growthCurve(currentTime);
            % invs = 1./HSCpool(3, :);
            % difprobs = invs/sum(invs);
            % difprobs = difprobs-(mean(difprobs));
            % difprobs = difprobs+(1-(1/overkill));
            % difprobs(difprobs < 0) = 0;
            % difprobs(difprobs > 1) = 1;
            % differen = binornd(1, difprobs(1, :));
            % HSCpool(:, differen(1, :) == 1) = 0;
            % HSCpool = HSCpool(:, HSCpool(1, :) > 0);

            w_prob = find(HSCpool(1, :));
            n = size(w_prob,2);
            w_prob(2, :) = (growthCurve(currentTime) / n) * (n * HSCpool(3,
w_prob(1, :)) / sum(HSCpool(3, w_prob(1, :)))); %weighed probabilities
            w_prob(2, w_prob(2,:) >= 1) = 1;
            w_prob(3, :) = binornd(1, w_prob(2, :)); % all cells compete for
            free niches HSCpool(:, w_prob(1, w_prob(3, :) == 0)) = 0; % all
            lost cells zeroed
            w_prob = [];
            HSCpool = HSCpool(:, HSCpool(1, :) > 0);


            if(isnumeric(nicheEffect) & isnumeric(nicheEffectInc))
                %disp(nicheEffect);
                %NICHE EFFECT
                w_prob = find(HSCpool(1, :));
                n = size(w_prob,2);
                w_prob(2, :) = nicheEffectLocal;
                w_prob(2, w_prob(2,:) >= 1) = 1;
                w_prob(3, :) = binornd(1, w_prob(2, :)); % all cells compete for
                free niches HSCpool(:, w_prob(1, w_prob(3, :) == 0)) = 0; % all
                lost cells zeroed
                %w_prob = [];
                HSCpool = HSCpool(:, HSCpool(1, :) > 0);
                nicheEffectLocal = nicheEffectLocal + (1 - nicheEffectLocal)*(1-exp(-
nicheEffectLocalInc*currentTime));
                nicheEffectStore(run, currentTime) = nicheEffectLocal;
            end



            % CELL AGES UPDATED
            HSCpool(2, :) = HSCpool(2, :) + 1;
        end
        % MEMORY CLEANUP
        newCells = [];

        % time = toc(timer);
        % hours = floor(time / 3600);
        % time = time - hours * 3600;
        % mins = floor(time / 60);
        % secs = time - mins * 60;
        % secs = round(secs);
        %
        % fprintf('Execution time (HH:MM:SS) - %d:%d:%d \n\n', hours, mins, secs);

    end

    % writing into global storage
    globalAneu = [globalAneu; aneuStore];
    globalDivStore = [globalDivStore; cellDivStore];
    % ============================


    for i = 1 : size(aneuStore, 2)
        averAneuStore(i) = mean(aneuStore(:, i));
    end

    if(isnumeric(nicheEffect) & isnumeric(nicheEffectInc))
        nicheEffectLocal = nicheEffect;
    end
```

```
    plot([1:maxTime], averAneuStore, 'Color', colorstring(c))
    %plot([1:maxTime], HSCs, 'Color', colorstring(c))
    hold on
    varToStr = @(x) inputname(1);
     leg = strcat(varToStr(aneuFit), '=', num2str(aneuFit),
     ':::', ... varToStr(aneuRate), '=', num2str(aneuRate),
                           ':::', ...
        varToStr(expCoeff), '=', num2str(expCoeff), ':::', ...
        varToStr(nicheEffect), '=', num2str(nicheEffectLocal), ':::',
        ... varToStr(nicheEffectInc), '=',
        num2str(nicheEffectLocalInc), ':::', ... varToStr(initHSC),
        '=', num2str(initHSC), ':::', ...
        varToStr(maxPool), '=', num2str(maxPool), ':::', ...
        varToStr(cellDivRate), '=', num2str(cellDivRate), ':::' ...
        );
    remn(1:legsLength-size(leg, 2)) = '<';

    legs = [legs; [leg, remn]];
    remn=[];

    time = toc(timer);
    hours = floor(time / 3600);
    time = time - hours * 3600;
    mins = floor(time / 60);
    secs = time - mins * 60;
    secs = round(secs);

    fprintf('Execution time (HH:MM:SS) - %d:%d:%d \n\n', hours, mins,
secs);
end

legend('show')
```

Figures

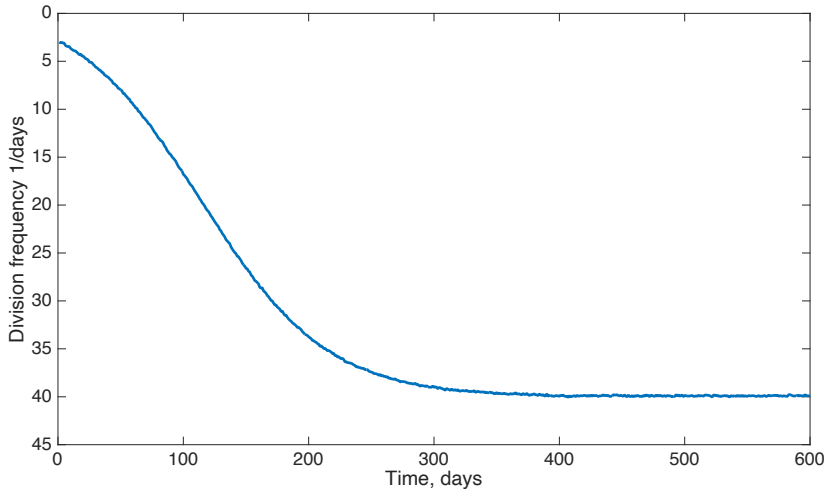


Fig. S1. Dynamics of changing average cell division rate over time (typical example)

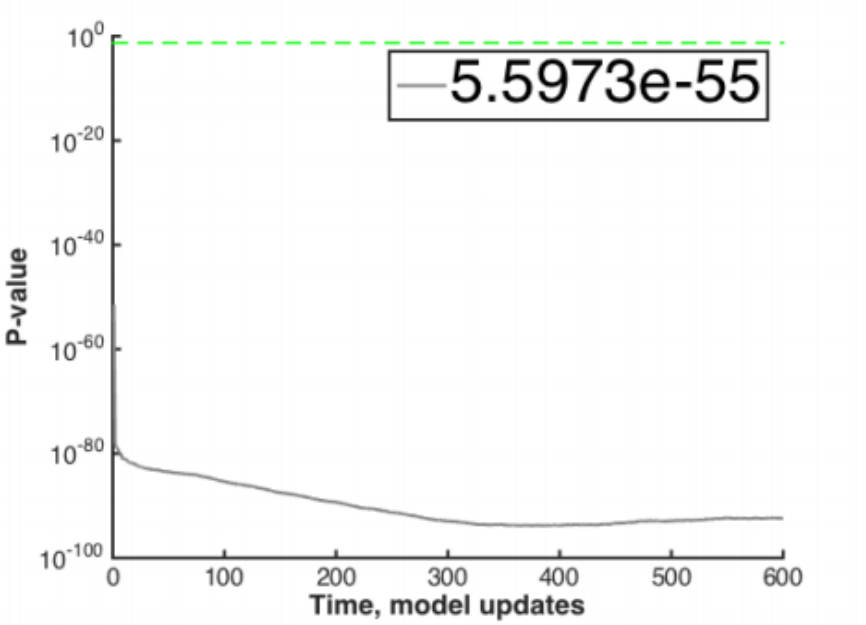


Fig. S2. Kruskal-Wallis analysis of the differences between the processes shown Fig. 1B. At each simulation time point, a p-value (Y-axis) was calculated by comparing the modeled processes each represented by a sample of 100 repeated runs using the Kruskal-Wallis method

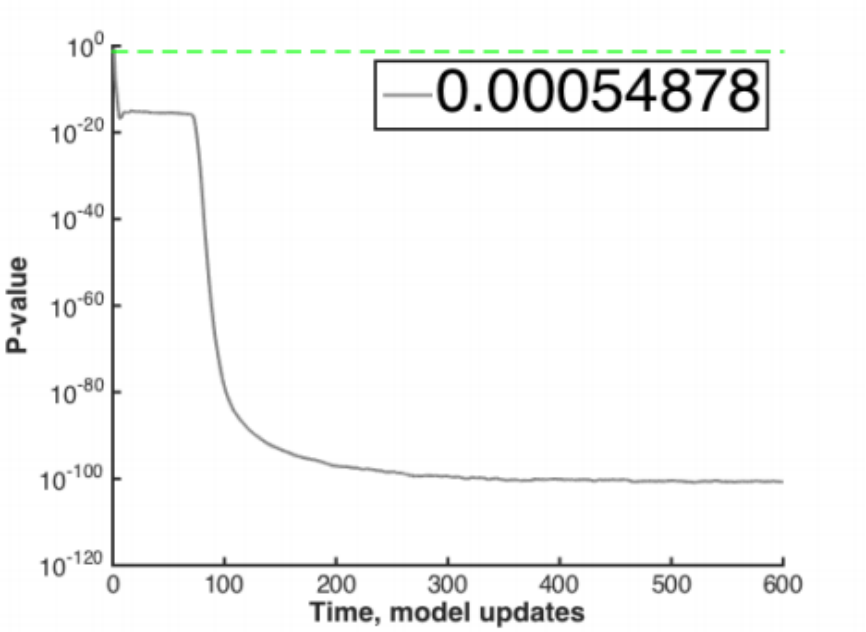


Fig. S3. Kruskal-Wallis analysis of the differences between the processes shown Fig. 1C. At each simulation time point, a p-value (Y-axis) was calculated by comparing the modeled processes each represented by a sample of 100 repeated runs using the Kruskal-Wallis method

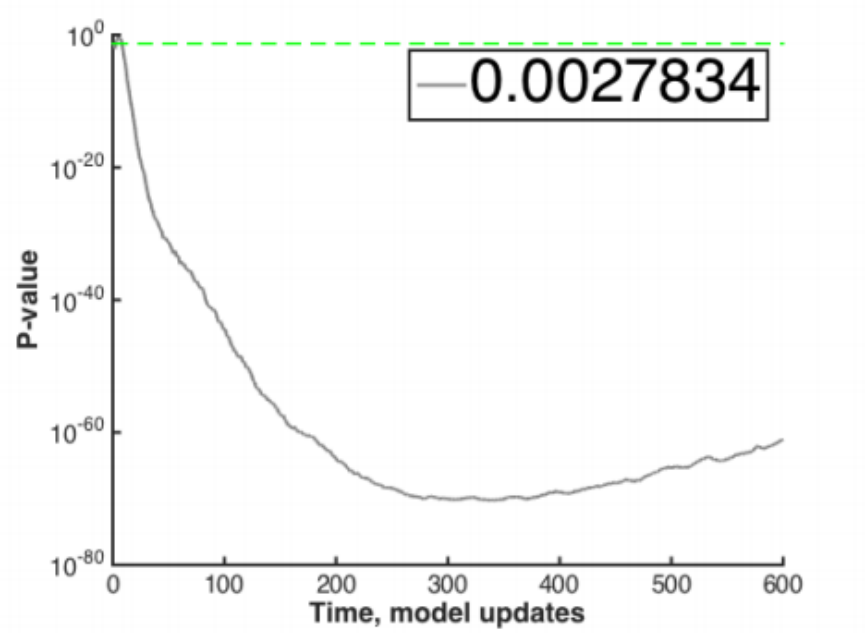


Fig. S4. Kruskal-Wallis analysis of the differences between the processes shown Fig. 1E. At each simulation time point, a p-value (Y-axis) was calculated by comparing the modeled processes each represented by a sample of 100 repeated runs using the Kruskal-Wallis method

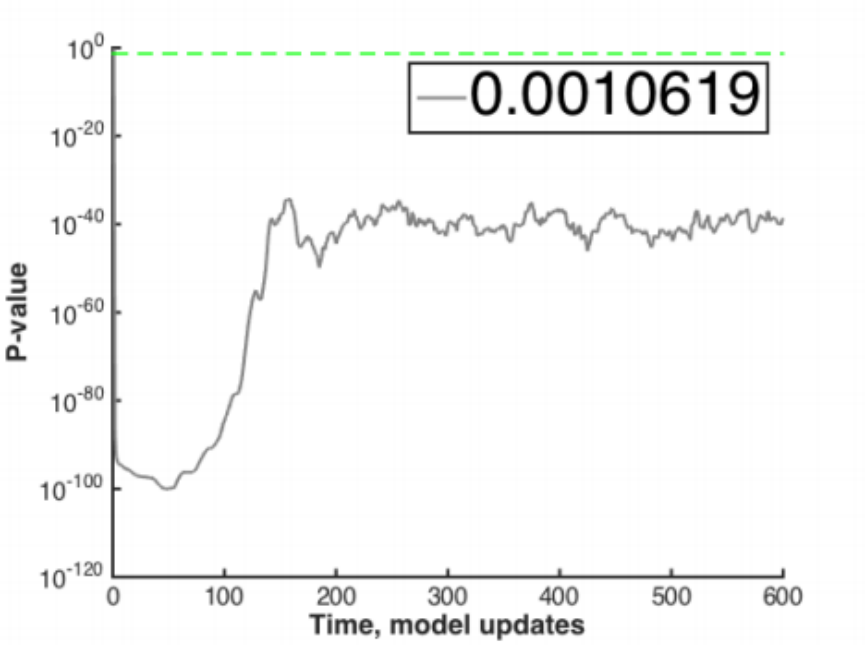


Fig. S5. Kruskal-Wallis analysis of the differences between the processes shown Fig. 1F. At each simulation time point, a p-value (Y-axis) was calculated by comparing the modeled processes each represented by a sample of 100 repeated runs using the Kruskal-Wallis method

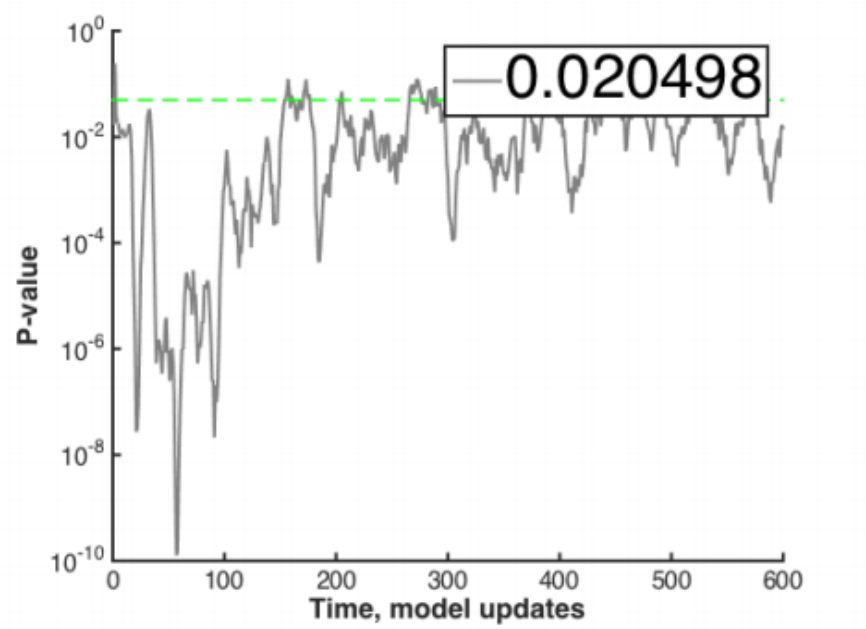


Fig. S6. Kruskal-Wallis analysis of the differences between the processes shown Fig. 1G. At each simulation time point, a p-value (Y-axis) was calculated by comparing the modeled processes each represented by a sample of 100 repeated runs using the Kruskal-Wallis method

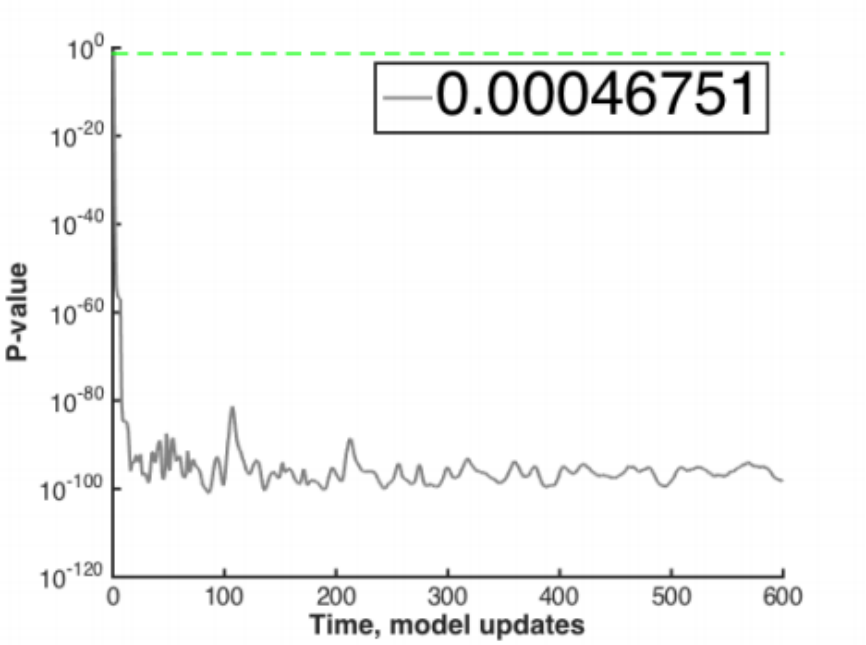


Fig. S7. Kruskal-Wallis analysis of the differences between the processes shown Fig. 1H. At each simulation time point, a p-value (Y-axis) was calculated by comparing the modeled processes each represented by a sample of 100 repeated runs using the Kruskal-Wallis method

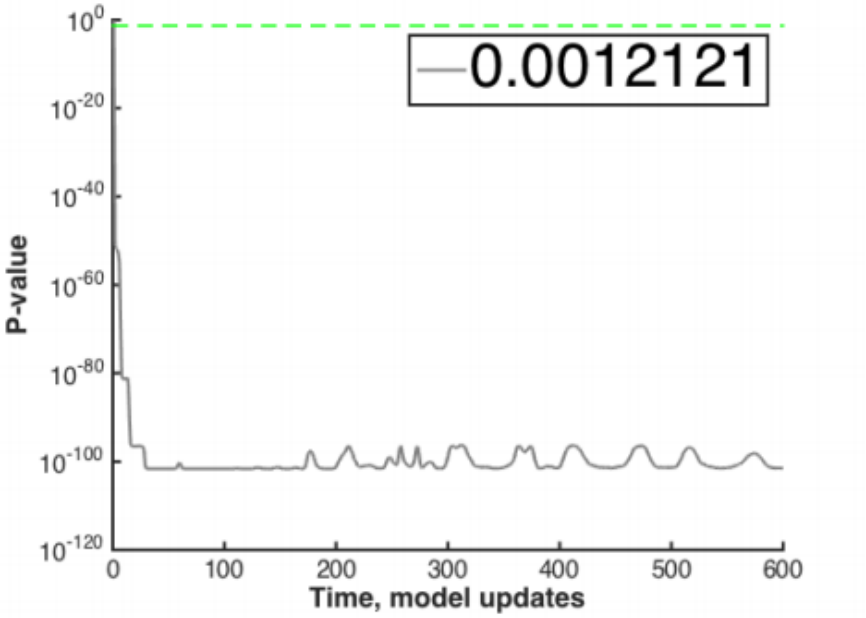


Fig. S8. Kruskal-Wallis analysis of the differences between the processes shown Fig. 1I. At each simulation time point, a p-value (Y-axis) was calculated by comparing the modeled processes each represented by a sample of 100 repeated runs using the Kruskal-Wallis method

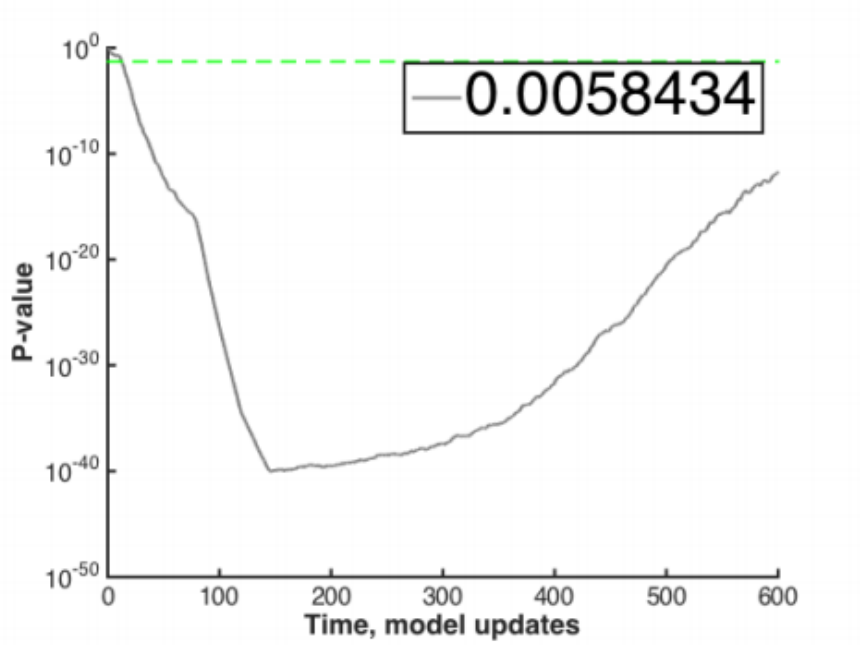


Fig. S9. Kruskal-Wallis analysis of the differences between the processes shown Fig. 2C. At each simulation time point, a p-value (Y-axis) was calculated by comparing the modeled processes each represented by a sample of 100 repeated runs using the Kruskal-Wallis method

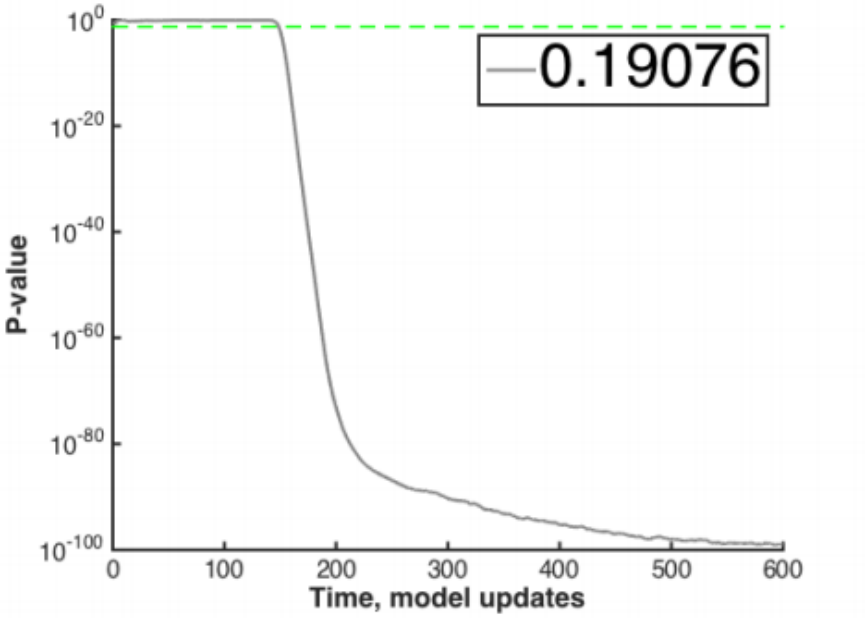


Fig. S10. Kruskal-Wallis analysis of the differences between the processes shown Fig. 2D. At each simulation time point, a p-value (Y-axis) was calculated by comparing the modeled processes each represented by a sample of 100 repeated runs using the Kruskal Wallis method

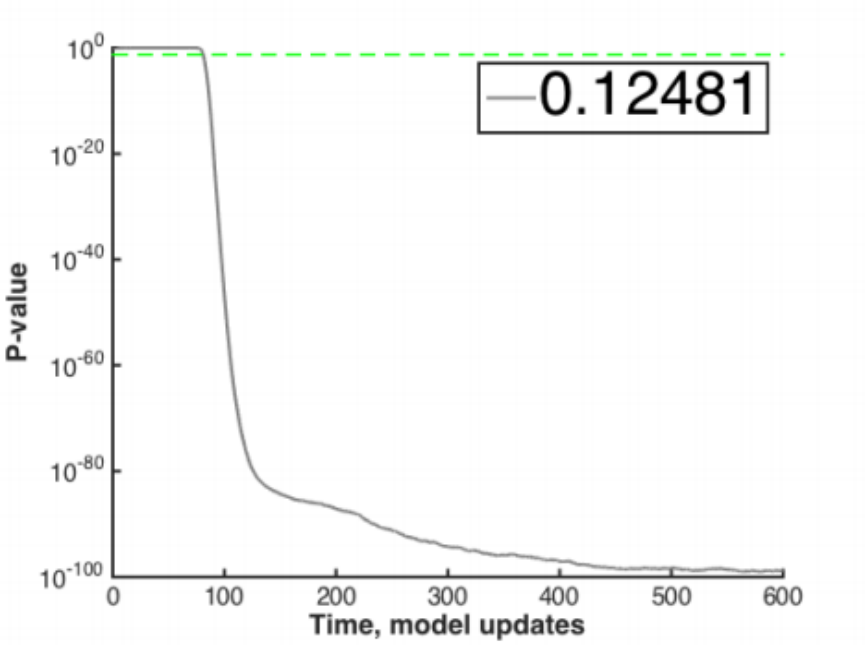


Fig. S11. Kruskal-Wallis analysis of the differences between the processes shown Fig. 2E. At each simulation time point, a p-value (Y-axis) was calculated by comparing the modeled processes each represented by a sample of 100 repeated runs using the Kruskal-Wallis method

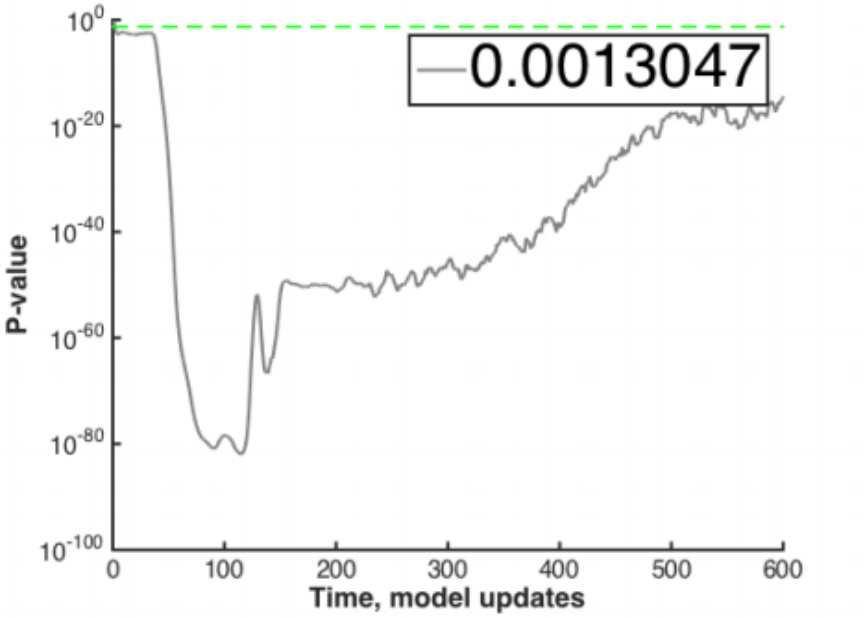


Fig. S12. Kruskal-Wallis analysis of the differences between the processes shown Fig. 3A. At each simulation time point, a p-value (Y-axis) was calculated by comparing the modeled processes each represented by a sample of 100 repeated runs using the Kruskal Wallis method

Aneuploidy Counts

Sample	# Aneuploid	# Euploid	Total Cells	Percent Aneuploid	Number of Animals	Days Post Recons-titution	Published in Pfau *et al.* 2016?
3 weeks FL	1	11	12	8.3	2	21	yes
3 weeks BM	1	17	18	5.6	2	21	yes
6 weeks FL	6	19	25	24.0	3	42	yes
6 weeks BM	6	19	25	24.0	3	42	yes
13 weeks FL	4	12	16	25.0	2	91	yes
13 weeks BM	3	13	16	18.8	2	91	yes
36 weeks FL	1	11	12	8.3	2	252	yes
34 weeks BM	2	10	12	16.7	2	238	yes
52 weeks FL	0	12	12	0.0	1	364	no
50 weeks BM	0	11	11	0.0	2	350	no

Chapter Four: Aneuploidy and a deregulated DNA damage response suggest haploinsufficiency in breast tissues of BRCA2 mutation carriers

Karaayvaz-Yildirim M, Silberman RE, Langenbucher A, Saladi SV, Ross KN, Zarcaro E, Desmond A, Yildirim M, Vivekanandan V, Ravichandran H, et al. 2020. Aneuploidy and a deregulated DNA damage response suggest haploinsufficiency in breast tissues of BRCA2 mutation carriers. *Sci Adv* **6**: 1–11.

Abstract

Women harboring heterozygous germline mutations of *BRCA2* have a 50-80% risk of developing breast cancer, yet the early pathogenesis of these cancers is poorly understood. To reveal early steps in *BRCA2*-associated carcinogenesis we analyzed sorted cell populations from freshly-isolated, non-cancerous breast tissues of *BRCA2* mutation carriers and matched controls. Single-cell whole-genome sequencing demonstrates that >25% of *BRCA2* carrier ($BRCA2^{mut/+}$) luminal progenitor (LP) cells exhibit sub-chromosomal copy number variations (CNVs), which are rarely observed in non-carriers. Correspondingly, primary $BRCA2^{mut/+}$ breast epithelia exhibit DNA damage together with attenuated replication checkpoint and apoptotic responses, associated with an age-associated expansion of the LP compartment. We provide evidence that these phenotypes do not require loss of the wild-type *BRCA2* allele. Collectively, our findings suggest that *BRCA2* haploinsufficiency and associated DNA damage precede histologic abnormalities in vivo. Employing such hallmarks of cancer predisposition will yield unanticipated opportunities for new risk assessment and prevention strategies in high-risk patients.

Introduction

Breast cancers arising in women who inherit heterozygous mutations in *BRCA2* are associated with a high prevalence of genomic alterations and aggressive clinical behavior (*1, 2*). Due to the high risk of these cancers in *BRCA2* mutation carriers, many such women elect to undergo bilateral mastectomy for breast cancer prevention. Yet despite the unmet need for more effective breast cancer prevention approaches in this setting, the stepwise evolution from an otherwise normal *BRCA2* heterozygous mutant ($BRCA2^{mut/+}$) cell to an invasive malignancy has not been defined. Homozygous loss of *BRCA2* is embryonic lethal (*3-5*), and acute loss in cultured cells rapidly leads to DNA damage and growth arrest or cell death (*6-8*). These observations suggest a multi-step pathogenesis in which homozygous *BRCA2* loss is not the earliest genetic event, but rather that the wild-type *BRCA2* allele may remain intact as early genetic changes accumulate. Critically however, this scenario leaves unresolved the nature and enabling mechanism for early cancer evolution. Haploinsufficiency for *BRCA2* has been proposed as a possible driver of early pathogenesis, but direct evidence for such an effect in the normal human mammary gland is inconsistent. Furthermore, heterozygous genetically engineered mouse models (GEMMs) of *BRCA2* are not tumor prone and therefore represent a poor model of precancerous evolution in this setting (*3-5, 8, 9*). While the *BRCA1* tumor suppressor shares many of these features (*9, 10*), the pathogenesis of *BRCA1*- versus *BRCA2*-associated breast cancers may differ in important ways, as the former are primarily hormone receptor (HR) and HER2-negative tumors, while the latter are primarily HR-positive (*11*).

We sought to unveil the earliest steps in the pathogenesis of *BRCA2*-associated breast tumors through detailed analysis of histologically normal glands from women harboring germline deleterious mutations who elected to undergo bilateral prophylactic mastectomy. Genomic analysis of individual cells revealed frequent polyclonal chromosomal damage, which

was most prevalent among the subset of epithelial cells that are the suspected cell-of-origin of these cancers. Corresponding defects in replication stress and DNA damage checkpoint responses in these same cells collectively define a previously unappreciated phenotype for *BRCA2* that precedes histologic abnormalities in the human breast. The discovery of these precancerous hallmarks paves the way for improving clinical risk prediction and cancer prevention in this population.

Results

Single-cell whole genome analysis reveals sub-chromosomal aneuploidy in BRCA2mut/+ human primary breast epithelial cells

We carried out detailed analysis of non-cancerous glands from *BRCA2* carriers who elected to undergo bilateral prophylactic mastectomy, using as controls tissues from women matched for age, menopausal status, and hormonal exposure electing cosmetic breast surgery (Fig. 1A and table S1). None of these women had a prior breast cancer diagnosis or chemotherapy exposure, and no occult cancers were detected upon histologic analysis of the tissues we analyzed (table S1). We employed established markers to carry out flow cytometry-based isolation and sorting of the three major epithelial cell sub-populations: mature luminal (ML), luminal progenitor (LP) and basal epithelial cells (Fig. 1A). Notably, data from GEMMs and gene expression analyses of human tumors have suggested that the cell of origin of *BRCA1*-associated breast cancer is the LP cell (*12, 13*), while *BRCA2*-associated tumors may arise from an LP-related cell or a more mature luminal cell (*14*).

Among the earliest events in cancer evolution are thought to be polyclonal somatic genomic alterations. Accordingly, we looked for the presence of somatic CNVs at high-resolution through single-cell whole-genome sequencing (WGS) of uncultured, flow-sorted primary LP and basal epithelial cells from *BRCA2* carriers and controls. Low-coverage WGS

provides sufficiently high resolution to identify sub-chromosomal CNVs as small as 10MB, and our methodology for single-cell whole-genome amplification and analysis has been previously validated (*15, 16*). We carried out WGS to an average depth between 0.1-0.05X, then used two independent algorithms (HMMcopy and DNAcopy) to assign and confirm copy number changes across the genome (*15, 16*). Previous studies employing this methodology have demonstrated that in unselected individuals the proportion of cells with any such CNVs is very low (<5% of cells) in normal epithelial and brain tissues (*16*). In contrast, among nearly 100 individual LP cells from a cohort of *BRCA2* carriers analyzed by WGS we observed that 27% demonstrated one or more CNVs of >10MB (Fig. 1, B to D). Applying this methodology to an equal number of basal breast epithelial cells from the same individuals also revealed a substantial excess of cells harboring CNVs (13%), although significantly less than the proportion of CNV-positive LP cells ($p=0.04$) (Fig. 1, B to D). By comparison, a parallel WGS analysis of sorted LP and basal cells from non-carriers revealed a single CNV in 90 cells (Fig. 1, B and C). As further validation of our sequencing and analysis pipelines we re-analyzed existing data from normal skin and brain cells sequenced on the same platform. The overall sequence quality was comparable between these cells and the breast epithelial cells, and we confirmed the low prevalence of CNV-positive cells in 142 skin and brain cells sequenced (fig. S1). Thus, breast epithelia and particularly LP cells from non-cancerous breast tissue of *BRCA2* carriers harbor frequent sub-chromosomal aneuploid events (Fig. 1D, fig. S2A).

One notable CNV we observed was duplication of the entire chromosome 1q arm, which is a common genomic abnormality in breast cancer (Fig. 2A) (*17*). The majority of the identified CNVs were sub-chromosomal haploid losses, consistent with the widespread pattern of losses observed in *BRCA2*-associated breast cancer (Fig. 2A and fig. S2B) (*1*). In some cases, identical

losses were shared between multiple cells of the same patient, a finding which could conceivably correspond to early clonal evolution (Fig. 2B). Importantly, none of the losses in any cell involved the *BRCA2* locus on chromosome 13 (Fig. 1B and fig. S2, B to E). Prior analyses of germline *BRCA2*-associated breast cancers have demonstrated that most genetic loss-of-heterozygosity (LOH) events for *BRCA2* itself are >10MB and therefore would have been detected by our analysis (*18*). This observation suggests that the wild type *BRCA2* allele is intact in our cases, implying that accumulation of sub-chromosomal aneuploidy may be a haploinsufficient phenotype. To confirm the integrity of the wild-type allele we performed targeted PCR amplification of the locus surrounding the patient-specific *BRCA2* mutation from individual cells. Although efficiency for detection of either allele was low, we did not observe a bias toward detection of the mutant allele alone in the cells analyzed (Fig. 2C and table S2). Taken together, these findings imply that sub-chromosomal aneuploidy is an early and potentially haploinsufficient phenotype in $BRCA2^{mut/+}$ breast epithelia.

$BRCA2^{mut/+}$ primary cells exhibit DNA damage and a deregulated replication stress response

The presence of viable aneuploid cells in $BRCA2^{mut/+}$ tissues suggested ongoing DNA damage and/or a deregulated stress/damage response. Thus, we next used an independent method to directly assess DNA damage in single cells, the comet assay. This assay employs cells embedded in agarose that are lysed then subjected to electrophoresis, causing broken DNA structures to migrate toward the anode, thus forming a comet tail (*19*). We briefly cultured freshly-collected cells from *BRCA2* carriers or controls under ultra-low attachment conditions (48-72h) to select for epithelial progenitor cells prior to plating (*20*). Consistently, cells from *BRCA2* carriers demonstrated increased DNA breaks at baseline compared to controls (Fig. 3A). Additionally, inducing replication stress by treatment with hydroxyurea (HU) led to further

increases in DNA damage in *BRCA2*$^{mut/+}$ cells, potentially reflecting the established role of *BRCA2* in protection of stalled replication forks (Fig. 3A) (*21*).

We then examined the response to this genomic stress by analyzing phosphorylation of CHK1, a central coordinator of the response to replication stress and DNA damage (*22, 23*). Cytospins of primary epithelial progenitor cultures prepared as above were stained for phosphorylated CHK1 at baseline or following 4 hours of exposure to HU. As anticipated, control primary epithelia exhibited increased CHK1 phosphorylation within 4 hours of HU treatment (Fig. 3B). In contrast, however, cells from *BRCA2* carriers exhibited a failure to activate CHK1 in response to HU, despite normal levels of total CHK1 protein (Fig. 3B and fig. S3A). DNA sequencing of these cells revealed the presence of both wild-type and mutant *BRCA2* alleles (Fig. 3C). These findings provide further support for a haploinsufficient phenotype of *BRCA2* in the response to genomic stress.

Because we observed a deregulated genomic stress response in vitro, we wanted to know whether this also occurs in vivo. Thus, we carried out RNA-seq analysis of freshly-sorted LP and basal epithelial cell populations from *BRCA2* carrier tissues or controls (Fig. 1A). Analysis of these data revealed enrichment in *BRCA2*$^{mut/+}$ LP cells of an established signature reflecting a failure of the ATR/CHK1-mediated replication stress checkpoint in non-transformed mammary epithelial cells (Fig. 3D and fig. S3B) (*24*). This replication stress response deficiency (RSRD) signature is known to predict future cancer risk (*24*), and it contains some of the top most differentially expressed genes between *BRCA2*$^{mut/+}$ and control LP cells (Fig. 3D). Among these are genes of potential relevance to HR-positive breast cancer (which comprise 80% of *BRCA2*-associated breast cancers), including the estrogen receptor target gene *HOXC4* and the GATA transcription factor binding partner gene *ZFPM1* (Fig. 3D) (*25, 26*). Furthermore, evaluation of

differentially expressed programs through Gene Set Enrichment Analysis revealed the highly significant deregulation of a radiation response signature in *BRCA2*$^{mut/+}$ LP cells (fig. S3C) (*27*). Notably, the differential expression of this signature between *BRCA2*$^{mut/+}$ and control cells was far more significant within the LP compared to the basal population, in keeping with the more frequent occurrence of CNVs among LP cells (fig. S3C). Again consistent with haploinsufficiency for *BRCA2*, the RNA-seq data showed no evidence for exclusive expression of the mutant allele in *BRCA2*$^{mut/+}$ LP cells (fig. S3D). Thus, *BRCA2*$^{mut/+}$ LP cells exhibit evidence of aberrant replication stress and DNA damage responses in vivo.

BRCA2$^{mut/+}$ *LP cells show increased TP53 activity and decreased NF-kB/SASP pathway expression*

We then turned to examine the downstream consequences of the DNA damage detected in LP cells of *BRCA2* carriers. A hallmark genetic event that cooperates with *BRCA2* deficiency in cancer pathogenesis is loss of *TP53*, suggesting that activation of TP53 may be an early barrier to malignant progression in this setting (*28*). We therefore hypothesized that the failed CHK1-dependent replication stress response we observed might ultimately lead to DNA double-strand breaks and thereby trigger TP53 activation through a CHK1-independent pathway (*29*). Indeed, recent studies suggest CHK1 is not required for TP53 activation in primary breast epithelial cells following DNA damage (*30*). RNA-seq analysis did suggest activation of TP53 in *BRCA2*$^{mut/+}$ LP cells, evidenced by the increased expression of multiple direct TP53 target genes (Fig. 4A) (*31, 32*). This in vivo effect was associated with a strong transcriptional profile indicating suppression of NF-kB signaling, including numerous cytokine and inflammatory factors associated with the senescence-associated secretory phenotype (SASP) (Fig. 4, B to D). TP53 is known to suppress the NF-kB/SASP response (*33, 34*), and this effect is emerging as a

relevant component of TP53-dependent tumor suppression given that accumulation of SASP-expressing cells is an established driver of tumorigenesis (*35*). We independently validated the corresponding alterations in NF-kB protein expression, demonstrating that the NFKB1 (p50) and NFKB2 (p52) subunits were expressed at lower levels in *BRCA2* carrier tissues compared to controls (Fig. 4C). Furthermore, knockdown of *BRCA2* in non-transformed mammary epithelial cells via lentiviral shRNA attenuated expression of the same cytokine and NF-kB targets genes that were downregulated in $BRCA2^{mut/+}$ progenitor cells in vivo (fig. S4A). Like the damage response signature (fig. S3C), deregulation of the SASP program was selective for LP cells in *BRCA2* carriers, as no significant suppression of SASP was observed in the corresponding basal epithelial cells of these same patients (fig. S4B). These results suggest that DNA damage and TP53 activation in $BRCA2^{mut/+}$ LP cells are associated with suppression of the NF-kB/SASP response.

Age-associated deregulation of breast epithelial cell proportions in BRCA2 carriers suggests expansion of a damaged LP cell population over time

Deregulated DNA damage and senescence/SASP responses in *BRCA2* LP cells might be expected to alter the proportion of these cells over time (*36*). We thus sought to address whether there were differences in the proportions of progenitor or other epithelial sub-populations in $BRCA2^{mut/+}$ tissues compared to controls. We collected a larger cohort of tissues from *BRCA2* carriers (N=26) and controls (N=28), then performed flow cytometry analysis on these specimens and plotted the proportions of each epithelial sub-population as a function of age for each cohort (Fig. 5A). In non-carrier controls no significant age-associated changes in the prevalence of these sub-populations were noted. In contrast, *BRCA2* carriers showed an age-associated expansion in the proportion of LP cells and a decline in the basal cell fraction (Fig. 5B

and fig. S5, A and B). These differences were not accounted for by demographic factors such as parity or menopausal status, as such factors were not associated with significant differences in epithelial cell proportions (fig. S5, C and D). Thus, DNA-damage and suppression of a senescence-associated program in *BRCA2*$^{mut/+}$ LP cells is accompanied by an age-associated expansion of this progenitor cell compartment (*36*).

Finally, we hypothesized that altered epithelial cell proportions and a deregulated NF-kB/SASP program in *BRCA2* carrier tissues may be associated with differences in cell proliferation and/or survival in vivo (*36*). We did not observe strong differences in proliferation assessed by Ki67 staining between these *BRCA2* carrier breast tissues and controls (not shown), prompting us to ask whether differences in cell survival might contribute to the age-associated expansion of the LP population in this context. We therefore carried out TUNEL staining, an established marker of apoptosis, in *BRCA2* carrier tissues and controls. The proportion of TUNEL-positive cells is well-documented in normal human breast epithelial tissues, and we observed a similar prevalence of these cells in the control tissues we tested (Fig. 5C) (*37*). In contrast, however, *BRCA2* mutation carrier tissues consistently showed a paucity of TUNEL-positive luminal epithelial cells across all patients tested, in keeping with established links between checkpoint and NF-kB suppression and a defective apoptotic response (Fig. 5C) (*38, 39*). Collectively, our findings suggest that non-cancerous *BRCA2*$^{mut/+}$ breast tissues exhibit *BRCA2* haploinsufficiency and an age-associated accumulation of DNA-damaged luminal epithelial progenitor cells bearing altered checkpoint and survival responses (Fig. 6).

Discussion

This study advances our understanding of early changes in *BRCA2*$^{mut/+}$ breast tissues, defining unanticipated phenotypes in this setting with implications for both cancer risk

assessment and prevention. The majority of the tissues we studied were deemed to be histologically normal by highly experienced breast pathologists, suggesting that the alterations we report precede clinically-defined cellular abnormalities (tables S1, S3). We present evidence that a failed replication stress response and DNA damage in *BRCA2*$^{mut/+}$ tissues result from haploinsufficiency for *BRCA2* rather than homozygous loss of function. While the presence of haploinsufficiency for either *BRCA1* or *BRCA2* in vivo has been controversial, our findings are in accord with data suggesting that LOH for the wild-type *BRCA2* is not universal in *BRCA2*-associated cancers (*40*). Our observations are also in keeping with a recent report that the BRCA2 protein is selectively susceptible to degradation by environmental aldehydes (*41*), an effect which could contribute to a haploinsufficient phenotype in cells with only one functional *BRCA2* allele. Nonetheless, we analyzed a relatively small number of cells and tissues, and it is difficult to definitively rule out LOH in a subset of cells. Thus, our study suggests rather than confirms haploinsufficiency for *BRCA2* as a potential initiating event for these cancers.

A prominent feature of the phenotype we have uncovered is frequent sub-chromosomal aneuploidy, most prevalent within the LP cell population. LP cells are a potential target cell for *BRCA2*-associated carcinogenesis in the breast, and indeed we observe instances of apparently clonally related genomic alterations among these cells. While our study does not prove they are direct cancer precursors, such alterations could conceivably represent the earliest somatic genetic abnormalities that underlie these malignancies. Notably, all the CNVs we identified were sub-chromosomal and therefore are to be distinguished from whole-chromosome gains and losses that are typically later events and are associated with TP53 inactivation (*42*).

Although the early genomic changes we observed are likely to include many passenger events, they nevertheless may provide a quantifiable hallmark of the pre-neoplastic *BRCA2*

carrier state. Tracking the prevalence of DNA-damaged cells in the clinical setting could possibly improve risk prediction for such women, who are faced with the difficult choice of whether to undergo mastectomy long before cancer develops. Finally, the *BRCA2* haploinsufficient phenotype we report may portend particular vulnerabilities of certain *BRCA2*$^{mut/+}$ cancer precursor cells. Accordingly, this work provides a foundation for future studies seeking to identify novel pharmacologic approaches to cancer prevention in this setting.

Materials and Methods

Human breast tissues

Fresh human breast tissues were obtained from Massachusetts General Hospital with approval by the local Institution Review Board and signed informed patient consent (Protocols 93-085 and 2008-P-1789). Samples were either normal breast tissues from reduction mammoplasties (confirmed by pathology) or non-cancerous breast tissues from prophylactic mastectomies of known BRCA1 or BRCA2 mutation carriers. All BRCA1/2 carrier status was determined through clinical germline genetic testing performed by commercial providers prior to tissue collection.

Mammary cell preparations

Tissue samples were minced and digested with collagenase/hyaluronidase (Stemcell technologies) in complete Epicult-B Medium supplemented with 0.48 μg/mL hydrocortisone (Stemcell technologies) overnight at 37 °C. The resulting suspensions were either cryopreserved or further sequentially digested with 0.25% trypsin, 5 mg/mL dispase and 1 mg/mL DNase I. Single cell suspensions were collected by filtration through a 40 μm cell strainer.

Cell staining and sorting

Cells were blocked with rat immunoglobulin (Jackson Immunolabs) and antibody to F_c receptor binding inhibitor (eBioscience) before incubation with the following primary antibodies: PE-conjugated anti-human CD31 (BD Pharmingen), PE-conjugated anti-human CD45 (BD Pharmingen), PE-conjugated anti-human CD235a (BD Pharmingen), BV650-conjugated anti-human EPCAM CD326 (Biolegend), biotin-conjugated anti-human ITGA6 (eBioscience). Where required, cells were incubated with APC-Cy7-conjugated streptavidin (BD Pharmingen). Cells were either stained with DAPI for viability or fixed with 1% PFA and stained with zombie aqua fixable viability kit (Biolegend). Viable cells were sorted on a FACSAria flow cytometer (Becton Dickinson). Data was analyzed using FlowJo software (Tree Star).

Single-Cell PCR for allele-specific LOH analysis

Microaspirated single cells were transferred into PCR tubes containing lysis buffer (water + 400 ng/ul Proteinase K + 17 μM SDS) and DNA was amplified by nested PCR using primers flanking BRCA2 mutations. Sanger sequencing was performed by the CCIB DNA Core Facility at Massachusetts General Hospital.

Primer sequences for patient 128 (Val3079PhefsX4) are as in the following:

1st PCR-Forward: TGGCGTCCATCATCAGATTT

1st PCR-Reverse: TCAGAGGTTCAAAGAGGCTTAC

2nd PCR-Forward: CAGATTTACCAGCCACGGGA

2nd PCR-Reverse: GCCAACTGGTAGCTCCAACTAA

Primer sequences for patient 140 (6027del4) are as in the following:

1st PCR-Forward: GGGCCACCTGCATTTAGGAT

1st PCR-Reverse: TGAGCTGGTCTGAATGTTCGT

2nd PCR-Forward: GCAGGTTGTTACGAGGCATT

2nd PCR-Reverse: CCTGGACAGATTTTCCACTTGC

Comet Assays

Single cell suspensions from patient samples were plated in ultralow-adherence plates in DMEM/F12 medium containing 5 μg/ml insulin, 10 ng/mL EGF, 5 ng/ml bFGF, 4 μg/ml heparin, 500 ng/ml hydrocortisone, B27, Glutamax and penicillin-streptomycin. Cells were either treated with hydroxyurea (10 mM, Sigma) for 4 h or left untreated, washed with PBS and alkaline comet assays were performed using TrevigenCometAssay kit, according to the manufacturer's instructions. Olive tail movement was quantified with ImageJ, 50 individual cells were quantified per condition.

Immunostaining

Immunofluorescence for paraffin sections and TUNEL staining were performed by Dana-Farber/Harvard Cancer Center Specialized Histopathology Core. For Immunofluorescence in cells, fixation was performed with methanol for 10 min followed by permeabilization in 0.1% TritonX100 for 2 min. Blocking was performed with 10% horse serum for 30 min and cells were further incubated with primary p-CHK1 antibody (Novus Biologicals) for 2 hr, washed with wash buffer (PBS +10% horse serum + 0.1% TritonX100), incubated with appropriate secondary antibody for 1hr and stained with DAPI. All immunofluorescence images were captured by a confocal microscope (Leica TCS SP8) and were analyzed by ImageJ.

Western Blotting

Snap-frozen tissues were homogenized using Precellys 24 homogenizer (Bertin Technologies). For total protein extraction, cells were lysed in RIPA buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% sodium deoxycholate, 0.1% (w/v) SDS, 1% (v/v) NP40, proteinase inhibitor cocktail, phosphatase inhibitor cocktail) for 30 min on ice. Western blotting was performed using NFKB p50 (Santa Cruz Biotechnologies) and NFKB p52 (Millipore) antibodies by standard protocol.

Single-Cell Copy Number Analysis

Fresh tissues were dissociated as described above and single cells were isolated by microaspiration. Genomic DNA was amplified and sequenced as described in Knouse et al. 2014. Fastqs were aligned using bwa-mem, with resulting bams sorted and duplicates marked using Picard. Coverage was then computed over 500kb bins across the entire genome. The count for each bin was then divided by the sum across all bins for the relative sample (to correct for library size), and then by the median for that genomic bin across all samples from the same batch. The

coverage profiles were then transformed into .wig files and fed into the R package HMMCopy for segmentation and CNV-calling. HMMCopy was run with e = 0.9999999 and nu = 5, with all other parameters set to default. A noise statistic termed "VS" was computed in the same manner as Knouse etal. 2014, with cells with values greater than or equal to 0.5 being excluded from the analysis. CNVs that mapped to the Y chromosome, were less than 10Mb in size, or had an absolute log2 ratio less than 0.4 were excluded from the analysis.

RNA-Seq Analysis

Total RNA from sorted cell populations was extracted using RNeasy FFPE kit, according to manufacturer's instructions. Libraries for ribosomally reduced RNA was prepared by Harvard Biopolymers Facility using directional RNA-seq Wafergen protocol. Libraries were sequenced on Illumina HiSeq 2000 at Next Generation Sequencing Core at Massachusetts General Hospital. TPM values were computed using Salmon and batch-corrected using ComBat. The two samples with the lowest total counts were excluded from the analysis. GSEA was run on the ComBat-corrected TPM values using phenotype permutation and default parameters. The heatmaps in figures 4B and 4D were made using the ComBat-corrected TPM values, subset to the comparison of interest, and transformed into z-scores by gene.

Other Statistical Methods

p Values were determined using the Student's unpaired t test unless indicated otherwise.

Data Availability

All data supporting the findings of this study are available within the article and its supplementary information files or upon request. The RNAseq and SC-WES data have been deposited in the dpGaP database under accession code XXXXX.

Acknowledgements

We thank the HSCI-CRM Flow Cytometry Core Facility for assistance with cell sorting and MGH DF/HCC Specialized Histopathology Service Core for immunostaining experiments. We thank Biopolymers Facility at Harvard Medical School for library processing of RNA samples and MGH Next Generation Sequencing Core for performing RNA-sequencing. We are grateful to MIT BioMicroCenter for performing genome sequencing reactions.

Funding
This work was supported by DOD/CDMRP grant BC140903 (L.W.E.), by the Tracey Davis Breast Cancer Research Fund (L.W.E.), by the Weissman Family MGH Research Scholar grant (L.W.E.), by the Susan G. Komen Foundation grant PDF16380794 (M.K.), by a Terri Brodeur Breast Cancer Foundation grant 2016D001483 (M.K.), and by the Howard Hughes Medical Institute (A.A.).

Author Contributions
M. K., A.A. and L.W.E. conceived and designed the study. M.C.S. contributed patient samples. M.K., R.E.S. and S.V.S. designed and performed experiments and interpreted data. H.R. and R.M. and V.V. performed experiments. M.K., S.V.S, E.Z., A.D., M.Y. performed data analysis. M.K, A.L., K.R., S.R. and M.L. performed bioinformatic analysis and interpreted data. L.W.E., A.A. and M.L. conceived experiments, interpreted data and provided funding. L.W.E. and M.K. wrote the manuscript. All authors approved the final submitted manuscript.

The authors declare no competing financial interests.

References

1. H. Davies *et al.*, HRDetect is a predictor of BRCA1 and BRCA2 deficiency based on mutational signatures. *Nat Med***23**, 517-525 (2017).
2. F. J. Couch, K. L. Nathanson, K. Offit, Two decades after BRCA: setting paradigms in personalized cancer care and prevention. *Science***343**, 1466-1470 (2014).
3. T. Ludwig, D. L. Chapman, V. E. Papaioannou, A. Efstratiadis, Targeted mutations of breast cancer susceptibility gene homologs in mice: lethal phenotypes of Brca1, Brca2, Brca1/Brca2, Brca1/p53, and Brca2/p53 nullizygous embryos. *Genes Dev***11**, 1226-1241 (1997).
4. S. K. Sharan *et al.*, Embryonic lethality and radiation hypersensitivity mediated by Rad51 in mice lacking Brca2. *Nature***386**, 804-810 (1997).
5. A. Suzuki *et al.*, Brca2 is required for embryonic cellular proliferation in the mouse. *Genes Dev***11**, 1242-1252 (1997).
6. A. Tutt *et al.*, Cell cycle and genetic background dependence of the effect of loss of BRCA2 on ionizing radiation sensitivity. *Oncogene***22**, 2926-2931 (2003).
7. W. Feng, M. Jasin, BRCA2 suppresses replication stress-induced mitotic and G1 abnormalities through homologous recombination. *Nat Commun***8**, 525 (2017).
8. F. Connor *et al.*, Tumorigenesis and a DNA repair defect in mice with a truncating Brca2 mutation. *Nat Genet***17**, 423-430 (1997).
9. L. C. Gowen, B. L. Johnson, A. M. Latour, K. K. Sulik, B. H. Koller, Brca1 deficiency results in early embryonic lethality characterized by neuroepithelial abnormalities. *Nat Genet***12**, 191-194 (1996).
10. R. Hakem *et al.*, The tumor suppressor gene Brca1 is required for embryonic cellular proliferation in the mouse. *Cell***85**, 1009-1023 (1996).
11. Pathology of familial breast cancer: differences between breast cancers in carriers of BRCA1 or BRCA2 mutations and sporadic cases. Breast Cancer Linkage Consortium. *Lancet***349**, 1505-1510 (1997).
12. G. Molyneux *et al.*, BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells. *Cell Stem Cell***7**, 403-417 (2010).
13. E. Lim *et al.*, Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat Med***15**, 907-913 (2009).
14. J. E. Visvader, J. Stingl, Mammary stem cells and the differentiation hierarchy: current status and perspectives. *Genes Dev***28**, 1143-1158 (2014).
15. K. A. Knouse, J. Wu, A. Amon, Assessment of megabase-scale somatic copy number variation using single-cell sequencing. *Genome Res***26**, 376-384 (2016).
16. K. A. Knouse, J. Wu, C. A. Whittaker, A. Amon, Single cell sequencing reveals low levels of aneuploidy across mammalian tissues. *Proc Natl Acad Sci U S A***111**, 13409-13414 (2014).
17. N. Cancer Genome Atlas, Comprehensive molecular portraits of human breast tumours. *Nature***490**, 61-70 (2012).
18. S. Nik-Zainal *et al.*, Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature***534**, 47-54 (2016).
19. X. Song *et al.*, alpha-MSH activates immediate defense responses to UV-induced oxidative stress in human melanocytes. *Pigment Cell Melanoma Res***22**, 809-818 (2009).
20. E. Nolan *et al.*, RANK ligand as a potential target for breast cancer prevention in BRCA1-mutation carriers. *Nat Med***22**, 933-939 (2016).
21. W. Feng, M. Jasin, Homologous Recombination and Replication Fork Protection: BRCA2 and More! *Cold Spring Harb Symp Quant Biol***82**, 329-338 (2017).
22. N. Walworth, S. Davey, D. Beach, Fission yeast chk1 protein kinase links the rad checkpoint pathway to cdc2. *Nature***363**, 368-371 (1993).
23. J. Bartek, J. Lukas, Chk1 and Chk2 kinases in checkpoint control and cancer. *Cancer Cell***3**, 421-429 (2003).
24. D. J. McGrail *et al.*, Defective Replication Stress Response Is Inherently Linked to the Cancer Stem Cell Phenotype. *Cell Rep***23**, 2095-2106 (2018).
25. T. Mai *et al.*, Estrogen receptors bind to and activate the HOXC4/HoxC4 promoter to potentiate HoxC4-mediated activation-induced cytosine deaminase induction, immunoglobulin class switch DNA recombination, and somatic hypermutation. *J Biol Chem***285**, 37797-37810 (2010).
26. N. M. Robert, J. J. Tremblay, R. S. Viger, Friend of GATA (FOG)-1 and FOG-2 differentially repress the GATA-dependent activity of multiple gonadal promoters. *Endocrinology***143**, 3963-3973 (2002).
27. S. A. Ghandhi, B. Yaghoubian, S. A. Amundson, Global gene expression analyses of bystander and alpha particle irradiated normal human lung fibroblasts: synchronous and differential responses. *BMC Med Genomics***1**, 63 (2008).
28. J. Jonkers *et al.*, Synergistic tumor suppressor activity of BRCA2 and p53 in a conditional mouse model for breast cancer. *Nat Genet***29**, 418-425 (2001).
29. J. Smith, L. M. Tho, N. Xu, D. A. Gillespie, The ATM-Chk2 and ATR-Chk1 pathways in DNA damage signaling and cancer. *Adv Cancer Res***108**, 73-112 (2010).
30. M. T. M. van Jaarsveld, D. Deng, E. A. C. Wiemer, Z. Zi, Tissue-Specific Chk1 Activation Determines Apoptosis by Regulating the Balance of p53 and p21. *iScience***12**, 27-40 (2019).
31. S. S. McDade *et al.*, Genome-wide characterization reveals complex interplay between TP53 and TP63 in response to genotoxic stress. *Nucleic Acids Res***42**, 6270-6285 (2014).

32. S. T. Younger, D. Kenzelmann-Broz, H. Jung, L. D. Attardi, J. L. Rinn, Integrative genomic analysis reveals widespread enhancer regulation by p53 in response to DNA damage. *Nucleic Acids Res***43**, 4447-4462 (2015).
33. J. P. Coppe *et al.*, Senescence-associated secretory phenotypes reveal cell-nonautonomous functions of oncogenic RAS and the p53 tumor suppressor. *PLoS Biol***6**, 2853-2868 (2008).
34. C. D. Wiley *et al.*, Small-molecule MDM2 antagonists attenuate the senescence-associated secretory phenotype. *Sci Rep***8**, 2410 (2018).
35. C. J. Sieben, I. Sturmlechner, B. van de Sluis, J. M. van Deursen, Two-Step Senescence-Focused Cancer Therapies. *Trends Cell Biol***28**, 723-737 (2018).
36. B. G. Childs, D. J. Baker, J. L. Kirkland, J. Campisi, J. M. van Deursen, Senescence and apoptosis: dueling or complementary cell fates? *EMBO Rep***15**, 1139-1153 (2014).
37. F. Feuerhake, W. Sigg, E. A. Hofter, P. Unterberger, U. Welsch, Cell proliferation, apoptosis, and expression of Bcl-2 and Bax in non-lactating human breast epithelium in relation to the menstrual cycle and reproductive history. *Breast Cancer Res Treat***77**, 37-48 (2003).
38. M. B. Kastan, J. Bartek, Cell-cycle checkpoints and cancer. *Nature***432**, 316-323 (2004).
39. S. J. Veuger, B. W. Durkacz, Persistence of unrepaired DNA double strand breaks caused by inhibition of ATM does not lead to radio-sensitisation in the absence of NF-kappaB activation. *DNA Repair (Amst)***10**, 235-244 (2011).
40. K. N. Maxwell *et al.*, BRCA locus-specific loss of heterozygosity in germline BRCA1 and BRCA2 carriers. *Nat Commun***8**, 319 (2017).
41. S. L. W. Tan *et al.*, A Class of Environmental and Endogenous Toxins Induces BRCA2 Haploinsufficiency and Genome Instability. *Cell***169**, 1105-1118 e1115 (2017).
42. L. Sansregret, C. Swanton, The Role of Aneuploidy in Cancer Evolution. *Cold Spring Harb Perspect Med***7**, (2017).
43. M. R. Ramsey *et al.*, FGFR2 signaling underlies p63 oncogenic function in squamous cell carcinoma. *J Clin Invest***123**, 3525-3538 (2013).

Figures

Fig. 1

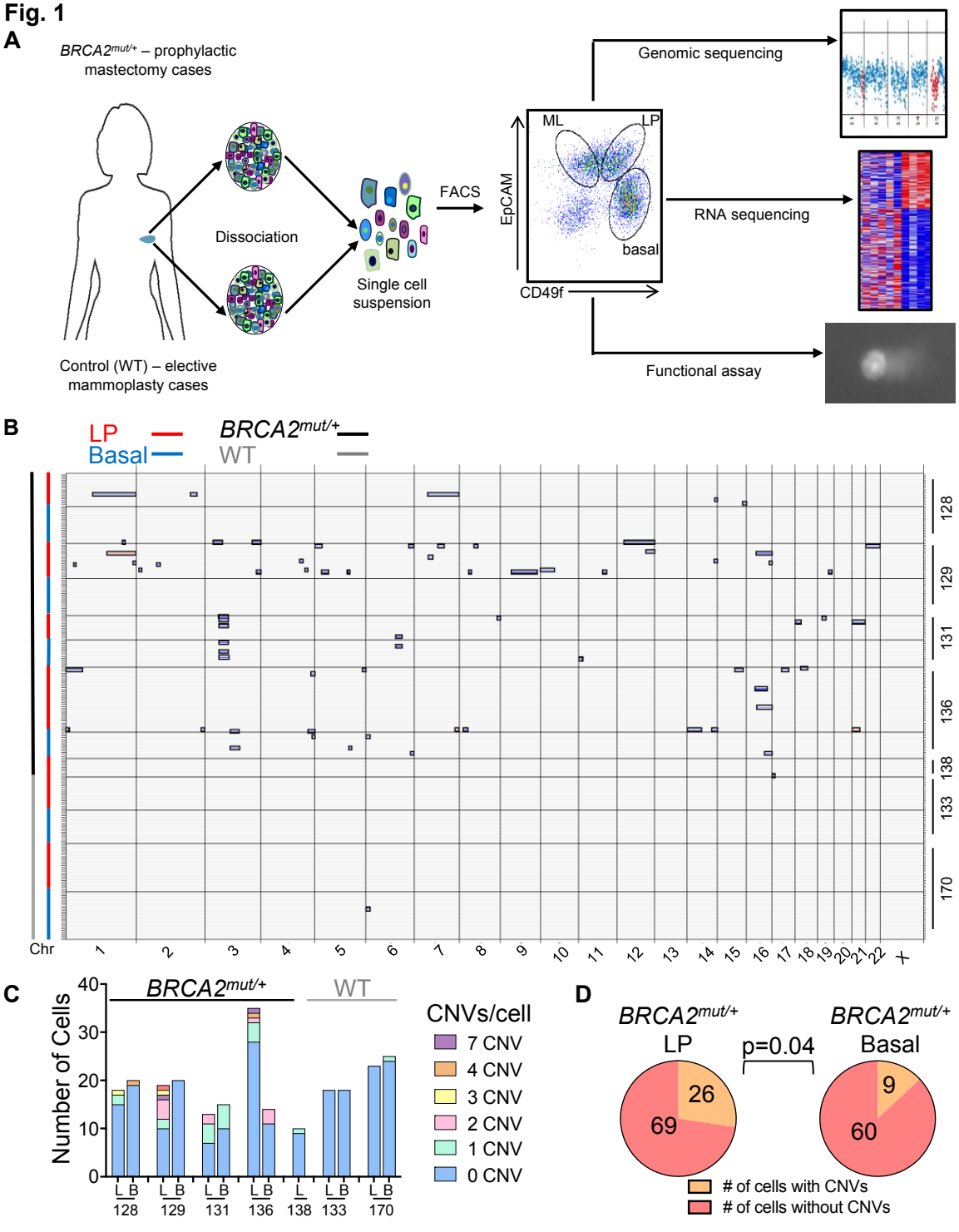


Fig. 1 Single-cell whole genome analysis of *BRCA2*$^{mut/+}$ human primary breast epithelial cells.

(A) Workflow depicts dissociation and isolation of human breast epithelial cells from *BRCA2* carrier (*BRCA2*$^{mut/+}$) prophylactic mastectomy and control (WT) elective mammoplasty cases for subsequent analyses as indicated. Dot plot at center shows representative flow cytometry sorting via CD49f and EpCAM of mature luminal (ML), luminal progenitor (LP), and basal epithelial cells.

(B) Summary of single-cell whole-genome sequencing (WGS) analysis of flow-sorted, primary uncultured breast epithelial cells. Copy Number Variation (CNV) calls for individual cells (rows) across the genome (x-axis; Chr, chromosome) are shown, with gains and losses boxed. Cell types and genotypes are indicated at top left, and individual patient ID numbers are indicated at right. In total, 252 sequenced breast epithelial cells from *BRCA2*$^{mut/+}$ (N=5) and control (N=2) tissue specimens are depicted.

(C) Bar chart depicting the prevalence of CNVs in LP (L) and basal (B) cells of *BRCA2* carrier and control (WT) patients. Color code depicts the number of CNVs identified per cell.

(D) LP cells from *BRCA2* carriers are significantly more likely than basal cells to harbor CNVs. P-value by Chi-square test.

Fig. 2

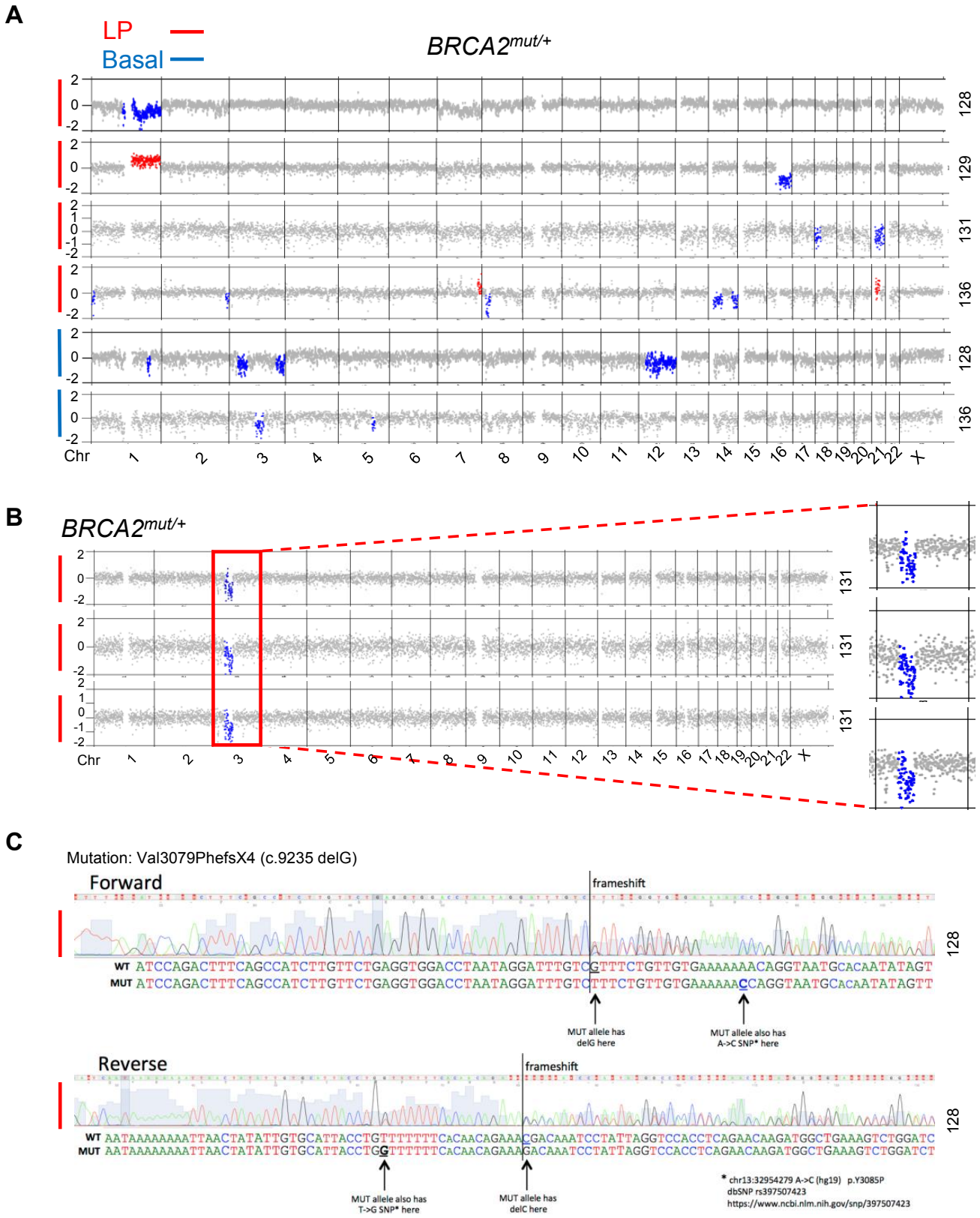

A
LP
Basal
BRCA2mut/+
Chr
B
BRCA2mut/+
Chr
C
Mutation: Val3079PhefsX4 (c.9235 delG)
Forward
frameshift
WT ATCCAGACTTTCAGCCATCTTGTTCTGAGGTGGACCTAATAGGATTTGTCGTTTCTGTTGTGAAAAAAACAGGTAATGCACAATATAGT
MUT ATCCAGACTTTCAGCCATCTTGTTCTGAGGTGGACCTAATAGGATTTGTCTTTCTGTTGTGAAAAAACCAGGTAATGCACAATATAGTT
MUT allele has delG here
MUT allele also has A->C SNP* here
Reverse
frameshift
WT AATAAAAAAAATTAACTATATTGTGCATTACCTGTTTTTTTCACAACAGAAACGACAAATCCTATTAGGTCCACCTCAGAACAAGATGGCTGAAAGTCTGGATC
MUT AATAAAAAAAATTAACTATATTGTGCATTACCTGGTTTTTTTCACAACAGAAAGACAAATCCTATTAGGTCCACCTCAGAACAAGATGGCTGAAAGTCTGGATCT
MUT allele also has T->G SNP* here
MUT allele has delC here
* chr13:32954279 A->C (hg19) p.Y3085P
dbSNP rs397507423
https://www.ncbi.nlm.nih.gov/snp/397507423

Fig. 2 Polyclonal, sub-chromosomal aneuploidy is a hallmark of *BRCA2*$^{mut/+}$ breast epithelial cells.

(A) Representative segmentation plots of individual LP (n = 4) and basal (n = 2) cells harboring CNVs from four *BRCA2* mutation carriers. Y-axis depicts normalized WGS read counts across the genome (x-axis). Red dots indicate region of gain, blue dots indicate losses. Patient ID numbers are indicated at right.

(B) Segmentation plots of 3 LP cells that share a clonal loss (red box) in a *BRCA2* carrier (Patient 131). Zoomed images of the clonal loss are shown at right.

(C) Representative chromatograms from single-cell PCR-based Sanger sequencing of genomic DNA in a *BRCA2*$^{mut/+}$ LP cell. The presence of a heterozygous SNP and the superimposition of sequences adjacent to the frameshift mutation suggest LOH has not occurred.

Fig. 3

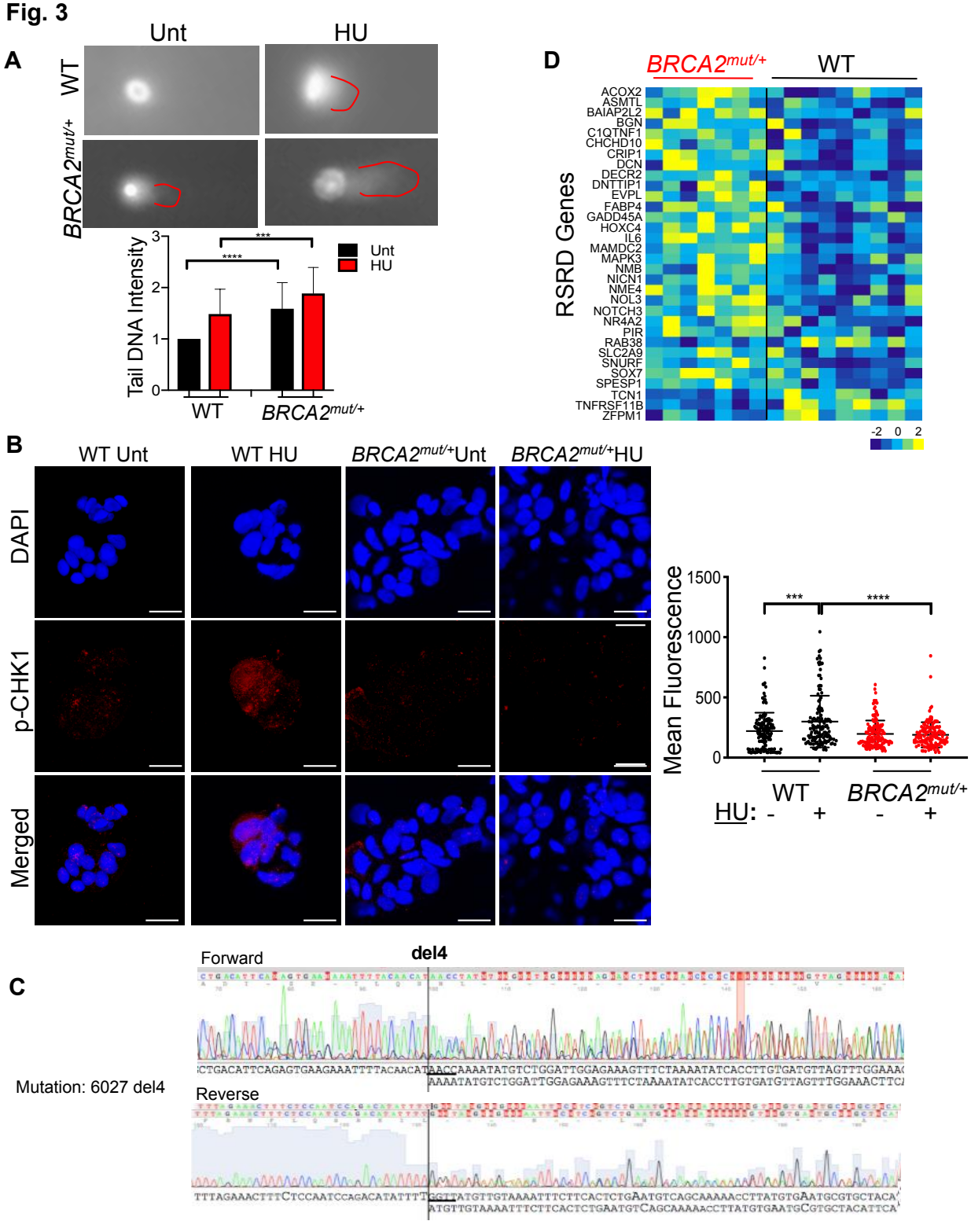

A
Unt
HU
WT
BRCA2mut/+
Tail DNA Intensity
Unt
HU
WT
BRCA2mut/+
D
BRCA2mut/+
WT
RSRD Genes
ACOX2
ASMTL
BAIAP2L2
BGN
C1QTNF1
CHCHD10
CRIP1
DCN
DECR2
DNTTIP1
EVPL
FABP4
GADD45A
HOXC4
IL6
MAMDC2
MAPK3
NMB
NICN1
NME4
NOL3
NOTCH3
NR4A2
PIR
RAB38
SLC2A9
SNURF
SOX7
SPESP1
TCN1
TNFRSF11B
ZFPM1
B
WT Unt
WT HU
BRCA2mut/+Unt
BRCA2mut/+HU
DAPI
p-CHK1
Merged
Mean Fluorescence
WT
BRCA2mut/+
HU: - + - +
C
Forward
del4
Reverse
Mutation: 6027 del4

Fig. 3 *BRCA2*$^{mut/+}$ breast epithelial cells exhibit DNA damage and an impaired replication stress checkpoint response.

(A) Representative images of comet assays performed on primary human breast epithelial cells isolated from control (WT) and *BRCA2*$^{mut/+}$ tissues. Red lines highlight "tail" of broken DNA. Graph below summarizes data from n = 3 patients per genotype, 50 cells per patient. Cells were either untreated (Unt) or treated with hydroxyurea (HU) for 4 h. Data are depicted as fold change in tail DNA intensity. p values by unpaired *t*-test. *** $p < 0.001$. Error bars indicate standard deviation.

(B) Representative confocal immunofluorescence staining of primary breast epithelial cells for p-CHK1 (Ser317) shows increased nuclear staining following HU treatment only in control (WT) but not in *BRCA2*$^{mut/+}$ cells. Graph at right summarizes nuclear fluorescence of individual cells (dots) (n =4 patients for control, n = 3 for *BRCA2*$^{mut/+}$; four fields counted per condition per patient). p values by unpaired *t*-test. *** $p < 0.001$, **** $p < 0.0001$. Horizontal lines indicate means and standard deviations. Scale bars represent 20 μm.

(C) Chromatograms depicting Sanger sequencing of a cytospin of primary breast cells assayed in (B) from a *BRCA2*$^{mut/+}$ patient. The superimposition of sequences adjacent to the frameshift mutation suggests LOH has not occurred.

(D) Heatmap of RNA-seq data from freshly-sorted cells shows differential expression of RSRD (replication stress response defect) genes (*24*) in *BRCA2*$^{mut/+}$ LP cells (n = 7 patients) compared to control (WT) LP cells (n = 9 patients). Columns correspond to individual patients.

Fig. 4

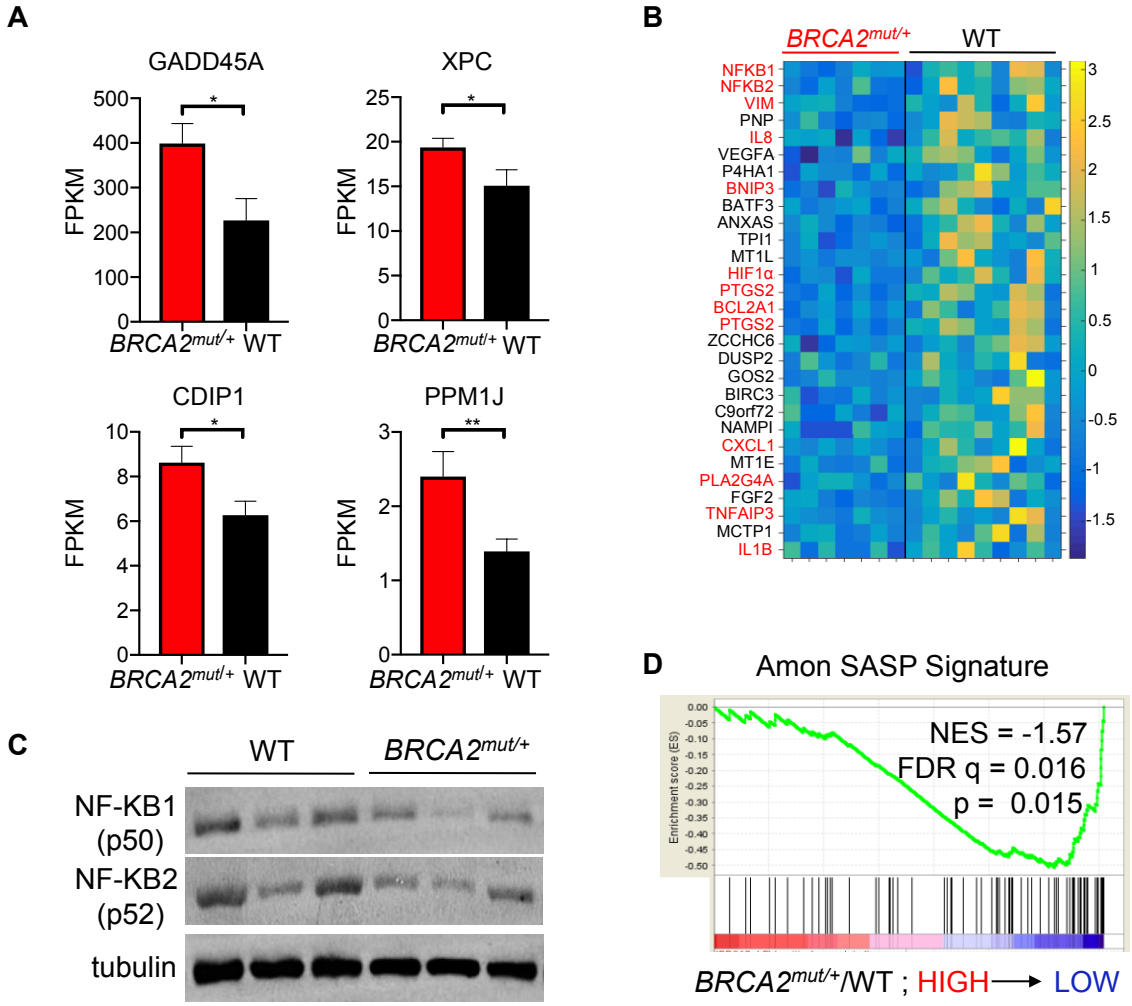


Fig. 4 *BRCA2*$^{mut/+}$ LP cells display increased TP53 activity and suppressed NF-kB/SASP pathway expression.

(A) Bar charts show the mean expression levels of canonical TP53 target genes in freshly-sorted *BRCA2* carrier LP cells (n = 7 patients) compared to controls (WT, n = 9 patients), assessed by RNA-seq. Error bars denote standard error of the mean. p values by Mann-Whitney test. * $p < 0.05$, ** $p < 0.01$.

(B) Heatmap depicts down-regulation of NF-kB/SASP pathway genes in *BRCA2* carrier LP cells compared to controls (WT), assessed by RNA-seq as in (A). Columns correspond to individual patients. Direct NF-kB target genes are highlighted in red.

(C) Western blot analysis shows that NFKB1 (p50) and NFKB2 (p52) subunits are expressed at lower levels in $BRCA2^{mut/+}$ breast tissues compared to control (WT) tissues (n =3 patients per genotype). β-tubulin serve as loading control.

(D) Negative enrichment of a SASP signature in GSEA analysis of RNA-seq data from freshly-sorted LP cells of *BRCA2* carriers (n = 7 patients) and controls (WT, n = 9 patients). NES, normalized enrichment score; FDR, false discovery rate.

Fig. 5

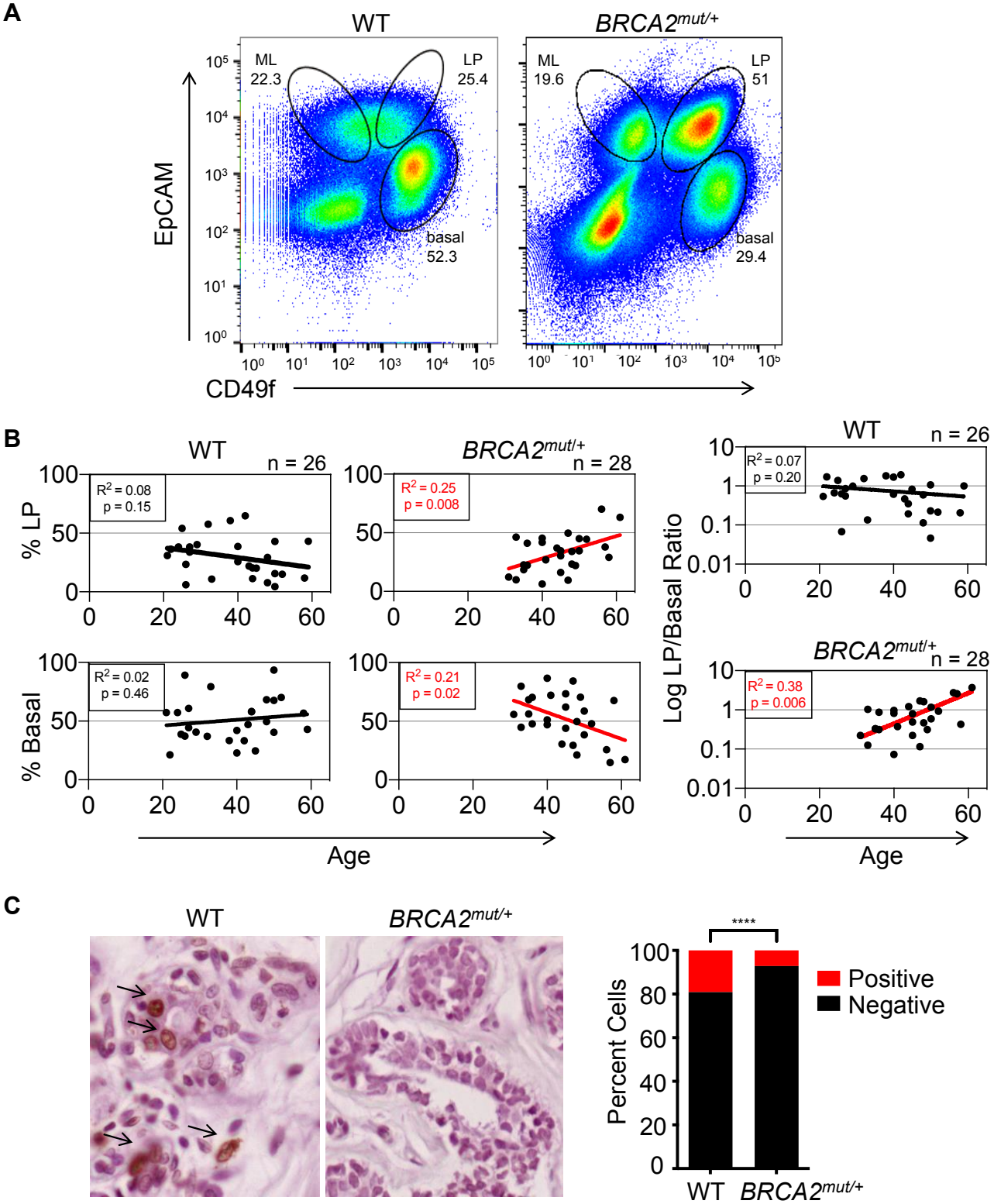

A
WT
BRCA2mut/+
ML
22.3
LP
25.4
basal
52.3
ML
19.6
LP
51
basal
29.4
EpCAM
CD49f
B
WT
n = 26
BRCA2mut/+
n = 28
% LP
R2 = 0.08
p = 0.15
R2 = 0.25
p = 0.008
% Basal
R2 = 0.02
p = 0.46
R2 = 0.21
p = 0.02
Age
Log LP/Basal Ratio
R2 = 0.07
p = 0.20
R2 = 0.38
p = 0.006
C
WT
BRCA2mut/+
****
Percent Cells
Positive
Negative

Fig. 5 Non-cancerous breast tissues of *BRCA2* mutation carriers demonstrate age-associated deregulation of epithelial cell proportions compared to controls.

(A) Representative flow cytometry analysis showing distinct epithelial subpopulations (basal, LP, mature luminal (ML)) isolated from breast tissues of control (WT) and *BRCA2* mutation carriers following sorting via CD49f and EpCAM staining. Numbers indicate percentages of each epithelial cell subpopulation.

(B) Linear regression analysis of LP and basal cell proportions by age for controls (WT) (n = 26 patients) and *BRCA2* carriers (n = 28 patients). The LP/basal ratio by patient provides additional validation as it accounts for technical factors that may have subtle effects on absolute cell numbers.

(C) TUNEL staining of representative control (WT) and *BRCA2* carrier tissues. Summary data obtained by counting four fields for 5 patients per genotype. ****p <0.0001 by Fisher's exact test.

Fig. 6

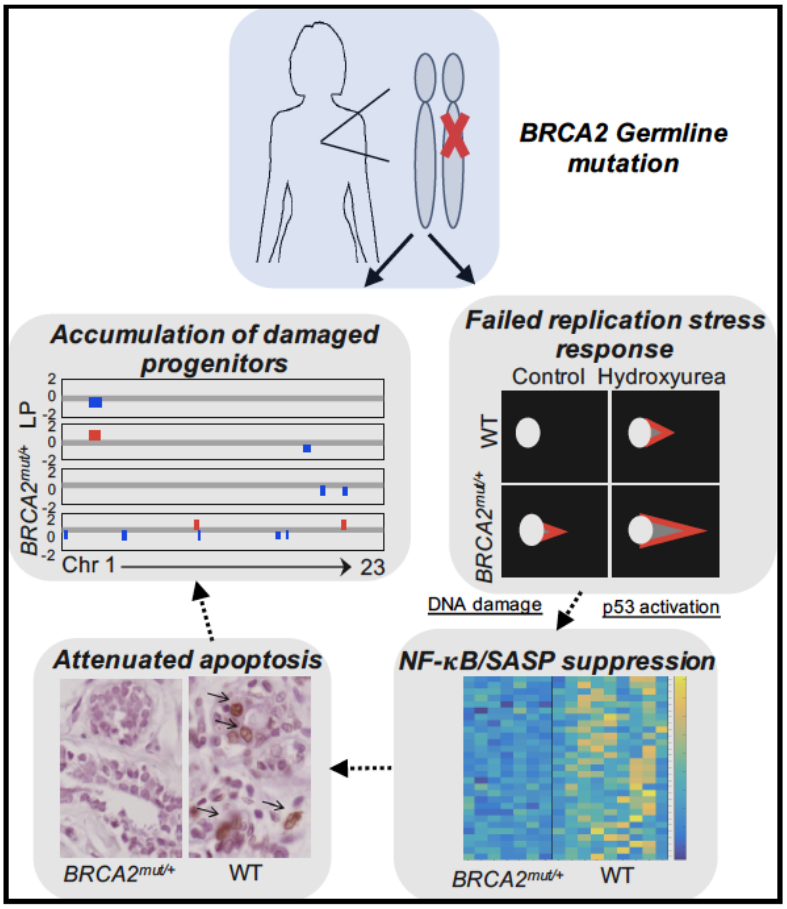


Fig. 6 Summary of findings in primary *BRCA2*$^{mut/+}$ breast tissues. Epithelial progenitor cells of heterozygous germline *BRCA2* carriers exhibit DNA damage, failed replication stress and damage responses, together with attenuated apoptosis. LOH analyses suggest these findings may reflect a haploinsufficient phenotype for *BRCA2* in vivo.

Chapter Five: Temporal analysis of copy number variations during tumor evolution

Silberman RE, Karaayvaz-Yildirim M, Roper J, Braverman J, Imada S, Schipperke E, Eng G, Kedrin D, Stone J, Demehri S, Ellisen L, Yilmaz Ö, Ben-David U, Amon A

RES performed and analyzed all experiments.
The experiments in Figure 1 were completed with assistance from UBD.
The experiments in Figure 2 were completed with assistance from JR, JB, and SI.
The samples of human tissues analyzed in Figure 3, Figure 5, and Figure S1 were collected with assistance from MK, ES, GE, DK, JS, SD, LE, and ÖY.

Abstract

Although aneuploidy, a state of genome imbalance, induces cellular stress and slows proliferation in untransformed cells, it is present in ~90% of tumors. To gain insight the role of aneuploidy in cancer, past studies have investigated the timing with which aneuploidy arises during tumorigenesis. Studies of colorectal carcinoma have argued that aneuploidy arises early in tumorigenesis, while more recent studies of lung and breast cancer have contradicted these findings. Here we assess the hypothesis that aneuploidy arises early in tumorigenesis. We define the karyotypes of samples of normal colons, early colorectal adenomas, late colorectal adenomas, and colorectal carcinomas. We find that whole chromosomal aneuploidy is only present in late colorectal adenomas and colorectal carcinomas. However, sub-chromosomal copy number variants (CNVs), a distinct form of genome imbalance, are present in every stage of colorectal carcinogenesis. The CNVs that appear throughout colorectal carcinogenesis are larger (>10 megabases (Mb)) than the size of CNVs common in healthy tissues. These large CNVs are absent from samples of normal colon. Modeling early adenoma formation in mice reveals that *APC* inactivation is insufficient to induce the accumulation of aneuploid cells but sufficient to generate large CNVs in early adenomas. We perform single-cell sequencing of samples of the early pre-malignant lesions seborrheic keratosis, melanotic nevi, and breast atypical hyperplasia. We find that while these lesions lack whole chromosomal aneuploidy, some contain large CNVs. We conclude that large sub-chromosomal CNVs arise early in tumorigenesis while whole chromosomal aneuploidy arises late in tumorigenesis. These findings propose that aneuploidy, but not large CNVs, faces negative selection during tumorigenesis.

Introduction

Aneuploidy—a state of genome imbalance which alters the copy numbers of whole chromosomes—is a hallmark of cancer. While aneuploidy is rare in healthy human tissues (Knouse et al. 2014), ~90% of tumors are aneuploid (Knouse et al. 2017). Aneuploidy increases cellular stress (Sheltzer et al. 2012; Oromendia et al. 2012; Santaguida et al. 2015), slows proliferation (Williams et al. 2008; Thorburn et al. 2013), and decreases cells' fitness (Sheltzer et al. 2011; Tang et al. 2011). Given its deleterious effects, it is surprising that aneuploidy characterizes a disease of hyper-proliferation and survival.

Aneuploidy appears to be both pro- and anti-tumorigenic in different contexts, depending on factors which include cell type, karyotype, and environmental conditions. Aneuploid mouse embryonic fibroblasts (MEFs) are no more tumorigenic than euploid MEFs (Sheltzer et al. 2017; Thomas et al. 2018), and trisomic mice and humans show decreased rates of solid tumor formation (Satgé et al. 1998; Baek et al. 2009; Reynolds et al. 2010; Yang and Reeves 2011). Mouse models that experience increased chromosomal instability (CIN), elevating the rate with which aneuploid cells are generated, show both enhanced (Foijer et al. 2014; Levine et al. 2017; Wijshake et al. 2012) and diminished (Weaver et al. 2007; Silk et al. 2013) tumor growth.

But aneuploid cells thrive under stress conditions (Rutledge et al. 2016; Pavelka et al. 2010; Yona et al. 2012). Moreover, because aneuploidy causes genome instability (Burrell et al. 2013; Lamm et al. 2016; Sheltzer et al. 2011), aneuploid populations are heterogeneous and unstable. Evolution to fitter, more tumorigenic karyotypes has been observed in aneuploid populations in culture and can drive aneuploid cells to emerge from euploid populations (Sheltzer et al. 2017; Thomas et al. 2018; Ben-David et al. 2014; Ben-David and Benvenisty 2012). These data suggest that even if most aneuploid cells lack the qualities which would pre-

dispose them to cancerous transformation, like enhanced proliferation and survival, aneuploid populations may evolve to generate rare, pro-tumorigenic karyotypes.

In order to understand the relationship between aneuploidy and cancer, researchers have sought to define when aneuploid cells accumulate during tumorigenesis. Analyses of pre-cancerous colorectal adenomas concluded that aneuploidy is present early in tumorigenesis, then progressively increases in late-stage adenomas and carcinomas (Vogelstein et al. 1988, 1989; Ried et al. 1996; Meijer et al. 1998). However, more recent studies of lung cancer and breast cancer have argued that aneuploidy arises late in tumorigenesis, at the carcinoma *in situ* stage (Teixeira et al. 2019; Casasent et al. 2018; Gao et al. 2016; Wang et al. 2014). These more recent reports did not assess the karyotypes of early pre-malignant lesions with single-cell precision, so we cannot exclude the possibility that early pre-malignant growths contain rare aneuploid cells, even if these cells fail to contribute significantly to the population of late pre-malignant lesions or primary tumors.

Colorectal cancer is not the only disease where aneuploidy is reported to arise early. Glioblastoma is characterized by the gain of chromosome 7 and this karyotype emerges early in the development of this disease (Gerstung et al. 2020). Similarly, loss of chromosome 3p drives tumorigenesis in clear cell renal cancer (Mitchell et al. 2018). Because these are both examples of highly recurrent, disease-specific karyotypes, these findings may indicate that distinct evolutionary paths exist for karyotypically homogeneous versus karyotypically heterogeneous cancers. Although colorectal carcinomas do share some common karyotype features, tending to gain chromosomes 13 and 20, they are much more heterogenous karyotypically than either glioblastoma or clear cell renal cancer (Knouse et al. 2017).

Should aneuploidy arise early in colorectal carcinoma, this would represent a unique evolutionary path among karyotypically heterogeneous cancers. One possible explanation for why colorectal carcinoma would have a different relationship with aneuploidy than other cancers is the prevalence of *APC* mutations in this disease. *APC* mutations are an early or initiating event in ~85% of spontaneous colorectal carcinomas (Kwong and Dove 2009). Colorectal carcinomas diverge between two mutually exclusive forms of genome instability. The ~85% of tumors which feature *APC* mutations have high CIN, while the remaining 15% are initiated by mutations in mismatch repair machinery and experience microsatellite instability (Zhang and Shay 2017). While *APC* is believed to drive tumorigenesis through dysregulation of the Wnt signaling pathway, *APC* also promotes microtubule stability (Kaplan et al. 2001). The possible link between *APC* mutations, microtubule stability, and CIN in colorectal carcinoma proposes the hypothesis that *APC* loss may induce CIN early in tumorigenesis, leading to the early accumulation of aneuploid cells and providing colorectal carcinoma with a unique pattern of karyotype evolution.

Here we evaluate the hypothesis that aneuploidy arises early in tumorigenesis. By karyotyping samples of normal colons, early- and late-stage colorectal adenomas, and colorectal carcinomas, we determine that chromosomal aneuploidy arises late in tumorigenesis in colorectal carcinoma. However, we observe that sub-chromosomal copy number variations (CNVs), a form of genome imbalance distinct from whole chromosomal aneuploidy, are absent in healthy colons but present in all stages of colorectal carcinogenesis. Using a mouse model of early tumorigenesis, we demonstrate that *APC* inactivation is insufficient to cause whole chromosomal aneuploidy to accumulate but can induce CNVs in cells of early adenomas. In order to determine whether the pre-malignancies of other tissues contain either type of genome imbalance, we

perform single-cell sequencing of pre-malignancies of the skin and breast. We demonstrate that whole chromosomal aneuploidy is absent in early pre-malignant lesions, but that some contain large CNVs. Our findings indicate that while chromosomal aneuploidy arises late in tumorigenesis, CNVs arise early in tumorigenesis.

Results

Aneuploidy arises late in the progression to colorectal carcinoma

Previous studies of karyotype evolution in colorectal carcinoma used the hybridization of DNA probes to determine the presence or absence of each chromosome (Vogelstein et al. 1989) or used comparative genomic hybridization to evaluate copy numbers across the whole genome (Meijer et al. 1998). Because these methods can overestimate the degree of aneuploidy in tissues (Knouse et al. 2014), we performed e-karyotyping, a computational method for defining karyotypes from gene expression microarray data (Ben-David et al. 2013). E-karyotyping simultaneously identifies both whole chromosomal aneuploidy and sub-chromosomal CNVs. Sub-chromosomal CNVs can range in size from a few kilobases (kb) to over a hundred megabases (Mb) (Henrichsen et al. 2009). While whole chromosomal aneuploidy results from failures in chromosome segregation (Gordon et al. 2012), CNVs result from many different types of DNA damage and damage repair failures (Conrad et al. 2010; Sharp et al. 2005; Hastings et al. 2009; Soto et al. 2018).

From publicly deposited microarray data, we performed e-karyotyping on samples of 38 healthy colons, 16 low grade adenomas, 13 high grade adenomas, and 27 colorectal carcinomas. In the development of colorectal carcinomas, low grade adenomas represent an earlier phase of tumorigenesis and carry a lower risk of malignant transformation while high grade adenomas are a later stage and considered high risk (Sievers et al. 2017).

Because our data came from bulk analyses of gene expression, we can only report on clonal imbalances that occur at a high degree. In order to quantify the amount of genome imbalance in each stage of tumorigenesis, we calculated the average genome imbalance (AGI) for each stage. First, we determined the total genome imbalance for each sample by summing the genomic distance in base pairs (bp) of all regions of the genome affected by either form of genome imbalance. A sample with no CNVs or whole chromosomal aneuploidy was assigned a total genome imbalance of 0 bp. We then averaged the total genome imbalances of all samples from the same disease stage to calculate the AGI of that stage. We calculated an AGI of $5.1x10^6$ bp (standard deviation (SD) $1.1x10^7$ bp) for low grade adenomas, $1.1x10^8$ bp (SD $1.7x10^8$ bp) for high grade adenomas, and $2.4x10^8$ bp (SD $2.2x10^8$ bp) for colorectal carcinomas (Fig. 1A). Each step of colorectal carcinogenesis was accompanied by a significant increase in AGI (Mann-Whitney U Test, $p < 0.05$). For context, these AGI scores represent 0%, 0.002%, 3.56%, and 7.77% of total genome size for each sample, respectively.

We observed no genome imbalances in any samples of normal colon (Fig. 1B). We did not expect to find any clonal whole chromosomal aneuploidy in these samples because even sub-clonal chromosome gains and losses are rare in healthy tissues (Knouse et al. 2014). We attribute the lack of CNVs identified in healthy colons to our high size threshold for CNV discovery. We only included CNVs greater than 10 Mb in our analysis because the risk of false identification increases dramatically for CNVs below this threshold (Ben-David et al. 2013). While small constitutional CNVs greater than 100 kb are present in over half of individuals, constitutive CNVs exceeding 1 Mb are only found in 1% of humans (Itsara et al. 2010). Large somatic CNVs are also rare. Less than 15% of the CNVs found in healthy brain and skin cells are greater than 10 Mb in size (Knouse et al. 2016).

We observed large (> 10 Mb) CNVs in all stages of colorectal carcinogenesis (Fig. 1B), indicating that large CNVs are a feature of pre-malignant and malignant colorectal growths. CNVs greater than 10 Mb appeared in early colorectal adenomas in the absence of whole chromosome gains and losses (Fig. 1C), suggesting that sub-chromosomal CNVs can arise before whole chromosomal aneuploidy during tumorigenesis.

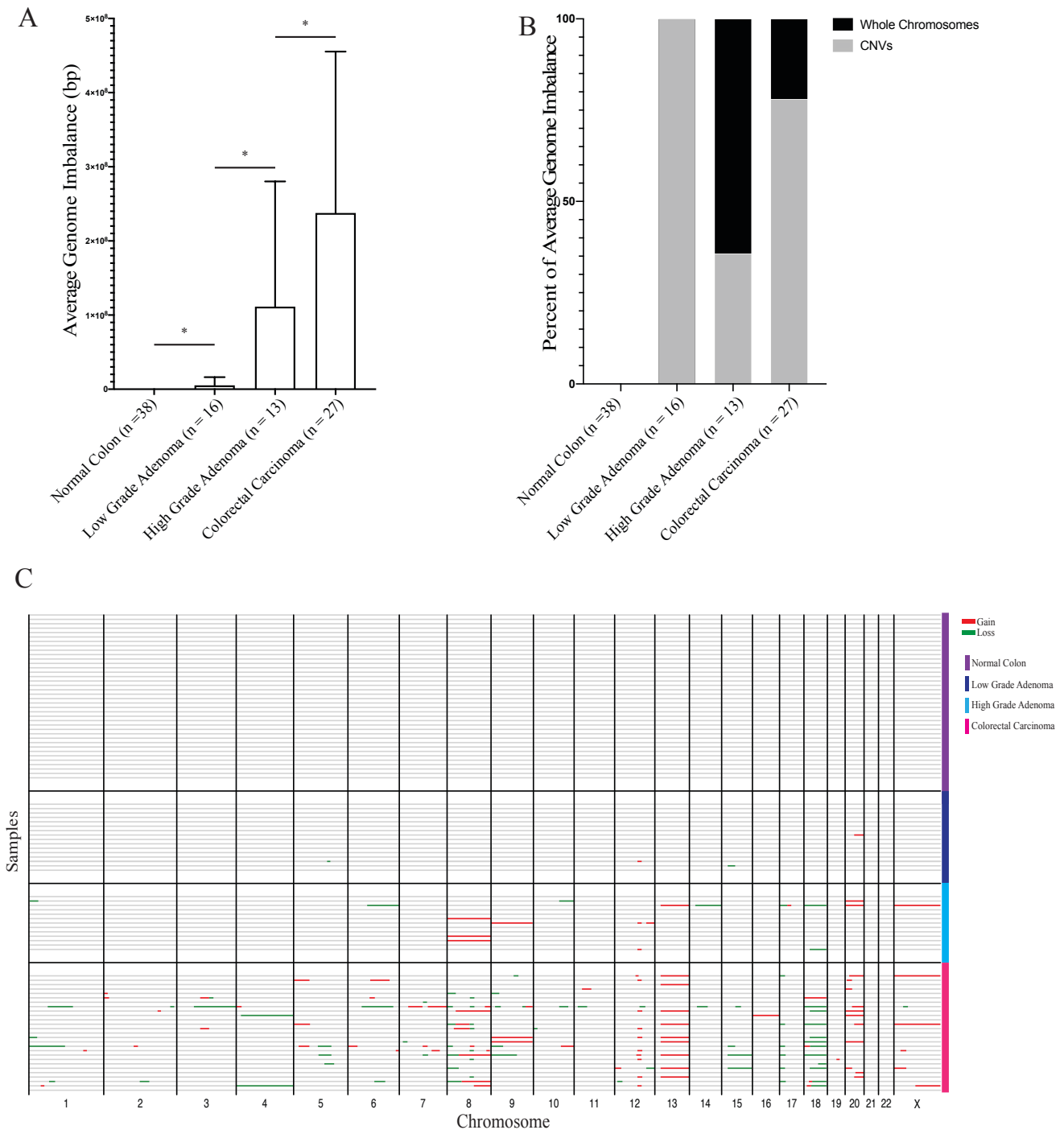


Figure 1: Whole chromosomal aneuploidy arises in late tumorigenesis in colorectal carcinomas

A. Average genome imbalance (AGI) measured in base pairs (bp). Averages plotted as boxes, with error bars depicting standard deviation. Adjacent groups compared by Mann-Whitney U test. * $p < 0.05$.

B. Fraction of AGI contributed by whole chromosomal aneuploidy or CNVs. Black bars represent the contribution to AGI from whole chromosomal aneuploidy and grey bars represent the contribution from CNVs.
C. Karyotypes of all samples. Each line represents one sample. Grey identifies regions with no copy number alterations, green identifies losses, and red identifies gains. Markers on the right y-axis indicate samples' tumor stage: normal colon (purple), low grade adenoma (dark blue), high grade adenoma (light blue), and colorectal carcinoma (pink).

Although we initially hypothesized that aneuploidy arises early in the development of colorectal carcinomas, our data did not support this hypothesis. Whole chromosomal aneuploidy was only present in late adenomas and colorectal carcinomas (Fig. 1B). We observed a dramatic shift in the source of genome imbalance between late pre-malignancy and invasive disease. Whole chromosomal aneuploidy accounted for 65% of the AGI of late pre-malignant adenomas but only 22% of the AGI of colorectal carcinomas (Fig. 1B). During the transition from late adenomas to colorectal carcinomas, the total genome imbalance attributed to whole chromosomal aneuploidy increased 1.5 fold while the total genome imbalance attributed to CNVs increased 9.7 fold. In colorectal carcinomas, whole chromosomal aneuploidy affected only 7 of the 23 chromosomes, while CNVs were found on 20 of 23 chromosomes (Fig. 1C).

In order to confirm that our results from e-karyotyping accurately depict the karyotypes of colorectal lesions, we performed single-cell sequencing on 1 high grade adenoma, following previously published methods (Knouse et al. 2014, 2016). We performed single-cell sequencing because this method directly quantifies DNA copy number, as opposed to e-karyotyping, which translates gene expression data into DNA copy number. The karyotypes of 20 single cells isolated from 1 high grade adenoma contained both whole chromosomal aneuploidy and CNVs (Sup. Fig. 1A). Single-cell sequencing permits the identification of CNVs as small as 2 Mb, allowing us to observe CNVs ranging from 2 Mb to 17 Mb in this sample of high grade adenoma. The karyotypes of individual high grade adenoma cells were similar to the karyotypes

we defined by bulk sampling high grade adenomas. All cells of the high grade adenoma gained chromosome 20, which was clonally gained in 2/16 high grade adenomas analyzed by e-karyotyping. We single-cell sequenced a sample of normal colon isolated from healthy tissue adjacent to the high grade adenoma. Of 20 cells, none contained genomic imbalances (Sup. Fig. 1A), which reflects our e-karyotyping of normal colon.

APC inactivation is insufficient to induce whole chromosomal aneuploidy in adenomas in mice

The bulk nature of e-karyotyping does not allow us to exclude the possibility that rare aneuploid cells exist in early pre-malignant colorectal adenomas. We remained interested in whether *APC* inactivation can drive the accumulation of aneuploid cells during early colorectal carcinogenesis. We could not collect samples of early human colorectal adenomas for single-cell sequencing, so we employed the $Apc^{fl/fl};Villin^{CreER}$ mouse model (Roper et al. 2017) to explore the impact of *APC* inactivation on karyotype evolution in early tumorigenesis.

Following colonoscopy-guided sub-mucosal injection of 100 μM tamoxifen into the distal colon, animals developed early colorectal adenomas within 2 weeks. *APC* inactivation was confirmed by immunohistochemistry demonstrating the nuclear localization of beta-catenin (Fig. 2A). We identified no genome imbalances in adenomas harvested 2 weeks or 4 weeks post-injection of tamoxifen (n = 2 mice, 42 cells), suggesting that *APC* inactivation in early adenomas is insufficient for the accumulation of aneuploid cells (Fig. 2B and Fig. 2C.).

But human adenomas rarely lose both copies of *APC* simultaneously. In order to create a more faithful model of tumorigenesis, we injected 100 μM tamoxifen into the sub-mucosa of the distal colon of animals bearing only one floxed allele of *APC*, then waited for loss of heterozygosity to occur to drive adenoma formation. We observed the presence of CNVs in the adenoma of 1 animal euthanized 3 months post-injection. Of 34 cells sequenced, 2 cells contained CNVs and the remaining 32 cells contained no genome imbalances (Fig. 2D). All of

the CNVs we identified in these cells were greater than 10 Mb in size. The observation that this early mouse adenoma contained CNVs greater than 10 Mb mirrors our findings in human early adenomas, where we also observed that some samples featured CNVs greater than 10 Mb (Fig. 1C). Because we discovered large CNVs in early adenomas of mice and humans using different karyotyping techniques, we are reluctant to directly compare the size or frequency of CNVs between these two experiments.

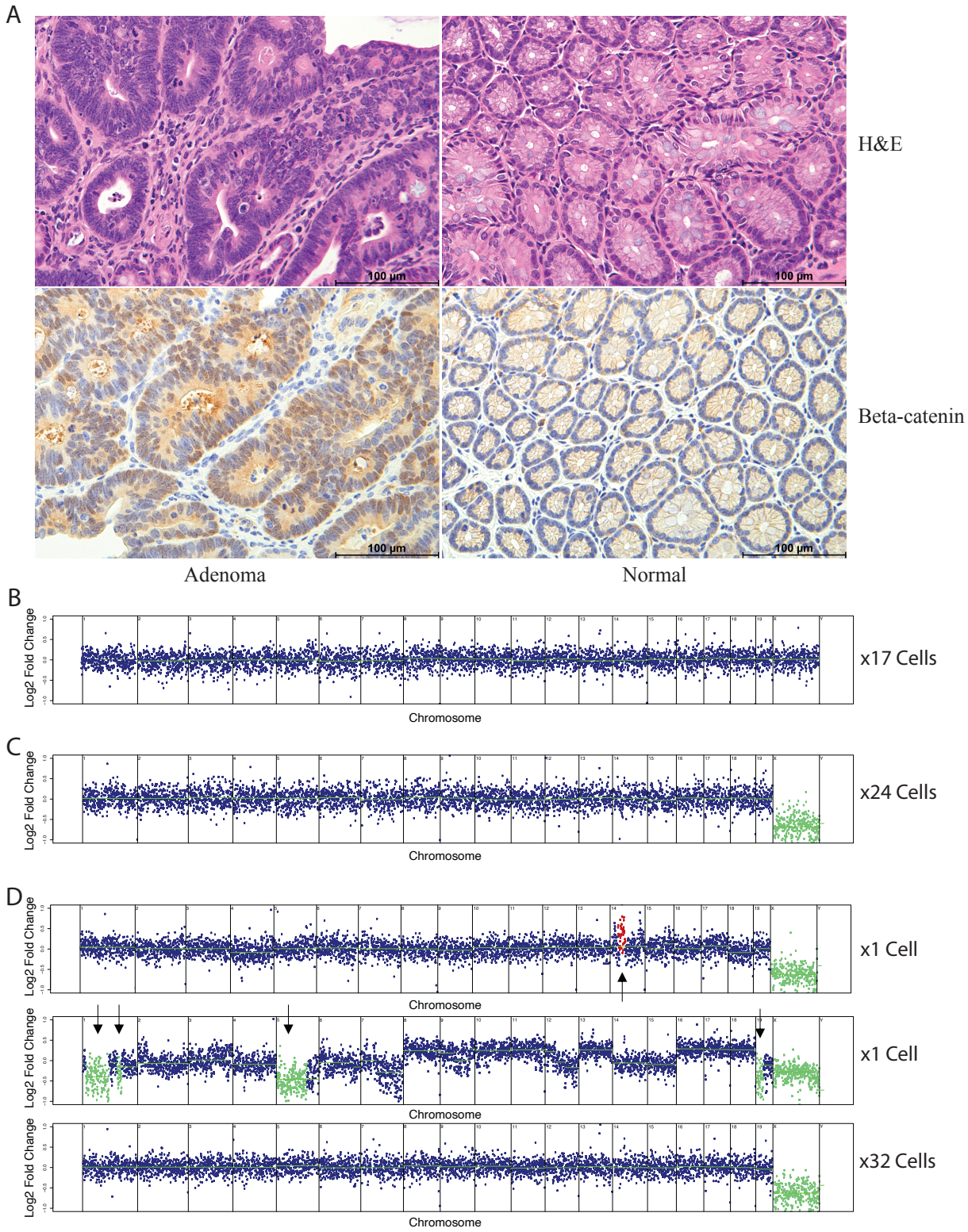


Figure 2: *APC* inactivation is insufficient to induce whole chromosomal aneuploidy in early adenomas in mice

A. Immunohistochemistry staining for beta-catenin in adenoma and normal colon from *Apc*$^{fl/fl}$;*Villin*^{CreER} mice.

B. Representative karyotype from single-cell sequencing of *Apc*$^{fl/fl}$;*Villin*^{CreER} adenoma harvested 2 weeks post-injection of 100 μM tamoxifen.
C. Representative karyotype from single-cell sequencing of *Apc*$^{fl/fl}$;*Villin*^{CreER} adenoma harvested 4 weeks post-injection of 100 μM tamoxifen.
D. Three cells from an *Apc*$^{fl/+}$;*Villin*^{CreER} adenoma harvested 3 months post-injection of 100 μM tamoxifen. 2 cells represent the only non-euploid karyotypes and 1 cell is representative of the 32 remaining euploid cells. Arrows indicate the sites of CNVs.

APC inactivation does not drive malignant transformation in mice. The adenomas we sequenced were benign, suggesting that hyperplastic growth alone is sufficient for the generation of CNVs. From these experiments we concluded that though *APC* inactivation cannot drive the accumulation of whole chromosomal aneuploidy in early tumorigenesis, it can drive the accumulation of cells with large CNVs in mice.

Early pre-malignant lesions lack whole chromosomal aneuploidy but feature large CNVs

In order to determine whether the presence of large CNVs in early tumorigenesis is specific to cells lacking *APC* function or to the intestinal epithelium, we defined the karyotypes of pre-malignancies of the breast and skin with single-cell precision. We single-cell sequenced 5 seborrheic keratoses, 3 breast atypical hyperplasias, and 5 melanotic nevi. Although single-cell sequencing permits the identification of CNVs as small as 2 Mb, we remained primarily interested in the distribution of large (>10 Mb) CNVs in pre-malignant lesions as compared to healthy tissues. Given our observation that large CNVs are absent in healthy colons but are present in early pre-malignant adenomas, we hypothesized that the presence of large CNVs may be a common feature of hyperplasias.

Seborrheic keratosis is an extremely common skin tumor—affecting 80-100% of people over the age of 50 (Engel et al. 1988; Yeatman et al. 1997; Herd et al. 2006)—which results from the overproliferation of immature keratinocytes (Phulari et al. 2014). Although basal cell carcinomas and other skin cancers can arise within seborrheic keratoses (Bedir et al. 2014;

Rajabi et al. 2012), most are considered low risk. Seborrheic keratoses are present in ~30% of people under the age of 30, indicating that these growths can persist for many decades (Gill et al. 2000). Because of this, seborrheic keratoses are an ideal substrate to study the relationship between long-term hyperplastic growth and genome imbalance.

Of the 5 seborrheic keratoses we analyzed, representing two individuals, none contained cells featuring whole chromosomal aneuploidy. However, 3 of 5 seborrheic keratoses contained CNVs larger than 10 Mb. Though the majority of CNVs we described were non-clonal, in one seborrheic keratosis every cell featured the loss of the same 24 Mb region on chromosome 10 (Fig. 3A). The size of CNVs in seborrheic keratoses was significantly larger than the size of CNVs in normal skin cells, which we collected from healthy skin adjacent to the seborrheic keratoses (Fig. 5A; Mann-Whitney Test, $p < 0.01$). We concluded that large CNVs are a feature of seborrheic keratoses, but not healthy skin.

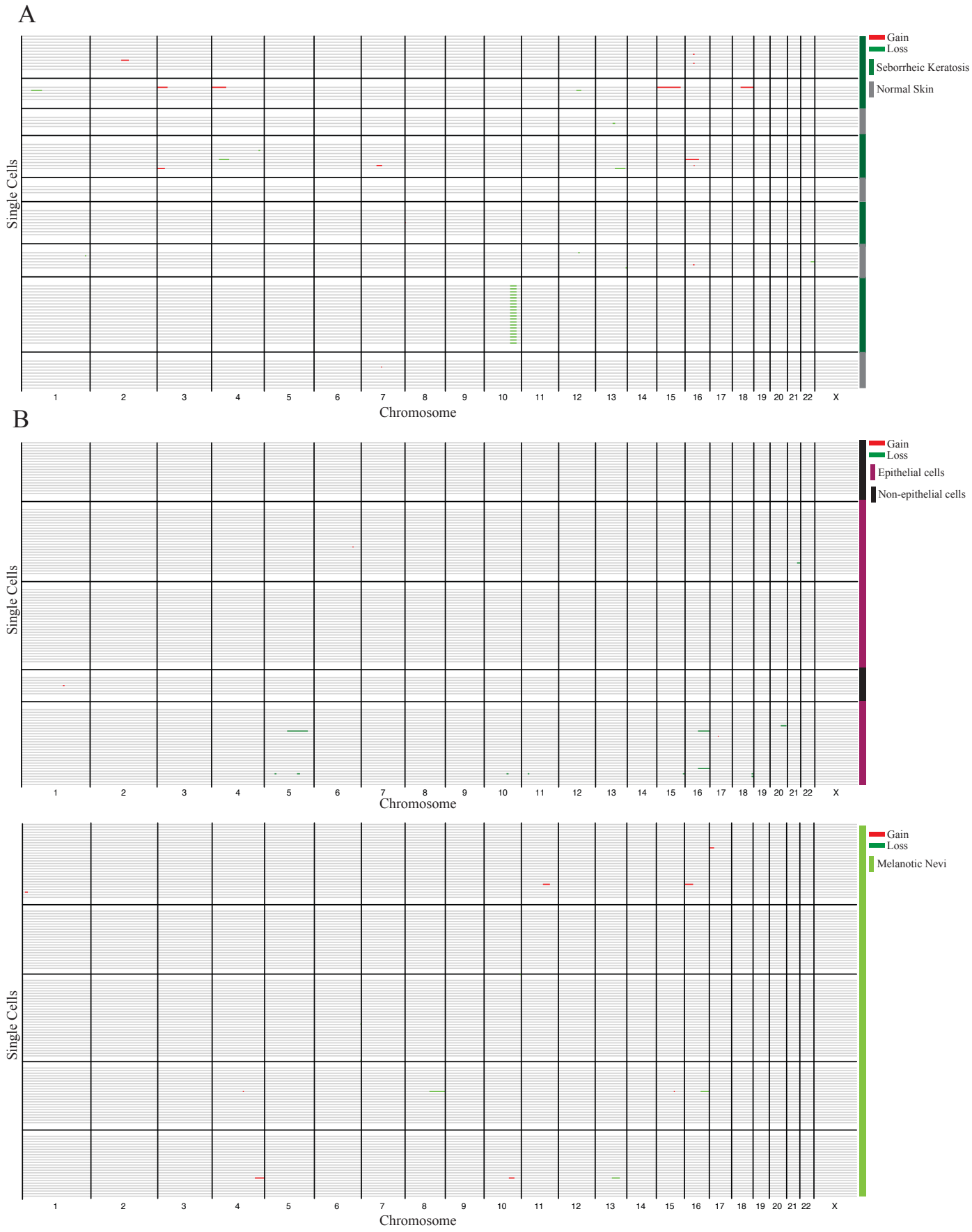

A
Gain
Loss
Seborrheic Keratosis
Normal Skin
Single Cells
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X
Chromosome
B
Gain
Loss
Epithelial cells
Non-epithelial cells
Single Cells
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X
Chromosome
Gain
Loss
Melanotic Nevi
Single Cells
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X
Chromosome

Figure 3: Early-premalignant cells lack whole chromosomal aneuploidy
A. Single-cell karyotypes of all samples. Each line represents one cell. Cells from the same sample are bounded by black lines. Grey identifies regions with no copy number alteration, green identifies losses, and red identifies gains. Markers on the right y-axis indicate samples' origin: seborrheic keratoses (green), healthy skin cells (grey).
B. Markers on the right y-axis indicate samples' origin: atypical hyperplasia epithelial cells (maroon) and atypical hyperplasia non-epithelial cells (black).
C. Markers on the right y-axis indicate samples' origin: melanotic nevi (light green)

Similar to colorectal cancer, invasive breast cancer develops through successive pre-malignant stages (Fig. 4). While atypical hyperplasia is a relatively early stage of tumorigenesis—occurring before the last stage of pre-malignancy, ductal carcinoma *in situ*—diagnosis of atypical hyperplasia caries a substantial cancer risk (Hartmann et al. 2015). The cumulative incidence of invasive breast cancer for women diagnosed with atypical hyperplasia is 30% after 25 years (Page et al. 2003).

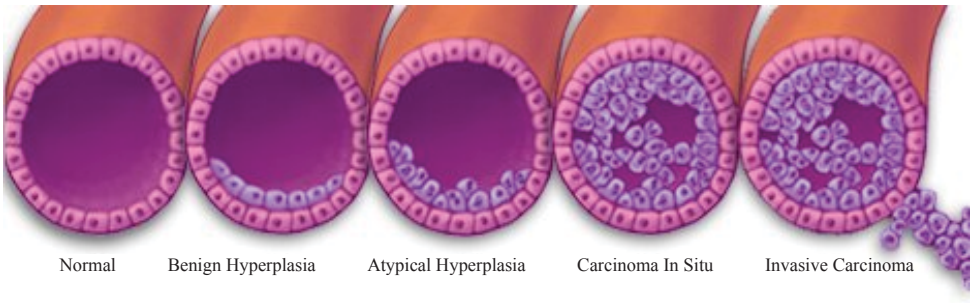


Figure 4: Invasive breast cancer develops through multiple pre-malignant stages. Schematic representation of the stages of tumorigenesis in invasive breast cancer. Tubular structures represent ducts. Figure adapted from the Mayo Clinic "Atypical hyperplasia of the breast."

We collected samples of breast atypical hyperplasia and used the epithelial cell surface marker EpCAM to identify epithelial cells, which are the cell type that compose atypical hyperplasia (Bombonati and Sgroi 2011). Of the 82 EpCAM-positive epithelial cells we sequenced, representing three individuals, none had gained or lost a whole chromosome (Fig. 3B). However, 5 cells, from two individuals, included CNVs greater than 10 Mb. Of 27 non-

epithelial control cells, 0 had CNVs greater than 10 Mb. The size of CNVs in epithelial cells from atypical hyperplasia were significantly greater than those in non-epithelial cells (Fig. 5B; Mann-Whitney U Test, $p < 0.05$). We conclude that CNVs greater than 10 Mb characterize some cases of breast atypical hyperplasia.

Melanotic nevi are a pre-cursor to the highly aneuploid (Knouse et al. 2017) disease cutaneous melanoma. Approximately one third of melanomas arise directly from melanotic nevi (Bevona et al. 2003; Lin et al. 2015). However, the risk that any single nevus will give rise to a melanoma is very low (Tsao et al. 2003). Melanotic nevi consist of melanocytes which have overproliferated (Damsky and Bosenberg 2017) frequently as a result of activating mutations in *BRAF* (Poynter et al. 2006; Roh et al. 2015; Piris et al. 2015). The development of melanoma follows similar steps as other invasive cancers, like colorectal carcinoma and invasive breast cancer. This progression initiates with benign nevi which can evolve into melanoma *in situ,* before terminating in invasive melanoma (Clark et al. 1984).

In order to determine whether melanotic nevi contain genome imbalances, we single-cell sequenced 5 melanotic lesions. Because we collected these samples based on clinical diagnosis, we assembled a collection of 5 melanotic lesions, from 5 individuals, which were later determined to include 2 melanotic nevi, 2 benign melanotic hyperplasias, and 1 case of lentigo. No cells isolated from melanotic nevi contained whole chromosomal aneuploidy (Fig. 3C) and 5 cells contained CNVs greater than 10Mb in size. There was no difference in the prevalence of cells with large CNVs between different sub-classes of melanotic lesion. There was also no significant difference in the size of CNVs in melanotic nevi as compared to healthy skin cells (Fig. 5A).

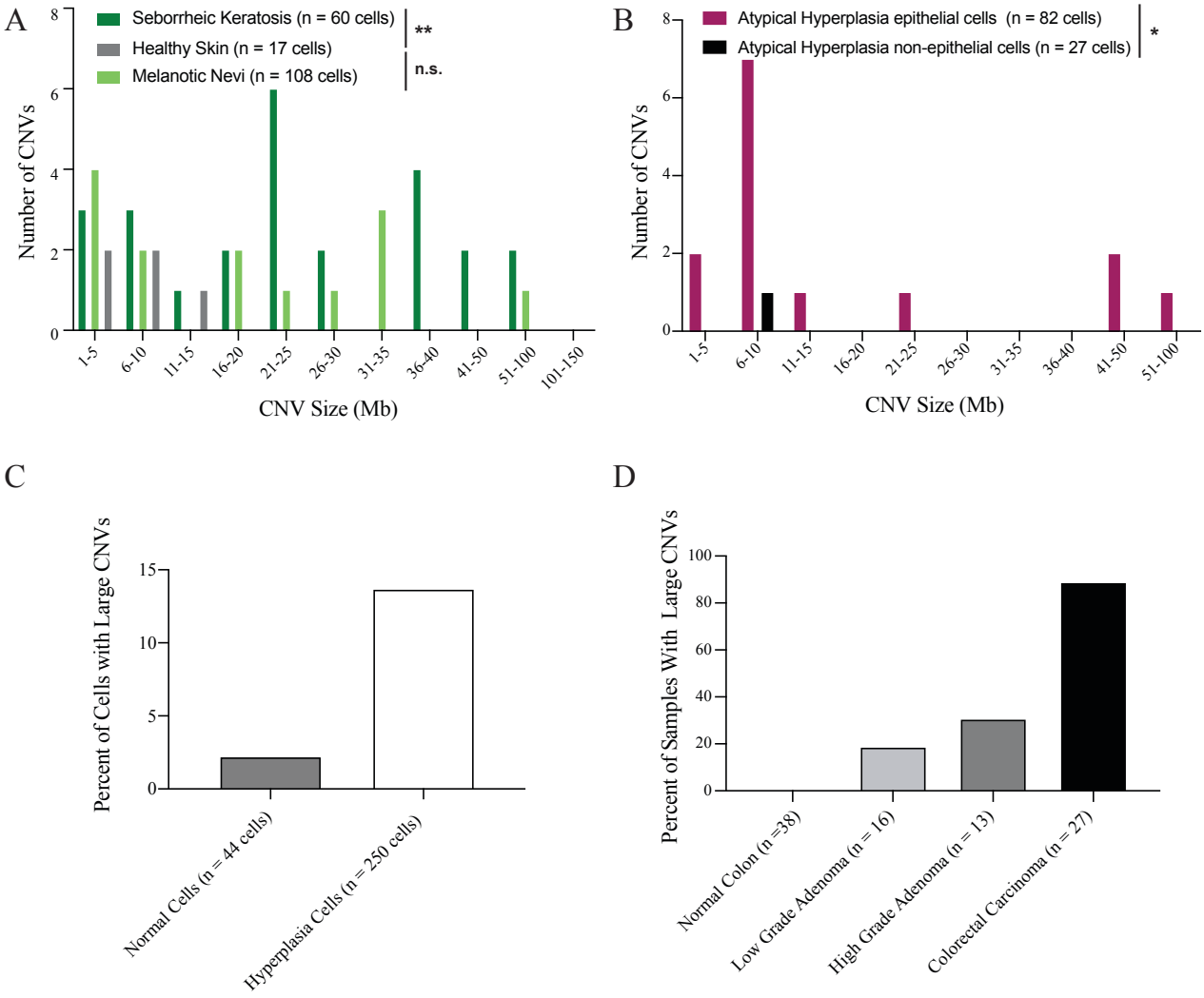


Figure 5: Some pre-malignant lesions contain large CNVs
A. Distribution of the size of CNVs in seborrheic keratoses, melanotic nevi, and healthy skin. ** $p < 0.01$. * $p < 0.05$. Not significant (n.s.).
B. Distribution of the size of CNVs in atypical hyperplasia epithelial and non-epithelial cells. * $p < 0.05$.
C. Graph depicting percent of cells featuring CNVs larger than 10 Mb as determined by single-cell sequencing.
D. Graph depicting percent of cells featuring CNVs larger than 10 Mb as determined by e-karyotyping.

Discussion

Here we provide evidence that whole chromosomal aneuploidy arises during late tumorigenesis in colorectal carcinomas. While we observed no clonal CNVs greater than 10 Mb in samples of normal colon, we observed large clonal CNVs in early colorectal adenomas, late

colorectal adenomas, and colorectal carcinomas, suggesting that large sub-chromosomal CNVs arise early in tumorigenesis and characterize all stages of colorectal carcinogenesis. We demonstrate that *APC* inactivation in mice is insufficient to cause the accumulation of cells with whole chromosomal aneuploidy in early colorectal adenomas. We also observed that loss of heterozygosity for *APC* can drive the accumulation of cells with large CNVs in early mouse adenomas. By single-cell sequencing samples of the early pre-malignancies seborrheic keratosis, melanotic nevi, and breast atypical hyperplasia, we observed that these early pre-malignant lesions lack whole chromosomal aneuploidy, but that some growths contain CNVs greater than 10 Mb.

Sub-chromosomal CNVs arise during early tumorigenesis

Single-cell sequencing healthy skin cells and non-epithelial breast cells from women with atypical hyperplasia confirmed that CNVs greater than 10 Mb are rare in healthy tissues. Of the 44 normal cells we sequenced, only ~2% contained a CNV larger than 10 Mb. In contrast, ~13% of all cells isolated from pre-malignant lesions featured CNVs greater than 10 Mb (Fig. 5C). While the results of our e-karyotyping cannot be directly compared to the results of single-cell sequencing, we noted that while normal colon samples lacked large CNVs, 19% of early colorectal adenomas contained clonal CNVs greater than 10 Mb (Fig. 5D). We conclude that CNVs greater than 10 Mb arise during early tumorigenesis.

CNVs larger than 10 Mb have been observed in the highly tumorigenic but histologically normal breast tissue of women harboring *BRCA2* mutations (Karaayvaz-Yildirim et al. 2020) and in Barrett's esophagus, a precursor of esophageal carcinoma (Ross-Innes et al. 2015). Longitudinal studies of individuals with Barrett's esophagus have described that pre-malignant growths with large CNVs can persist for many decades without transforming into esophageal carcinoma. In Barrett's esophagus cases which do transform into esophageal carcinoma, whole-

chromosomal aneuploidy arises shortly before malignant transformation (Killcoyne et al. 2020; Martinez et al. 2018). This suggests that even though large CNVs are a common feature of early pre-malignancy, they may not drive malignant transformation.

Even if large CNVs do not independently drive tumorigenesis, we observed that large CNVs progressively increase in late pre-malignant adenomas and colorectal carcinomas. Analyses of the karyotypes of tumors from many cancers have indicated that CNVs can enrich the genome for pro-proliferation gene copies and deplete anti-proliferation gene copies (Beroukhim et al. 2010), suggesting that CNVs can enhance tumorigenesis by altering the copy numbers of key genes. The presence of highly clonal large CNVs in some early colorectal adenomas and seborrheic keratoses may indicate that these large CNVs provide a fitness benefit, allowing cells to propagate within the early pre-malignant cell population. Highly clonal large CNVs are also found in late colorectal adenomas and colorectal carcinomas, suggesting these large CNVs may also provide a fitness benefit at later stages. To this point, we identified many large CNVs that are present in multiple samples of colorectal carcinoma. We speculate that these recurrent CNVs may indicate the presence of key genes whose altered copy numbers enhance colorectal carcinoma cells' fitness.

Whole chromosomal aneuploidy arises during late tumorigenesis

Because the majority of coding mutations are tolerated in tumors, it has been proposed that positive selection dominates tumor evolution while negative selection has a much smaller impact (Martincorena et al. 2017). However, we propose that whole chromosomal aneuploidy undergoes negative selection during tumor evolution. Whole chromosomal copy number alterations were much less diverse than sub-chromosomal CNVs, affecting just 7 of 23 chromosomes in colorectal carcinomas while CNVs occurred on 20 of 23 chromosomes. While we saw some novel whole chromosomal copy number alterations in colorectal carcinomas as

compared to late adenomas, the karyotype diversity of whole-chromosomal aneuploidy in pre-malignant lesions was similar to that of mature tumors. Conversely, the diversity of CNVs expanded dramatically during the transition from late pre-malignancy to invasive disease.

Whole chromosomal aneuploidy also generated a narrow range of total genome imbalance. While we observed a 9.7 fold increase in the total genomic distance affected by CNVs during the transition from late pre-malignancy to invasive disease, we observed only a modest 1.5 fold increase in the total genomic distance affected by whole chromosomal aneuploidy. From these data, we speculate that aneuploidy faces negative selection, restricting the diversity of whole chromosomal copy number alterations tolerated during tumor evolution and limiting the total contribution to genome imbalance generated by whole chromosomal aneuploidy.

Our proposal that aneuploidy faces negative selection during tumorigenesis aligns with our understanding of the impact of aneuploidy on non-transformed cells, where it induces cellular stress (Sheltzer et al. 2012; Torres et al. 2007; Oromendia et al. 2012) and slows proliferation (Williams et al. 2008; Thorburn et al. 2013). Previous studies have shown that aneuploid cells are less fit *in vivo* and are selected against in regenerating tissues (Pfau et al. 2016). We propose that the deleterious effect of aneuploidy on cell fitness persist during early tumorigenesis and that negative selection could explain the lack of whole chromosomal aneuploidy in early pre-malignant lesions.

We speculate that the emergence of whole chromosomal aneuploidy in late tumorigenesis may coincide with the development of aneuploidy tolerating mutations (Torres et al. 2010), or with a shift towards environmental conditions that favor the accumulation of aneuploid cells. Although they are less fit under normal conditions, aneuploid cells thrive under stress conditions

(Rutledge et al. 2016; Pavelka et al. 2010; Yona et al. 2012). We speculate that during late tumorigenesis, stress conditions in the microenvironment such as hypoxia (Muz et al. 2015), tumor acidification (Huber et al. 2017), or the buildup of reactive oxygen species (Weinberg et al. 2019) may allow aneuploid cells to accumulate.

Studies of esophageal carcinoma have demonstrated that whole chromosomal aneuploidy arises over a brief period in late tumorigenesis immediately preceding malignant transformation (Martinez et al. 2018; Ross-Innes et al. 2015). Evidence from invasive breast cancer indicates that that aneuploidy evolves in a punctuated burst, rapidly forming clones which then stably grow to form the primary tumor (Gao et al. 2016). Our observation that whole chromosomal aneuploidy arises in late tumorigenesis in colorectal carcinoma and occupies a narrow window of karyotype diversity and total genome imbalance supports these findings. Our observation that large CNVs are present throughout tumorigenesis and progressively expand in number and diversity proposes that whole chromosomal aneuploidy and sub-chromosomal CNVs follow distinct evolutionary paths during tumorigenesis (Fig. 6). We speculate that these evolutionary differences reflect the different fitness consequences of these two types of genome imbalance, with whole chromosomal aneuploidy carrying a more severe fitness defect that large CNVs.

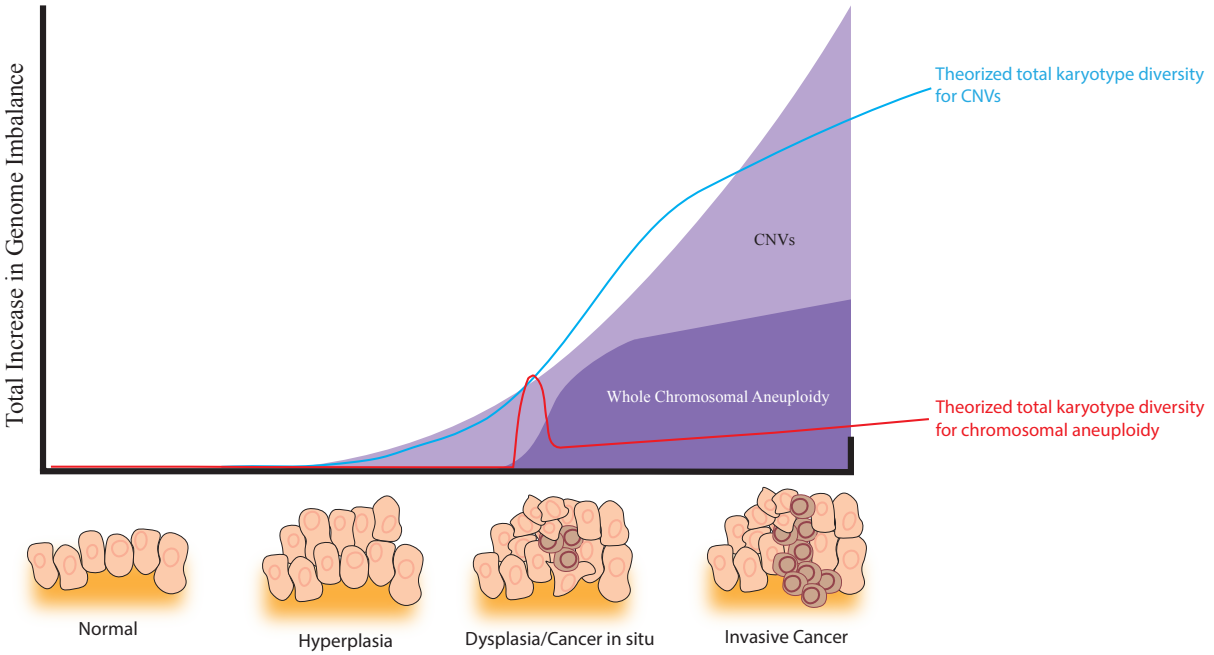


Figure 6: Sub-chromosomal CNVs accumulates gradually throughout tumorigenesis while whole chromosomal aneuploidy arises over a short burst. This model reflects the hypothesis that CNVs carry only a minor fitness penalty compared to whole chromosomal aneuploidy.

Methods

E-karyotyping

Microarray data were accessed via Gene Expression Omnibus (GSE37364), and translated into karyotypes using methods previously described (Ben-David et al. 2013). The original gene expression profiling was performed with Affymetrix HGU133 Plus 2.0 mircroarrays.

Tissue Sources

Samples of seborrheic keratosis and high-grade colorectal adenoma were isolated during autopsy at Massachusetts General Hospital through the Department of Pathology. Melanotic nevi were collected at Massachusetts General Hospital through the Cutaneous Biology Research Center. Samples of atypical hyperplasia were collected at the Dana-Farber Cancer Institute

through the Breast Oncology Center. All human procedures were approved by the Massachusetts Institute of Technology Committee of the Use of Humans as Experimental Subjects.

Isolating Human Cells

Before isolation of single cells, melanotic nevi were dissected using a Zeiss Stemi DV4 microscope so that only the brown-marked melanotic region was isolated. Single cells were isolated from seborrheic keratoses, high-grade adenoma, and melanocytic nevi by mincing the tissue in ice cold Hanks Balanced Salt Solution (HBSS) (Thermo Fisher Scientific #14025092). The tissue was then incubated in Dulbecco's Modified Eagle Medium (DMEM) (Thermo Fisher Scientific #11995-065) with 1 mg/mL collagenase (Millipore Sigma # C0130) for 30-90 minutes at 37°C on a rotating drum. Every 15-30 minutes the tissues were pipetted up and down with a 1000 mL Rainin pipette using a filtered pipette tip cut off at the end to widen the opening.

Once the mixture appeared cloudy and the pieces of tissue had shrunk or disappeared, the mixture was pipetted up and down for 1 minute with a glass Pasteur pipette. The mixture was passed through a 70 μM nylon strainer (Corning #431751), centrifuged at 300g for 5 minutes at 4°C, and resuspended in DMEM supplemented with 10% fetal bovine serum 2 mM glutamine, and 100 U/mL penicillin and streptomycin. Cells from seborrheic keratoses, high-grade adenoma, and melanocytic nevi were then isolated based on morphology by microaspiration, as described previously (Knouse et al. 2014).

Atypical hyperplasia samples were dissociated as previously described (Karaayvaz-Yildirim et al. 2020). After isolation, cells were resuspended in PBS with 5% FBS and stained for EpCAM (BD #324205). EpCAM-positive cells were isolated using a BD FACSAria II cell sorter. DAPI exclusion was used as a viability marker. Following collection of EpCAM-positive

cells, individual cells were collected for single-cell sequencing by micro-aspiration, as previously described (Knouse et al. 2014)

Single-cell sequencing

Single-cell sequencing and copy number analyses were completed as previously described (Knouse et al. 2014, 2016). As in Knouse et al. (2014), we excluded any cell which had a variability score (VS)—calculated by averaging the standard deviation of corrected read copies across the three most variable autosomes—greater than 0.34. Knouse et al. (2016) described that CNVs as small as 2 Mb could be identified in cells with a VS < 0.26. In this study, we established additional the cutoffs for determining what size of CNVs could be identified in cells with where 0.26 < VS <0.34. These cutoffs were identified based on probability of calling false positive CNVs at various VSs. At VS < 0.26 CNVs greater than 2 Mb were accepted, at VS < 0.27 CNVs greater than 5 Mb were accepted, at VS < 0.28 CNVs greater than 10 Mb were accepted, and at VS < 0.29 CNVs greater than 15 Mb were accepted, at VS < 0.30 CNVs greater than 20 Mb were accepted, at VS < 0.32 CNVs greater than 30 Mb were accepted.

For tissues featuring highly recurrent CNVs which did not meet the cutoffs in every cell that appeared to contain this imbalance, the sequences from all individual cells were pooled then analyzed as a single entity. Any CNVs which were identified from these pooled data were confirmed to be highly recurrent. This allowed us to call CNVs that just missed cutoffs in individual cells when there was strong evidence that these CNVs were present in all cells.

Mouse strains

Using a previously published method and mouse strain, $Apc^{fl/fl};Villin^{CreER}$ or $Apc^{fl/+};Villin^{CreER}$ mice were injected with 100 uM tamoxifen by sub-mucosal injection (Roper et

al. 2017). All animal studies and procedures were approved by the MIT Institutional Animal Care and Use Committee.

Isolating Mouse Intestinal Cells

After euthanasia, mouse intestines were removed and opened longitudinally. Adenomas and samples of adjacent normal colon were harvested. The samples were then agitated in cold PBS until all fecal matter was washed from the tissue. Samples were then placed in ice cold PBS with 5 mM EDTA and placed, in ice, on a shaker for 1 hour. The epithelium was scraped away from the underlying connective tissue using metal forceps and resuspended in PBS by pipetting up and down with a 10 mL pipette. The samples were centrifuged at 150g for 5 min at 4°C, then resuspended in DMEM with 200 U/mL type II collagenase (Thermo Fisher Scientific #17101015) and 125 ug/mL dispase (STEMCELL Technologies #07913). Samples were incubated at 37°C for 1 hour, then pipetted up and down for 1 minute with a glass Pasteur pipette. The sample was centrifuged at 300g for 5 min at 4°C, filtered through a 70 μM strainer, and resuspended in PBS with 5% FBS. Cells were stained for EpCAM (Thermo Fisher Scientific #12-5791-82) and single cells were deposited into 96-well PCR plates containing 10 uL of lysis mix from the Sigma-Aldrich WGA v4 kit using a BD Biosciences FACSAria II cell sorter. DAPI exclusion was used as a viability marker.

Immunohistochemistry

Fresh tissues were fixed in 10% formalin in PBS for 16-24 hours, then embedded in paraffin. Tissues were sectioned into 4 μM slices. For ß-catenin staining, antigen retrieval was performed in Borg Decloaker RTU (BioCare Medical #BD1000) at pH 6 at 97°C for 20 minutes. The following antibodies were used: mouse monoclonal ß-catenin (1:200; BD Biosciences #610154), mouse-on-mouse-HRP-polymer (BioCare Medical #MM620).

References

Baek KH, Zaslavsky A, Lynch RC, Britt C, Okada Y, Siarey RJ, Lensch MW, Park IH, Yoon SS, Minami T, et al. 2009. Down's syndrome suppression of tumour growth and the role of the calcineurin inhibitor DSCR1. *Nature* **459**: 1126–1130.

Bedir R, Yurdakul C, Güçer H, Ehitog IS¸, Lu ˘. 2014. Basal Cell Carcinoma Arising within Seborrheic Keratosis Dermatology Section. *J Clin Diagnostic Res* **8**: 6–07.

Ben-David U, Arad G, Weissbein U, Mandefro B, Maimon A, Golan-Lev T, Narwani K, Clark AT, Andrews PW, Benvenisty N, et al. 2014. Aneuploidy induces profound changes in gene expression, proliferation and tumorigenicity of human pluripotent stem cells. *Nat Commun* **5**: 1–11.

Ben-David U, Benvenisty N. 2012. High prevalence of evolutionarily conserved and species-specific genomic aberrations in mouse pluripotent stem cells. *Stem Cells* **30**: 612–622.

Ben-David U, Mayshar Y, Benvenisty N. 2013. Virtual karyotyping of pluripotent stem cells on the basis of their global gene expression profiles. *Nat Protoc* **8**: 989–997.

Beroukhim R, Mermel CH, Porter D, Wei G, Raychaudhuri S, Donovan J, Barretina J, Boehm JS, Dobson J, Urashima M, et al. 2010. The landscape of somatic copy-number alteration across human cancers. *Nature* **463**: 899–905.

Bevona C, Goggins W, Quinn T, Fullerton J, Tsao H, Corona R. 2003. Cutaneous Melanomas Associated with Nevi. *Arch Dermatol* **139**: 1620–1624.

Bombonati A, Sgroi DC. 2011. The molecular pathology of breast cancer progression. *J Pathol* **223**: 308–318.

Burrell RA, McClelland SE, Endesfelder D, Groth P, Weller MC, Shaikh N, Domingo E, Kanu N, Dewhurst SM, Gronroos E, et al. 2013. Replication stress links structural and numerical cancer chromosomal instability. *Nature* **494**: 492–496.

Casasent AK, Schalck A, Gao R, Sei E, Long A, Pangburn W, Casasent T, Meric-Bernstam F, Edgerton ME, Navin NE. 2018. Multiclonal Invasion in Breast Tumors Identified by Topographic Single Cell Sequencing. *Cell* **172**: 205-217.e12.

Clark WH, Elder DE, Guerry D, Epstein MN, Greene MH, Van Horn M. 1984. A study of tumor progression: The precursor lesions of superficial spreading and nodular melanoma. *Hum Pathol* **15**: 1147–1165.

Conrad DF, Pinto D, Redon R, Feuk L, Gokcumen O, Zhang Y, Aerts J, Andrews TD, Barnes C, Campbell P, et al. 2010. Origins and functional impact of copy number variation in the human genome. *Nature* **464**: 704–712.

Damsky WE, Bosenberg M. 2017. Melanocytic nevi and melanoma: Unraveling a complex relationship. *Oncogene* **36**: 5771–5792.

Engel A, Johnson ML, Haynes SG. 1988. Health Effects of Sunlight Exposure in the United States: Results From the First National Health and Nutrition Examination Survey, 1971-1974. *Arch Dermatol* **124**: 72–79.

Foijer F, Xie SZ, Simon JE, Bakker PL, Conte N, Davis SH, Kregel E, Jonkers J, Bradley A, Sorger PK. 2014. Chromosome instability induced by Mps1 and p53 mutation generates aggressive lymphomas exhibiting aneuploidy-induced stress. *Proc Natl Acad Sci U S A* **111**: 13427–13432.

Gao R, Davis A, McDonald TO, Sei E, Shi X, Wang Y, Tsai PC, Casasent A, Waters J, Zhang H, et al. 2016. Punctuated copy number evolution and clonal stasis in triple-negative breast cancer. *Nat Genet* **48**: 1119–1130.

Gerstung M, Jolly C, Leshchiner I, Dentro SC, Gonzalez S, Rosebrock D, Mitchell TJ, Rubanova Y, Anur P, Yu K, et al. 2020. The evolutionary history of 2,658 cancers. *Nature* **578**: 122–128.

Gill D, Dorevitch A, Marks R. 2000. The prevalence of seborrheic keratoses in people aged 15 to 30 years: Is the term senile keratosis redundant? *Arch Dermatol* **136**: 759–762.

Gordon DJ, Resio B, Pellman D. 2012. Causes and consequences of aneuploidy in cancer. *Nat Rev Genet* **13**: 189–203.

Hartmann LC, Degnim AC, Santen RJ, Dupont WD, Ghosh K. 2015. Atypical hyperplasia of the breast-risk assessment and management options. *N Engl J Med* **372**: 78–89.

Hastings PJ, Ira G, Lupski JR. 2009. A Microhomology-Mediated Break-Induced Replication Model for the Origin of Human Copy Number Variation ed. I. Matic. *PLoS Genet* **5**: e1000327.

Henrichsen CN, Vinckenbosch N, Zöllner S, Chaignat E, Pradervand S, Schütz F, Ruedi M, Kaessmann H, Reymond A. 2009. Segmental copy number variation shapes tissue transcriptomes. *Nat Genet* **41**: 424–429.

Herd RM, Cooper EJ, Hunter JAA, Mclaren K, Chetty U, Watson ACH, Gollock J. 2006. Cutaneous malignant melanoma. Publicity, screening clinics and survival-the Edinburgh experience 1982-90. *Br J Dermatol* **132**: 563–570.

Huber V, Camisaschi C, Berzi A, Ferro S, Lugini L, Triulzi T, Tuccitto A, Tagliabue E, Castelli C, Rivoltini L. 2017. Cancer acidity: An ultimate frontier of tumor immune escape and a novel target of immunomodulation. *Semin Cancer Biol* **43**: 74–89.

Itsara A, Wu H, Smith JD, Nickerson DA, Romieu I, London SJ, Eichler EE. 2010. De novo rates and selection of large copy number variation. *Genome Res* **20**: 1469–1481.

Kaplan KB, Burds AA, Swedlow JR, Bekir SS, Sorger PK, Näthke IS. 2001. A role for the Adenomatous Polyposis Coli protein in chromosome segregation. *Nat Cell Biol* **3**: 429–432.

Karaayvaz-Yildirim M, Silberman RE, Langenbucher A, Saladi SV, Ross KN, Zarcaro E, Desmond A, Yildirim M, Vivekanandan V, Ravichandran H, et al. 2020. Aneuploidy and a deregulated DNA damage response suggest haploinsufficiency in breast tissues of BRCA2 mutation carriers. *Sci Adv* **6**: 1–11.

Killcoyne S, Gregson E, Wedge DC, Woodcock DJ, Eldridge MD, de la Rue R, Miremadi A, Abbas S, Blasko A, Kosmidou C, et al. 2020. Genomic copy number predicts esophageal cancer years before transformation. *Nat Med*.

Knouse KA, Davoli T, Elledge SJ, Amon A. 2017. Aneuploidy in Cancer: Seq-ing Answers to Old Questions. *Annu Rev Cancer Biol* **1**: 335–354.

Knouse KA, Wu J, Amon A. 2016. Assessment of megabase-scale somatic copy number variation using single-cell sequencing. *Genome Res* **26**: 376–384.

Knouse KA, Wu J, Whittaker CA, Amon A. 2014. Single cell sequencing reveals low levels of aneuploidy across mammalian tissues. *Proc Natl Acad Sci U S A* **111**: 13409–13414.

Kwong LN, Dove WF. 2009. APC and its modifiers in colon cancer. *Adv Exp Med Biol* **656**: 85–106.

Lamm N, Ben-David U, Golan-Lev T, Storchová Z, Benvenisty N, Kerem B. 2016. Genomic Instability in Human Pluripotent Stem Cells Arises from Replicative Stress and Chromosome Condensation Defects. *Cell Stem Cell* **18**: 253–261.

Levine MS, Bakker B, Boeckx B, Moyett J, Lu J, Vitre B, Spierings DC, Lansdorp PM, Cleveland DW, Lambrechts D, et al. 2017. Centrosome Amplification Is Sufficient to Promote Spontaneous Tumorigenesis in Mammals. *Dev Cell* **40**: 313-322.e5.

Lin WM, Luo S, Muzikansky A, Lobo AZC, Tanabe KK, Sober AJ, Cosimi AB, Tsao H, Duncan LM. 2015. Outcome of patients with de novo versus nevus-associated melanoma. In *Journal of the American Academy of Dermatology*, Vol. 72 of, pp. 54–58, Mosby Inc.

Martincorena I, Raine KM, Gerstung M, Dawson KJ, Haase K, Van Loo P, Davies H, Stratton MR, Campbell PJ. 2017. Universal Patterns of Selection in Cancer and Somatic Tissues. *Cell* **171**: 1029-1041.e21.

Martinez P, Mallo D, Paulson TG, Li X, Sanchez CA, Reid BJ, Graham TA, Kuhner MK, Maley CC. 2018. Evolution of Barrett's esophagus through space and time at single-crypt and whole-biopsy levels. *Nat Commun* **9**: 1–12.

Meijer GA, Hermsen MA, Baak JP, van Diest PJ, Meuwissen SG, Beliën JA, Hoovers JM, Joenje H, Snijders PJ, Walboomers JM. 1998. Progression from colorectal adenoma to carcinoma is associated with non-random chromosomal gains as detected by comparative genomic hybridisation. *J Clin Pathol* **51**: 901–9.

Mitchell TJ, Turajlic S, Rowan A, Nicol D, Farmery JHR, O'Brien T, Martincorena I, Tarpey P, Angelopoulos N, Yates LR, et al. 2018. Timing the Landmark Events in the Evolution of Clear Cell Renal Cell Cancer: TRACERx Renal. *Cell* **173**: 611-623.e17.

Muz B, de la Puente P, Azab F, Azab AK. 2015. The role of hypoxia in cancer progression, angiogenesis, metastasis, and resistance to therapy. *Hypoxia* **3**: 83.

Oromendia AB, Dodgson SE, Amon A. 2012. Aneuploidy causes proteotoxic stress in yeast. *Genes Dev* **26**: 2696–2708.

Page DL, Schuyler PA, Dupont WD, Jensen RA, Plummer WD, Simpson JF. 2003. Atypical lobular hyperplasia as a unilateral predictor of breast cancer risk: A retrospective cohort study. *Lancet* **361**: 125–129.

Pavelka N, Rancati G, Zhu J, Bradford WD, Saraf A, Florens L, Sanderson BW, Hattem GL, Li R. 2010. Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. *Nature* **468**: 321–325.

Pfau SJ, Silberman RE, Knouse KA, Amon A. 2016. Aneuploidy impairs hematopoietic stem cell fitness and is selected against in regenerating tissues in vivo. *Genes Dev* **30**: 1395–1408.

Phulari RGS, Buddhdev K, Rathore R, Patel S. 2014. Seborrheic keratosis. *J Oral Maxillofac Pathol* **18**: 327–330.

Piris A, Mihm MC, Hoang MP. 2015. BAP1 and BRAFV600E expression in benign and malignant melanocytic proliferations. *Hum Pathol* **46**: 239–245.

Poynter JN, Elder JT, Fullen DR, Nair RP, Soengas MS, Johnson TM, Redman B, Thomas NE, Gruber SB. 2006. BRAF and NRAS mutations in melanoma and melanocytic nevi. *Melanoma Res* **16**: 267–273.

Rajabi P, Adibi N, Nematollahi P, Heidarpour M, Eftekhari M, Siadat AH. 2012. Bowenoid transformation in seborrheic keratosis: A retrospective analysis of 429 patients. *J Res Med Sci* **17**: 217–21.

Reynolds LE, Watson AR, Baker M, Jones TA, D'Amico G, Robinson SD, Joffre C, Garrido-Urbani S, Rodriguez-Manzaneque JC, Martino-Echarri E, et al. 2010. Tumour angiogenesis is reduced in the Tc1 mouse model of Downs syndrome. *Nature* **465**: 813–817.

Ried T, Knutzen R, Steinbeck R, Blegen H, Schröck E, Heselmeyer K, du Manoir S, Auer G. 1996. Comparative genomic hybridization reveals a specific pattern of chromosomal gains and losses during the genesis of colorectal tumors. *Genes, Chromosom Cancer* **15**: 234–245.

Roh MR, Eliades P, Gupta S, Tsao H. 2015. Genetics of melanocytic nevi. *Pigment Cell Melanoma Res* **28**: 661–672.

Roper J, Tammela T, Cetinbas NM, Akkad A, Roghanian A, Rickelt S, Almeqdadi M, Wu K, Oberli MA, Sánchez-Rivera FJ, et al. 2017. In vivo genome editing and organoid transplantation models of colorectal cancer and metastasis. *Nat Biotechnol* **35**: 569–576.

Ross-Innes CS, Becq J, Warren A, Cheetham RK, Northen H, O'Donovan M, Malhotra S, di Pietro M, Ivakhno S, He M, et al. 2015. Whole-genome sequencing provides new insights into the clonal architecture of Barrett's esophagus and esophageal adenocarcinoma. *Nat Genet* **47**: 1038–1046.

Rutledge SD, Douglas TA, Nicholson JM, Vila-Casadesús M, Kantzler CL, Wangsa D, Barroso-Vilares M, Kale SD, Logarinho E, Cimini D. 2016. Selective advantage of trisomic human cells cultured in non-standard conditions. *Sci Rep* **6**: 1–12.

Santaguida S, Vasile E, White E, Amon A. 2015. Aneuploidy-induced cellular stresses limit autophagic degradation. *Genes Dev* **29**: 2010–2021.

Satgé D, Sommelet D, Geneix A, Nishi M, Malet P, Vekemans M. 1998. A tumor profile in Down syndrome. *Am J Med Genet* **78**: 207–216.

Sharp AJ, Locke DP, McGrath SD, Cheng Z, Bailey JA, Vallente RU, Pertz LM, Clark RA, Schwartz S, Segraves R, et al. 2005. Segmental duplications and copy-number variation in the human genome. *Am J Hum Genet* **77**: 78–88.

Sheltzer JM, Blank HM, Pfau SJ, Tange Y, George BM, Humpton TJ, Brito IL, Hiraoka Y, Niwa O, Amon A. 2011. Aneuploidy Drives Genomic Instability in Yeast. *Science (80-)* **333**: 1026–1030.

Sheltzer JM, Ko JH, Replogle JM, Habibe Burgos NC, Chung ES, Meehl CM, Sayles NM, Passerini V, Storchova Z, Amon A. 2017. Single-chromosome Gains Commonly Function as Tumor Suppressors. *Cancer Cell* **31**: 240–255.

Sheltzer JM, Torres EM, Dunham MJ, Amon A. 2012. Transcriptional consequences of aneuploidy. *Proc Natl Acad Sci* **109**: 12644–12649.

Sievers CK, Grady WM, Halberg RB, Pickhardt PJ. 2017. New insights into the earliest stages of colorectal tumorigenesis. *Expert Rev Gastroenterol Hepatol* **11**: 723–729.

Silk AD, Zasadil LM, Holland AJ, Vitre B, Cleveland DW, Weaver BA. 2013. Chromosome missegregation rate predicts whether aneuploidy will promote or suppress tumors. *Proc Natl Acad Sci* **110**: E4134–E4141.

Soto M, García-Santisteban I, Krenning L, Medema RH, Raaijmakers JA. 2018. Chromosomes trapped in micronuclei are liable to segregation errors. *J Cell Sci* **131**.

Tang Y-C, Williams BR, Siegel JJ, Amon A. 2011. Identification of Aneuploidy-Selective Antiproliferation Compounds. *Cell* **144**: 499–512.

Teixeira VH, Pipinikas CP, Pennycuick A, Lee-Six H, Chandrasekharan D, Beane J, Morris TJ, Karpathakis A, Feber A, Breeze CE, et al. 2019. Deciphering the genomic, epigenomic, and transcriptomic landscapes of pre-invasive lung cancer lesions. *Nat Med* **25**: 517–525.

Thomas R, Marks DH, Chin Y, Benezra R. 2018. Whole chromosome loss and associated breakage–fusion–bridge cycles transform mouse tetraploid cells. *EMBO J* **37**: 201–218.

Thorburn RR, Gonzalez C, Brar GA, Christen S, Carlile TM, Ingolia NT, Sauer U, Weissman JS, Amon A. 2013. Aneuploid yeast strains exhibit defects in cell growth and passage through START. *Mol Biol Cell* **24**: 1274–1289.

Torres EM, Dephoure N, Panneerselvam A, Tucker CM, Whittaker CA, Gygi SP, Dunham MJ, Amon A. 2010. Identification of Aneuploidy-Tolerating Mutations. *Cell* **143**: 71–83.

Torres EM, Sokolsky T, Tucker CM, Chan LY, Boselli M, Dunham MJ, Amon A. 2007. Effects of aneuploidy on cellular physiology and cell division in haploid yeast. *Science (80-)* **317**: 916–924.

Tsao H, Bevona C, Goggins W, Quinn T. 2003. The transformation rate of moles (melanocytic nevi) into cutaneous melanoma: A population-based estimate. *Arch Dermatol* **139**: 282–288.

Vogelstein B, Fearon ER, Hamilton SR, Kern SE, Preisinger AC, Leppert M, Smits AMM, Bos JL. 1988. Genetic Alterations during Colorectal-Tumor Development. *N Engl J Med* **319**: 525–532.

Vogelstein B, Fearon ER, Kern SE, Hamilton SR, Preisinger AC, Nakamura Y, White R. 1989. Allelotype of colorectal carcinomas. *Science* **244**: 207–11.

Wang Y, Waters J, Leung ML, Unruh A, Roh W, Shi X, Chen K, Scheet P, Vattathil S, Liang H, et al. 2014. Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature* **512**: 155–160.

Weaver BAA, Silk AD, Montagna C, Verdier-Pinard P, Cleveland DW. 2007. Aneuploidy Acts Both Oncogenically and as a Tumor Suppressor. *Cancer Cell* **11**: 25–36.

Weinberg F, Ramnath N, Nagrath D. 2019. Reactive oxygen species in the tumor microenvironment: An overview. *Cancers (Basel)* **11**.

Wijshake T, Malureanu LA, Baker DJ, Jeganathan KB, van de Sluis B, van Deursen JM. 2012. Reduced Life- and Healthspan in Mice Carrying a Mono-Allelic BubR1 MVA Mutation ed. P. Hasty. *PLoS Genet* **8**: e1003138.

Williams BR, Prabhu VR, Hunter KE, Glazier CM, Whittaker CA, Housman DE, Amon A. 2008. Aneuploidy Affects Proliferation and Spontaneous Immortalization in Mammalian Cells. *Science (80-)* **322**: 703–709.

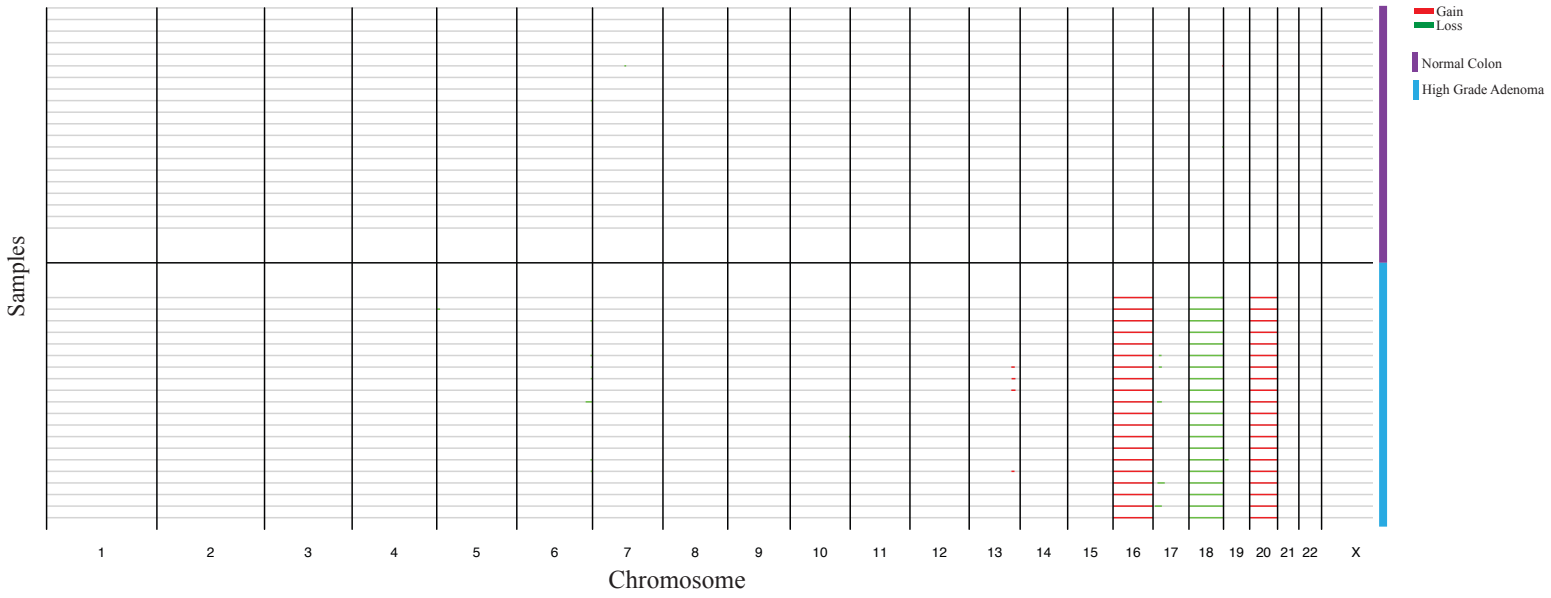
Yang A, Reeves RH. 2011. Increased survival following tumorigenesis in Ts65Dn mice that model down syndrome. *Cancer Res* **71**: 3573–3581.

Yeatman JM, Kilkenny M, Marks R. 1997. The prevalence of seborrhoeic keratoses in an Australian population: does exposure to sunlight play a part in their frequency? *Br J Dermatol* **137**: 411–414.

Yona AH, Manor YS, Herbst RH, Romano GH, Mitchell A, Kupiec M, Pilpel Y, Dahan O. 2012. Chromosomal duplication is a transient evolutionary solution to stress. *Proc Natl Acad Sci U S A* **109**: 21010–21015.

Zhang L, Shay JW. 2017. Multiple Roles of APC and its Therapeutic Implications in Colorectal Cancer. *J Natl Cancer Inst* **109**.

Supplemental Materials



Sup. Figure 1: Single-cell sequencing of high grade adenoma reveals highly clonal chromosomal aneuploidy and large CNVs

A. Single-cell karyotypes of high grade adenoma and healthy colon. Each line represents one cell. Cells from the same sample are bounded by black lines. Grey identifies regions with no copy number alteration, green identifies losses, and red identifies gains. Markers on the right y-axis indicate samples' origin: normal colon (purple) and high-grade adenoma (light blue)

Chapter Six: Conclusions and Future Directions

Conclusions and Future Directions

Aneuploidy decreases hematopoietic stem cells' fitness and is selected against in vivo

The relative fitness of aneuploid hematopoietic stem cells (HSCs) was evaluated using *in vivo* competition assays. Trisomy 19, trisomy 16, and BubR1$^{H/H}$ HSCs—which have depleted levels of the spindle assembly checkpoint protein BubR1, inducing chromosomal instability (CIN) (Baker et al. 2004)—were co-injected with an equal number of euploid competitor HSCs into lethally irradiated recipients. The aneuploid HSCs and euploid competitor HSCs bore different isoforms of the pan-leukocyte cell surface marker, CD45, allowing the progeny of these HSCs to be differentiated in the peripheral blood of recipients by the binding of isoform-specific antibodies.

Trisomy 16 HSCs were much less fit than euploid HSCs are were rapidly outcompeted. However, trisomy 19 HSCs and BubR1$^{H/H}$ HSCs showed minor or no fitness defects, respectively. We were surprised to observe that BubR1$^{H/H}$ HSCs showed no fitness defects in *in vivo* competition assays because single-cell sequencing (Knouse et al. 2014, 2016) revealed that ~39% of the peripheral blood generated by BubR1$^{H/H}$ HSCs 16 weeks after reconstitution was aneuploid.

Although they did not show severe fitness defects in *in vivo* competition assays, the reconstitution potential of trisomy 19 and BubR1$^{H/H}$ HSCs were reduced upon serial reconstitution. In serial reconstitutions, aneuploid HSCs or HSCs from euploid littermates were injected into lethally irradiated recipients and the HSCs' contribution to the recipients' peripheral blood was quantified using isoform-specific antibodies. Bone marrow from these primary recipients was later used to reconstitute lethally irradiated mice, called secondary recipients. This process was repeated to generate tertiary and, in the case of trisomy 19 HSCs, quaternary recipients. Trisomy 19 and BubR1$^{H/H}$ HSCs were effective during primary and secondary

reconstitutions. However, trisomy 19 HSCs contributed less of the peripheral blood of quaternary recipients than HSCs from euploid littermates. While BubR1$^{+/+}$ HSCs contributed an average of 75% of the peripheral blood of tertiary recipients at 16 weeks post-reconstitution, BubR1$^{H/H}$ HSCs failed to contribute to the peripheral blood of tertiary recipients at all. From these experiments, we concluded that aneuploidy reduces HSC fitness *in vivo*.

This conclusion is supported by previous observations of aneuploid organisms. Aneuploidy is the leading cause of miscarriage in humans (Jia et al. 2015) and all constitutive trisomies are embryonic lethal in mice (Williams et al. 2008). In addition, comparing the fitness of trisomy 19 and trisomy 16 HSCs revealed that the more genes included on the trisomic chromosome, the more severe the fitness defects observed in trisomic HSCs. This reflects the rates of aneuploidy observed in the human population. Only the trisomies of the three chromosomes which carry the fewest genes—chromosomes 13, 18, and 21—survive to birth in humans. Individuals trisomic for the chromosome with the fewest genes, chromosome 21, are the only humans with constitutive aneuploidy to survive to adulthood (Hassold and Jacobs 1984). This correlation between size of the aneuploid chromosome and the severity of the fitness penalty due to aneuploidy is inferred to exist in cancer cells. The likelihood that a chromosome arm will be gained or lost in tumors correlates with the number of genes on that chromosome arm, suggesting that cells that gain or lose larger chromosome arms face more severe fitness consequences and do not expand in the tumor population (Duijf et al. 2013).

Like aneuploid yeast (Torres et al. 2007) and trisomic mouse embryonic fibroblasts (Williams et al. 2008), trisomy 16 and trisomy 19 HSCs showed evidence for decreased proliferative potential. In contrast, BubR1$^{H/H}$ HSCs did not reveal proliferation defects. While we attributed the decreased serial reconstitution potential of trisomy 19 HSCs to proliferation

defects, we could not conclude that decreased proliferation potential caused BubR1$^{H/H}$ HSCs to lose their reconstitution potential during serial reconstitutions.

Rather, we hypothesized that the pool of BubR1$^{H/H}$ HSCs might become increasingly aneuploid over the course of serial reconstitutions. Perhaps the proliferative challenge of serially regenerating the hematopoietic compartment of multiple mice caused the HSC pool to become so enriched for aneuploid cells that the HSC pool could no longer contribute to the peripheral blood of recipients. To address this possibility, we single-cell sequenced the peripheral blood of a secondary recipient 57 weeks post-reconstitution. At the time of sampling, BubR1$^{H/H}$ HSCs contributed ~70% of the peripheral blood of this secondary recipient. We were surprised to observe that all 17 BubR1$^{H/H}$ derived peripheral blood cells were euploid. This was unexpected because almost 40% of the peripheral blood cells generated by BubR1$^{H/H}$ HSCs 16 weeks into a primary reconstitution were aneuploid.

We hypothesized that these conflicting results were the result of the dynamics of HSC expansion during reconstitution. We observed a high prevalence of aneuploid cells early in reconstitution, during a time of rapid expansion, but we observed no aneuploid cells later in reconstitution, when the hematopoietic system has returned to a steady state, similar to normal hematopoiesis. Following the evolutionary principal of the "population flush" which dictates that rapidly expanding populations experience relaxed purifying selection, we hypothesized that immediately after reconstitution rapid HSC expansion may relax purifying selection, allowing aneuploidy to be tolerated and aneuploid cells to accumulate. Then, after HSC expansion, as proliferation slows and purifying selection resumes, these aneuploid cells would be eliminated.

To test this hypothesis, we used single-cell sequencing to measure the fraction of aneuploid cells present in the peripheral blood of lethally irradiated recipients of BubR1$^{H/H}$ HSCs

at 3 weeks, 6 week, 12 weeks, and 36 weeks post-reconstitution. We observed that the presence of aneuploid cells increased immediately following reconstitution, then decreased. We found that ~8% of peripheral blood cells were aneuploid at 3 weeks post-reconstitution, ~24% were aneuploid at 6 weeks post-reconstitution, ~22% were aneuploid at 12 weeks post-reconstitution, and ~12% were aneuploid 36 weeks post-reconstitution. Our findings indicated that a higher level of aneuploidy is tolerated in the peripheral blood of BubR1$^{H/H}$ HSC recipients during the period of rapid expansion that follows reconstitution. Then aneuploid cells were depleted from the peripheral blood once rapid proliferation ceases and the hematopoietic compartment returns to a steady state.

The observation that aneuploid cells persist in a rapidly expanding population but are eliminated from populations approaching a proliferative steady state predicts that adult BubR1$^{H/H}$ tissues should contain different levels of aneuploidy depending on the proliferative state of each tissue. Tissues that form through rapid proliferation during embryogenesis but are largely non-proliferative in adults should harbor higher levels of aneuploidy than tissues that regenerate in adults, and thus experience continuous purifying selection.

To test this prediction, we single-cell sequenced the tissues of adult BubR1$^{H/H}$ animals. We sequenced peripheral blood, skin, and intestine—which are all regenerative in adults—and liver and brain, which are largely non-proliferative in adults (Zimmermann 2004; Campisi and D'Adda Di Fagagna 2007). We observed that proliferative tissues contained fewer aneuploid cells than non-proliferative tissues. We concluded that aneuploidy is selected against in tissues that regenerate in adults.

To determine if proliferation status predicts the prevalence of aneuploid cells in human tissues, we have compiled single-cell sequencing data from histologically normal tissues of adult

humans. While the cells of regenerative tissues—blood, skin, intestine, and breast—lack aneuploid cells, non-proliferative tissues—brain and liver—contain aneuploid cells (Fig. 1). It is important to note that 3 of 4 aneuploid hepatocytes were tetraploid, which could lessen the fitness penalty due to aneuploidy (Storchova and Kuffer 2008) and may account for some of the liver's increased tolerance of aneuploidy. In the future, it may prove informative to single-cell sequence additional non-proliferative human tissues to determine if the proliferative status of adult tissues predicts their tolerance of aneuploid cells.

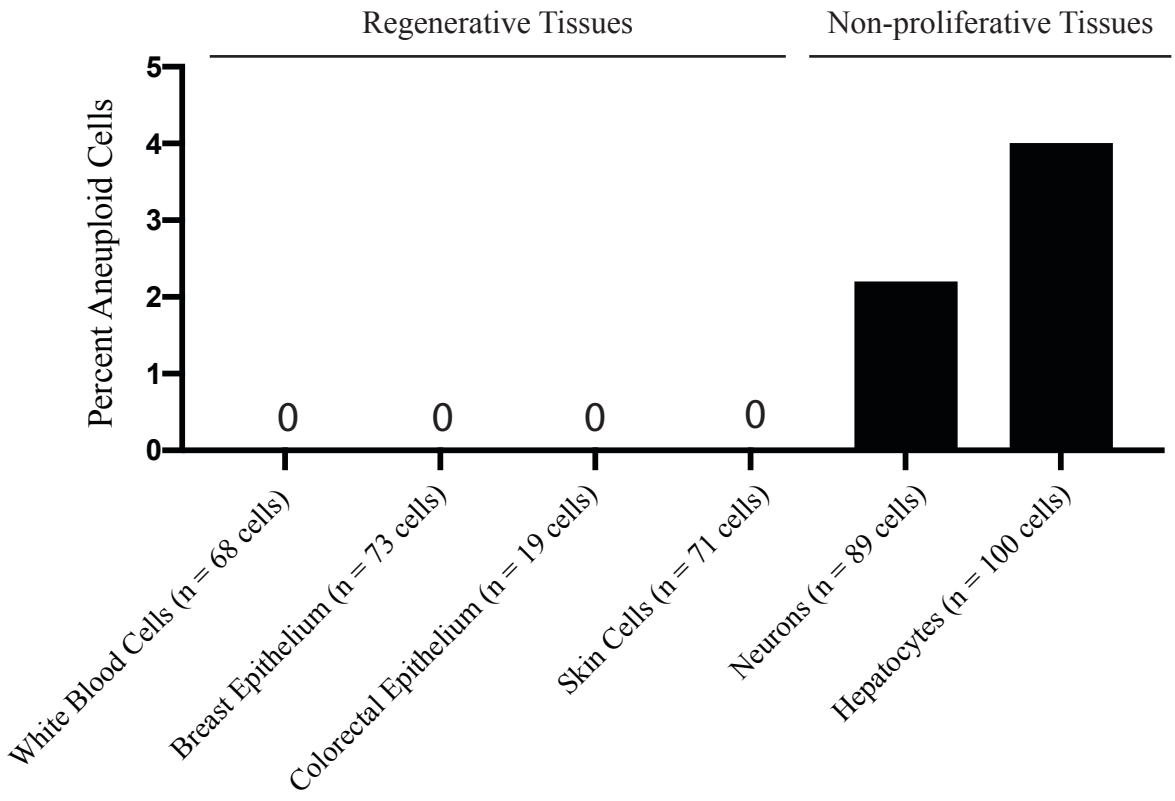


Figure 1: Healthy human cells from regenerative and non-proliferative tissues contain different levels of aneuploidy. Graph showing the percent of aneuploid cells. Tissues where no aneuploid cells were identified are marked with a "0." Data for white blood cells and breast epithelial cells come from experiments described in Chapter 4; data for intestinal epithelial cells and 18 of 71 skin cells come from experiments in Chapter 5; data for 53 of 71 skin cells, all neurons, and all hepatocytes come from Knouse et al. (2014).

Future studies could also explore the mechanism by which aneuploid cells are eliminated from regenerating tissues. Aneuploidy causes DNA damage (Burrell et al. 2013; Lamm et al. 2018; Sheltzer et al. 2011; Janssen et al. 2011) and p53 activation (Thompson and Compton 2008). DNA damage and p53 activation can trigger apoptosis (Vousden and Lane 2007). While it is known that different tissues regularly encounter different types of DNA damage due to their physical environment or function and thus have biases in their usage of the DNA damage repair machinery (Sun et al. 2019), it is not known how tissue-specific responses to DNA damage or p53 activation affect the prevalence of aneuploid cells in tissues. Experiments evaluating the role of p53 and other components of the DNA damage response machinery in the clearance of aneuploid cells in a tissue-specific manner may offer insight into the observation that aneuploid cells are found at different frequencies in different tissues.

A second possibility for how aneuploid cells are cleared from proliferative tissues is through immune intervention. Highly aneuploid cells are targeted and destroyed by natural killer cells *in vitro* (Wang et al. 2020; Santaguida et al. 2017) and micronuclei—a hallmark of aneuploid cells—activate the cell intrinsic immune surveillance pathway via activation of the cGAS-STING pathway (Harding et al. 2017; MacKenzie et al. 2017). However, the relationship between aneuploid cells and the immune system remains controversial, with some arguing that aneuploidy enhances immune evasion in tumors (Davoli et al. 2017). Whether the immune system plays a role in the clearance of immune cells and how this relationship may change during tumorigenesis are topics which require investigation.

The strength of purifying selection determines the prevalence of aneuploid cells following hematopoietic reconstitution

We observed that aneuploid cells accumulate in the peripheral blood of mice for a short period following reconstitution with chromosomally unstable HSCs but are eliminated later on. We used computational modeling to determine if altered purifying selection can account for this peak in the prevalence of aneuploid cells post-reconstitution. To recreate the process of hematopoietic reconstitution virtually, we simplified the hematopoietic system to two components: the HSC niche and HSCs. The HSC niche is composed of both hematopoietic and mesenchymal cell types which regulate the renewal, differentiation, and proliferation of HSCs (Morrison and Scadden 2014). *In vivo*, the immediate progeny of HSCs, transit amplifying cells, perform the bulk of the divisions that give rise to all other hematopoietic cells but have limited self-renewal (Boulais and Frenette 2015). For our simulated reconstitutions, we grouped HSCs and immediate HSC progeny together to form a single population of cells—which we called HSCs—which gives rise to all other hematopoietic cells. Because the karyotypes of the cells of this progenitor group affect the karyotypes of all progeny, our model is only concerned with the rates of aneuploidy in this population.

Our model consisted of 10,000 virtual stem cell niches that could each be filled with 1 HSC, allowing 10,000 total HSCs to populate the system, which is similar to the total number of HSCs in mice (Abkowitz et al. 2002). Simulations began with just 100 HSCs occupying the 10,000 available niches. This is similar to the number of chromosomally unstable HSCs injected in our reconstitutions. Initially, simulated HSCs expanded rapidly (once per 3 days). Once the population of HSCs expanded to fill all 10,000 available niches, division slowed to once per 40 days, which is similar to the rate of steady state divisions *in vivo* (Catlin et al. 2011). Our model

allowed us to alter additional variables like the rate of production of aneuploid cells and the degree of fitness defect due to aneuploidy.

In our simulated reconstitutions, increasing the fitness penalty due to aneuploidy 100 fold (from -0.1% to -10%) only decreased the number of aneuploid cells that accumulated post-reconstitution by ~40%. This suggested that purifying selection is relaxed immediately after reconstitution because increasing the fitness penalty due to aneuploidy did not result in a proportionate elimination of aneuploid cells from the population. When the fitness penalty due to aneuploidy was -0.1%, aneuploid cells were never purged from the population, continuing to persist after rapid HSC division ceased and the hematopoietic system returned to a steady state. When the fitness penalty due to aneuploidy was -10%, aneuploid cells were rapidly and completely purged once the hematopoietic system returned to a steady state. This indicated that purifying selection was involved in the clearance of aneuploid cells after the cessation of rapid HSC expansion.

We then sought to determine which factors affect the strength of purifying selection in this model. Increasing the rate of HSC division in a population with a fixed size—so in the absence of population expansion—did not produce an aneuploid peak. This suggested that population expansion, not just rapid HSC division, drives the relaxed purifying selection which permits the accumulation of aneuploid cells.

We also explored the possibility that damage to the bone marrow niche can contribute to the accumulation of aneuploid cells post-reconstitution. During reconstitution, recipient animals are lethally irradiated in order to ablate their HSCs, damaging the HSC niche (Cao et al. 2011). This damage may reduce the niche's ability to support stemness, the capacity for self-renewal, in HSCs. We added a variable that allowed us to set the probability that an HSC would be deleted

from the HSC pool, regardless of that HSC's fitness, creating events where a niche did not maintain HSC stemness. We observed that when HSCs expanded in the context of a damaged niche, the height and duration of the aneuploid peak was greater than that generated by HSCs expanding in the context of a healthy niche. We concluded that both rapid HSC expansion and decreased niche health could create conditions where purifying selection is relaxed post-reconstitution, allowing aneuploid cells to accumulate.

It would be interesting to test some of the predictions of our computational modeling *in vivo*. The results of our simulation suggested that injecting a larger number of HSCs at the start of the reconstitution, decreasing the total number of divisions required to return the hematopoietic system to steady state, would decrease the duration of the aneuploid peak post-reconstitution. We also modeled the impact that the rate of the generation of aneuploid cells had on the height and duration of the aneuploid peak. Our modeling suggested that increasing the rate with which aneuploid cells are produced by reconstituting mice with HSCs from mouse models with higher levels of CIN (Silk et al. 2013) would increase the height and duration of the aneuploid peak.

This research has implications for tumorigenesis. Bone marrow transplantation is associated with an increased risk of post-transplantation leukemia, though this may be explained at least in part by radiation-induced mutations and suppression of the immune system (Forrest et al. 2003). It is possible that purifying selection is also relaxed post-transplantation allowing less fit pre-leukemic cells to persist and undergo malignant transformation. Outside the context of bone marrow transplantation, the persistence of pre-leukemic cells under conditions of relaxed purifying selection have been predicted to explain the high rates of leukemia in children. We believe that cancer develops through the sequential acquisition of multiple pro-tumorigenic

genetic changes (Balmain et al. 2003; Vogelstein and Kinzler 2004). So it is surprising that leukemia is common in young children, whose cells have undergone fewer divisions than adults and likely accumulated fewer mutations. Stochastic modeling has indicated that during childhood the rapid expansion of HSCs and small HSC pool produce conditions where genetic drift has a greater impact on population dynamics than purifying selection. This genetic drift can allow pre-leukemic cells to persist, creating the opportunity for pre-leukemic cells to acquire additional mutations, possibly driving leukemogenesis (Rozhok et al. 2016). Although leukemogenesis may be enhanced post-transplantation and in children because of relaxed purifying selection, the conditions of post-transplantation hematopoiesis and hematopoietic expansion during normal development differ in many ways. For example, the stem cell niche is damage in bone marrow transplant (Cao et al. 2011) but likely remains intact during development.

To determine if the proposed relaxation of purifying selection during hematopoietic expansion in children has the same outcome as the relaxation of purifying selection post-reconstitution, it may be informative to single-cell sequence the bone marrow and peripheral blood of developing, chromosomally unstable mice. If less fit aneuploid cells persist in the hematopoietic system of developing CIN mice, but are cleared in CIN adults, this would suggest that purifying selection is relaxed during hematopoietic expansion during development. This would indicate that while niche damage contributed to the relaxation of purifying selection in our virtual reconstitutions, niche damage is not necessary for the relaxation of purifying selection during development. This would offer us insight into the mechanisms that maintain purifying selection in tissues and how these mechanisms are subverted in tumorigenesis.

Cells from BRCA2 mutation carriers have sub-chromosomal copy number variants and a deregulated DNA damage response

Women who carry heterozygous mutations in the gene *BRAC2* face an increased risk of developing breast cancer (Petrucelli et al. 1993). To explore the role of genome imbalance in the early pathogenesis of breast cancer in *BRCA2* mutated breast tissue, we performed single-cell sequencing on cells from histologically normal breast tissue from *BRCA2* carriers and non-carriers. We observed that a significant fraction of *BRCA2* carrier breast cells exhibit large (>10 megabase (Mb)) sub-chromosomal copy number variations (CNVs), which were rare in non-carriers. CNVs are a form of genome imbalance that are distinct from whole chromosomal aneuploidy. While CNVs arise through failures in any of several genome integrity maintenance pathways (Conrad et al. 2010; Sharp et al. 2005; Hastings et al. 2009; Soto et al. 2018), they do not result from chromosome mis-segregation, which produces whole chromosomal aneuploidy (Gordon et al. 2012). CNVs can range in size from a few kilobases to over 100 Mb, though CNVs greater than 10 Mb are rare in healthy tissues (Knouse et al. 2016).

We confirmed that loss of heterozygosity at the *BRCA2* locus had not occurred in these *BRCA2* carrier breast cells with large CNVs, suggesting that *BRCA2* may be haploinsufficient in breast tissue. In support of this, *BRCA2* carrier breast cells exhibited DNA damage and an attenuated replication checkpoint. We concluded that *BRCA2* is haploinsufficient in breast tissue, and that the accumulation of cells with large CNVs precedes tumorigenesis in this context.
In light of these findings and in order to evaluate if *BRCA2* haploinsufficiency is tissue-specific, we have single-cell sequenced white blood cells from *BRCA2* carriers and non-carriers. If *BRCA2* haploinsufficiency is specific to breast cells, this would have significant implications for our understanding why *BRCA2* mutation carriers face a tissue-specific increase in cancer risk. We discovered that the blood of *BRCA2* carriers also contains large CNVs (Fig. 2A). When we

compared the distribution of the size of CNVs in *BRCA2* carrier breast cells and non-carrier breast cells, we observed a highly significant difference (Mann-Whitney U Test, $p < 0.001$), but when we compared the distribution of the size of CNVs in *BRCA2* carrier breast cells to *BRCA2* carrier blood cells there was no was no significant difference (Mann-Whitney U Test, not significant (n.s.)). These distributions are not normalized by the number of CNVs per cell in each tissue, so do not describe the rate of CNVs in tissues. Rather, they indicate that the size of CNVs present in the breast and blood cells of *BRCA2* carriers are similar.

In order to quantitatively compare the degree of imbalance caused by CNVs in these tissues, we calculated the average genome imbalance (AGI) of each tissue. First, we quantified the total genomic distance in base pairs (bp) which was affected by CNVs in each cell. Then, we averaged these total imbalance values for each tissue. We were surprised to observe that *BRCA2* carrier blood cells have a significantly higher AGI than *BRCA2* carrier breast cells (Fig. 2B; Mann-Whitney U Test, $p < 0.05$). However, we found no significant differences between the AGI of *BRCA2* carrier blood cells and non-carrier blood cells (Mann-Whitney U Test; n.s.). We cannot conclude that *BRCA2* mutation causes a significant increase in genome imbalance in blood cells. However, we note that the lack of significant difference between these groups may be a consequence of the smaller number of blood cells in our dataset. We also observed that non-carrier blood cells have significantly higher AGI than non-carrier breast cells (Mann-Whitney U Test; $p < 0.01$). From these data, we conclude that normal blood cells have a higher degree of genome imbalance than normal breast cells. We speculate that breast tissue may be exquisitely sensitive to perturbations in genome imbalance.

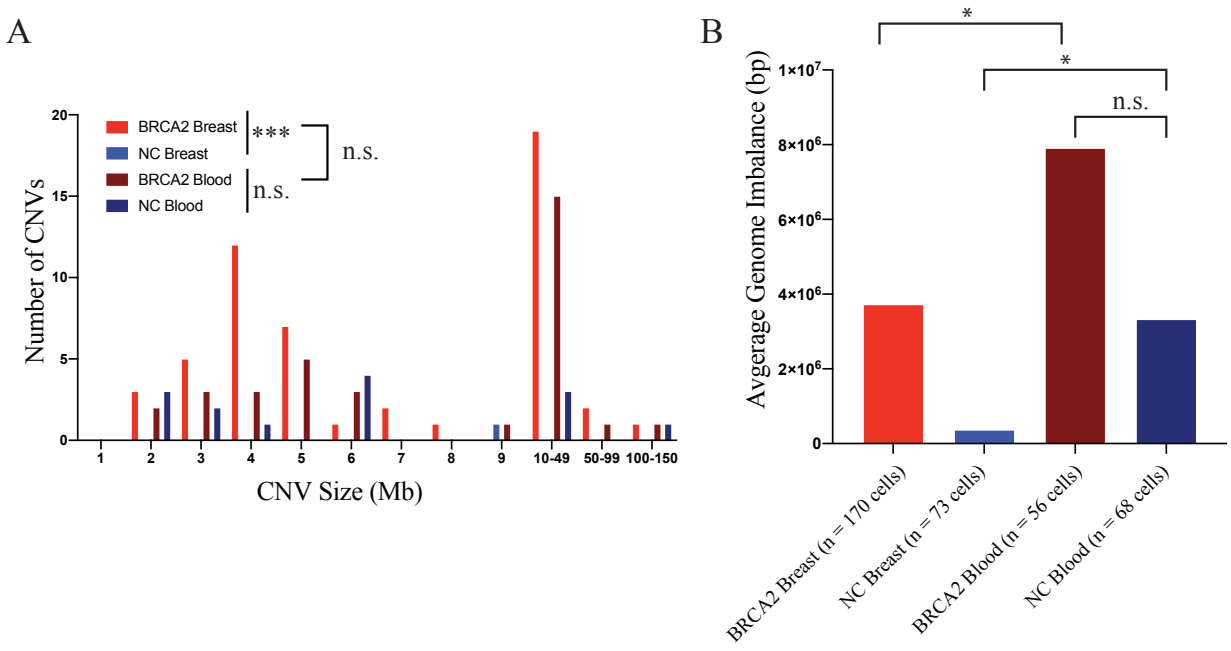


Figure 2: *BRCA2* mutated blood cells have large CNVs.
A. Distribution of the size of CNVs in non-carrier (NC) and *BRCA2* carrier blood and breast cells. Statistical comparisons by Mann-Whitney U Test: *** $p < 0.001$; not significant (n.s.). B. Average genome imbalance (AGI) in base pairs (bp) for non-carrier (NC) and *BRCA2* carrier blood and breast cells. Statistical comparisons performed by Mann-Whitney U Test: * $p < 0.05$; not significant (n.s.). *BRCA2* Breast (n = 7 individuals), NC Breast (n = 4 individuals), *BRCA2* Blood (n = 3 individuals), NC Blood (n = 4 individuals).

We propose that *BRCA2* haploinsufficiency may be specific to breast cells, though further research is needed to confirm that *BRCA2* carrier white blood cells maintain full *BRCA2* function. It would also be prudent to single-cell sequence additional *BRCA2* carrier and non-carrier white blood cells to confirm that the lack of significant difference in AGI between these two groups is not a consequence of smaller sample size.

One possible explanation for why loss of one functional copy of *BRCA2* might have a greater effect on breast cells than blood cells is the role of estrogen signaling in *BRCA2* function. Estrogen signaling leads to the phosphorylation and stabilization of the BRCA2 protein (Malone et al. 2009). Based on our observation that non-carrier breast cells have lower levels of genome imbalance than non-carrier blood cells, we speculate that estrogen signaling may stabilize the

BRCA2 protein in breast cells, enhancing the DNA damage repair capacity of this tissue and limiting the degree of genome imbalance in breast cells. Tissues without sensitivity to estrogen signaling lack this stabilization of BRCA2, possibly resulting in a higher baseline for genome imbalance. Perhaps mutating one copy of *BRCA2* depletes the pool of functional BRCA2 protein below the level required to maintain the typically high standard of genome integrity in breast cells but has a lesser impact on tissues where the BRCA2 pool not maintained by estrogen signaling. Elucidating the relationship between estrogen signaling and *BRCA2* haploinsufficiency would have important implications for our understanding why of *BRCA2* mutation drives tumorigenesis in a tissue-specific manner.

While large CNVs arise in early tumorigenesis, whole chromosomal aneuploidy arises in late tumorigenesis

In order to gain insight into the role of genome imbalance during tumorigenesis, we investigated the timing with which whole chromosomal aneuploidy and sub-chromosomal CNVs arise during tumorigenesis. We performed e-karyotyping, a computational method which translates bulk gene expression data into DNA copy numbers, to define the karyotypes of samples of normal colons, early colorectal adenomas, late colorectal adenomas, and colorectal carcinomas. We determined that clonal whole chromosomal aneuploidy is restricted to late pre-malignant colorectal adenomas and colorectal carcinomas. These data are in agreement with previous analyses of lung cancer (Teixeira et al. 2019), breast cancer (Casasent et al. 2017; Gao et al. 2016), and esophageal carcinoma (Ross-Innes et al. 2015; Martinez et al. 2018; Killcoyne et al. 2020), which all concluded that aneuploidy arises in late tumorigenesis.

While clonal CNVs greater than 10 Mb were absent from samples of normal colon, they were present in every stage of colorectal carcinogenesis including early adenomas, which lacked

clonal whole chromosomal aneuploidy. We concluded that clonal sub-chromosomal CNVs arise before clonal whole chromosomal aneuploidy in the development of colorectal carcinomas.

The quantity of genome imbalance generated by whole chromosomal aneuploidy and the diversity of chromosomes affected by whole chromosomal aneuploidy underwent a modest increase during the transition from late pre-malignant adenomas to colorectal carcinomas. Conversely, the quantity of genome imbalance attributed to CNVs and the distribution of CNVs across the genome increased dramatically in the transition from late pre-malignancy to invasive disease. While the observation that the vast majority of mutations are tolerated in tumors has led others to propose that negative selection has only a minor role in shaping cancer genomes (Martincorena et al. 2017), we propose that whole chromosomal aneuploidy faces significant negative selection, limiting the quantity and diversity of whole chromosomal copy number alterations in colorectal carcinomas. CNVs appear to be less limited by negative selection and likely confer a lesser fitness penalty. This is compatible with our observation that smaller aneuploid chromosomes confer less severe fitness penalties, as CNVs typically affect smaller numbers of genes than gaining or losing whole chromosomes. In addition to limiting the spectrum of aneuploidy seen in late stages of tumorigenesis, we propose that negative selection contributes to the lack of clonal aneuploidy we observed in early pre-malignant lesions. Given our observation that aneuploid cells are selected against in regenerating tissues, it follows that aneuploid pre-cancerous cells may be cleared from early pre-malignant growths by purifying selection.

In order to determine if sub-clonal, large CNVs or sub-clonal whole chromosomal aneuploidy are features of early pre-malignant lesions outside the context of intestinal adenomas, we used single-cell sequencing to karyotype cells from pre-malignancies of the skin and breast.

We analyzed 5 samples of seborrheic keratosis, a long-lived skin tumor which can give rise to basal cell carcinoma in rare cases (Bedir et al. 2014), and found that the size of CNVs in seborrheic keratoses is significantly larger than the size of CNVs found in healthy skin cells. We analyzed 3 samples of breast atypical hyperplasia, an early precursor of invasive breast cancer (Hartmann et al. 2015). We found the size of CNVs identified in the epithelial cells—which include hyperplasia cells (Bombonati and Sgroi 2011)—were significantly larger than those in non-epithelial cells.

We analyzed 5 samples of melanotic nevi, an early precursor of cutaneous melanoma (Clark et al. 1984). We did not find a significant difference in the size CNVs between melanotic nevus cells and healthy skin cells. One possible explanation for why melanotic nevi do not contain as many large CNVs while other pre-malignancies do is that melanotic nevi are growth arrested. While the mechanisms behind this arrest continue to be studied, it is accepted that the vast majority of nevi cease proliferating after a brief period of expansion (Damsky and Bosenberg 2017). This observation proposes that melanotic nevi may not experience a prolonged enough period of active hyperproliferation to allow for the accumulation of cells with large CNVs.

We found no cells containing whole chromosomal aneuploidy in any samples of early pre-malignant lesions, affirming that whole chromosomal aneuploidy arises in late tumorigenesis. This is in support of our proposal that whole chromosomal aneuploidy is selected against in early pre-malignant growths. We conclude that sub-chromosomal CNVs arise in early tumorigenesis, while whole chromosomal aneuploidy arises in late tumorigenesis. We have previously described that large CNVs are a characteristic of histologically normal breast tissue in women with *BRCA2* mutations and that large CNVs are present in the blood of both *BRCA2*

carriers and non-carriers. Large CNVs have also been found in Barrett's esophagus, an early precursor of esophageal carcinoma (Ross-Innes et al. 2015; Martinez et al. 2018; Killcoyne et al. 2020). The presence of large CNVs in the blood of *BRCA2* carriers indicates that large CNVs may not directly promote tumorigenesis, since *BRCA2* carriers do not have a significant risk of developing leukemia (Iqbal et al. 2016). Similarly, many cases of Barrett's esophagus which contain large CNVs never evolve to esophageal carcinoma. Instead, transformation from Barrett's esophagus to esophageal carcinoma is preceded by the emergence of whole chromosomal aneuploidy, which occurs shortly before malignant transformation (Killcoyne et al. 2020; Martinez et al. 2018; Ross-Innes et al. 2015). We conclude that large CNVs may not independently drive tumorigenesis, even though they are common in pre-malignant growths.

The observation that aneuploidy arises in late tumorigenesis begs the questions of what triggers aneuploid cells to accumulate. Loss of tissue architecture during tumorigenesis may impair chromosome segregation fidelity (Knouse et al. 2018) increasing the rate of chromosome mis-segregation and increasing the opportunity for evolution to a pro-tumorigenic karyotype. Altered interactions with the immune system may permit the persistence of aneuploid cells in late pre-malignant lesions (Davoli et al. 2017). Alternatively, environmental stress in late tumorigenesis may create conditions which permit the accumulation of aneuploid cells. Aneuploid cells thrive under stress conditions (Rutledge et al. 2016; Pavelka et al. 2010; Yona et al. 2012) and aneuploidy can facilitate adaptation to acute stress (Chen et al. 2015). We speculate that during late tumorigenesis, stress conditions such as hypoxia (Muz et al. 2015), acidification (Huber et al. 2017), or the buildup of reactive oxygen species (Weinberg et al. 2019) may allow aneuploid cells to accumulate.

In order to investigate this possibility, it may prove fruitful to study the effects of environmental conditions on the karyotypes of early adenoma cells using an intestinal organoid system. As part of our work exploring karyotype evolution in colorectal carcinomas, we have already demonstrated that simultaneous inactivation of both copies of *APC* in the $Apc^{fl/fl};Villin^{CreER}$ mouse model (Roper et al. 2017) generates early adenomas which lack whole chromosomal aneuploidy. We determined that inducing adenoma formation by inactivating one copy of *APC*, the waiting for loss of heterozygosity to occur resulted an early adenoma that lacked whole chromosomal aneuploidy by featured large CNVs. Given that *APC* mutations are an early or initiating event in ~85% of spontaneous colorectal carcinomas (Kwong and Dove 2009) and that we found that human early adenomas lack whole chromosomal aneuploidy but can feature large, clonal CNVs, we believe this model provides a faithful representation of karyotype evolution in early tumorigenesis in human colorectal adenomas.

Cells from the $Apc^{fl/fl};Villin^{CreER}$ mouse model readily grow in organoids (Roper et al. 2017) and organoids maintain genome stability in culture (Huch et al. 2015). The organoid setting allows for many different manipulations of cells' environments like varying nutrient conditions, altering oxygen content, co-culturing with immune cells, or drug treatment. This system provides the opportunity to determine which conditions, if any, trigger the accumulation of aneuploid cells. By studying the behavior of any evolved aneuploid organoids, we could investigate how the emergence of aneuploidy may contribute to tumorigenesis.

Summary

We have studied the dynamics of aneuploid cells in tissues from points all along the spectrum of tumorigenesis—from normal tissues to pre-malignant lesions to malignant tumors—in order to gain insight into why aneuploid cells accumulate during tumorigenesis. The work

presented in this thesis reveals that aneuploid cells experience fitness defects *in vivo*, and are selected against in normal, regenerating tissues. We also describe that purifying selection can drive the clearance of aneuploid cells from regenerating tissues. By studying the timing with which two distinct forms of genome imbalance—whole chromosomal aneuploidy and sub-chromosomal CNVs—arise during tumorigenesis, we have determined that sub-chromosomal CNVs arise in early tumorigenesis, while whole chromosomal aneuploidy arises in late tumorigenesis. We speculate that—unlike CNVs—whole chromosomal aneuploidy may decrease cells' fitness and face negative selection. This suggests that the same purifying selection which purges aneuploid cells from normal, regenerating tissues, may be active during early tumorigenesis. Our work suggests that whole chromosomal aneuploidy has a limited role in early tumorigenesis and motivates future research into the role of aneuploidy in late tumorigenesis.

References

Abkowitz JL, Catlin SN, McCallie MT, Guttorp P. 2002. Evidence that the number of hematopoietic stem cells per animal is conserved in mammals. *Blood* **100**: 2665–2667.

Baker DJ, Jeganathan KB, Cameron JD, Thompson M, Juneja S, Kopecka A, Kumar R, Jenkins RB, de Groen PC, Roche P, et al. 2004. BubR1 insufficiency causes early onset of aging-associated phenotypes and infertility in mice. *Nat Genet* **36**: 744–749.

Balmain A, Gray J, Ponder B. 2003. The genetics and genomics of cancer. *Nat Genet* **33**: 238–244.

Bedir R, Yurdakul C, Güçer H, Ehitog IS¸, Lu ˘. 2014. Basal Cell Carcinoma Arising within Seborrheic Keratosis Dermatology Section. *J Clin Diagnostic Res* **8**: 6–07.

Bombonati A, Sgroi DC. 2011. The molecular pathology of breast cancer progression. *J Pathol* **223**: 308–318.

Boulais PE, Frenette PS. 2015. Making sense of hematopoietic stem cell niches. *Blood* **125**: 2621–2629.

Burrell RA, McGranahan N, Bartek J, Swanton C. 2013. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature* **501**: 338–345.

Campisi J, D'Adda Di Fagagna F. 2007. Cellular senescence: When bad things happen to good cells. *Nat Rev Mol Cell Biol* **8**: 729–740.

Cao X, Wu X, Frassica D, Yu B, Pang L, Xian L, Wan M, Lei W, Armour M, Tryggestad E, et al. 2011. Irradiation induces bone injury by damaging bone marrow microenvironment for stem cells. *Proc Natl Acad Sci U S A* **108**: 1609–1614.

Casasent AK, Edgerton M, Navin NE. 2017. Genome evolution in ductal carcinoma in situ: invasion of the clones. *J Pathol* **241**: 208–218.

Catlin SN, Busque L, Gale RE, Guttorp P, Abkowitz JL. 2011. The replication rate of human hematopoietic stem cells in vivo. *Blood* **117**: 4460–4466.

Chen G, Mulla WA, Kucharavy A, Tsai HJ, Rubinstein B, Conkright J, McCroskey S, Bradford WD, Weems L, Haug JS, et al. 2015. Targeting the adaptability of heterogeneous aneuploids. *Cell* **160**: 771–784.

Clark WH, Elder DE, Guerry D, Epstein MN, Greene MH, Van Horn M. 1984. A study of tumor progression: The precursor lesions of superficial spreading and nodular melanoma. *Hum Pathol* **15**: 1147–1165.

Conrad DF, Pinto D, Redon R, Feuk L, Gokcumen O, Zhang Y, Aerts J, Andrews TD, Barnes C, Campbell P, et al. 2010. Origins and functional impact of copy number variation in the human genome. *Nature* **464**: 704–712.

Damsky WE, Bosenberg M. 2017. Melanocytic nevi and melanoma: Unraveling a complex relationship. *Oncogene* **36**: 5771–5792.

Davoli T, Uno H, Wooten EC, Elledge SJ. 2017. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science (80-)* **355**.

Duijf PHG, Schultz N, Benezra R. 2013. Cancer cells preferentially lose small chromosomes. *Int J Cancer* **132**: 2316–2326.

Forrest DL, Nevill TJ, Naiman SC, Le A, Brockington DA, Barnett MJ, Lavoie JC, Nantel SH, Song KW, Shepherd JD, et al. 2003. Second malignancy following high-dose therapy and autologous stem cell transplantation: Incidence and risk factor analysis. *Bone Marrow Transplant* **32**: 915–923.

Gao R, Davis A, McDonald TO, Sei E, Shi X, Wang Y, Tsai PC, Casasent A, Waters J, Zhang H, et al. 2016. Punctuated copy number evolution and clonal stasis in triple-negative breast cancer. *Nat Genet* **48**: 1119–1130.

Harding SM, Benci JL, Irianto J, Discher DE, Minn AJ, Greenberg RA. 2017. Mitotic progression following DNA damage enables pattern recognition within micronuclei. *Nature* **548**: 466–470.

Hartmann LC, Degnim AC, Santen RJ, Dupont WD, Ghosh K. 2015. Atypical hyperplasia of the breast-risk assessment and management options. *N Engl J Med* **372**: 78–89.

Hassold TJ, Jacobs PA. 1984. Trisomy in man. *Annu Rev Genet* **18**: 69–97.

Hastings PJ, Ira G, Lupski JR. 2009. A Microhomology-Mediated Break-Induced Replication Model for the Origin of Human Copy Number Variation ed. I. Matic. *PLoS Genet* **5**: e1000327.

Huber V, Camisaschi C, Berzi A, Ferro S, Lugini L, Triulzi T, Tuccitto A, Tagliabue E, Castelli C, Rivoltini L. 2017. Cancer acidity: An ultimate frontier of tumor immune escape and a novel target of immunomodulation. *Semin Cancer Biol* **43**: 74–89.

Huch M, Gehart H, Van Boxtel R, Hamer K, Blokzijl F, Verstegen MMA, Ellis E, Van Wenum M, Fuchs SA, De Ligt J, et al. 2015. Long-term culture of genome-stable bipotent stem cells from adult human liver. *Cell* **160**: 299–312.

Iqbal J, Nussenzweig A, Lubinski J, Byrski T, Eisen A, Bordeleau L, Tung NM, Manoukian S, Phelan CM, Sun P, et al. 2016. The incidence of leukaemia in women with BRCA1 and BRCA2 mutations: An International Prospective Cohort Study. *Br J Cancer* **114**: 1160–1164.

Janssen A, Van Der Burg M, Szuhai K, Kops GJPL, Medema RH. 2011. Chromosome segregation errors as a cause of DNA damage and structural chromosome aberrations. *Science (80-)* **333**: 1895–1898.

Jia CW, Wang L, Lan YL, Song R, Zhou LY, Yu L, Yang Y, Liang Y, Li Y, Ma YM, et al. 2015. Aneuploidy in early miscarriage and its related factors. *Chin Med J (Engl)* **128**: 2772–2776.

Killcoyne S, Gregson E, Wedge DC, Woodcock DJ, Eldridge MD, de la Rue R, Miremadi A, Abbas S, Blasko A, Kosmidou C, et al. 2020. Genomic copy number predicts esophageal cancer years before transformation. *Nat Med*.

Knouse KA, Lopez KE, Bachofner M, Amon A. 2018. Chromosome Segregation Fidelity in Epithelia Requires Tissue Architecture. *Cell* **175**: 200-211.e13.

Knouse KA, Wu J, Amon A. 2016. Assessment of megabase-scale somatic copy number variation using single-cell sequencing. *Genome Res* **26**: 376–384.

Knouse KA, Wu J, Whittaker CA, Amon A. 2014. Single cell sequencing reveals low levels of aneuploidy across mammalian tissues. *Proc Natl Acad Sci U S A* **111**: 13409–13414.

Kwong LN, Dove WF. 2009. APC and its modifiers in colon cancer. *Adv Exp Med Biol* **656**: 85–106.

Lamm N, Masamsetti VP, Read M, Biro M, Cesare A. 2018. ATR and mTOR regulate F-actin to alter nuclear architecture and repair replication stress. *bioRxiv* 451708.

MacKenzie KJ, Carroll P, Martin CA, Murina O, Fluteau A, Simpson DJ, Olova N, Sutcliffe H, Rainger JK, Leitch A, et al. 2017. CGAS surveillance of micronuclei links genome instability to innate immunity. *Nature* **548**: 461–465.

Malone JL, Nelson AC, Lieberman R, Anderson S, Holt JT. 2009. Oestrogen-mediated phosphorylation and stabilization of BRCA2 protein in breast. *J Pathol* **217**: 380–388.

Martincorena I, Raine KM, Gerstung M, Dawson KJ, Haase K, Van Loo P, Davies H, Stratton MR, Campbell PJ. 2017. Universal Patterns of Selection in Cancer and Somatic Tissues. *Cell* **171**: 1029-1041.e21.

Martinez P, Mallo D, Paulson TG, Li X, Sanchez CA, Reid BJ, Graham TA, Kuhner MK, Maley CC. 2018. Evolution of Barrett's esophagus through space and time at single-crypt and whole-biopsy levels. *Nat Commun* **9**: 1–12.

Morrison SJ, Scadden DT. 2014. The bone marrow niche for haematopoietic stem cells. *Nature* **505**: 327–334.

Muz B, de la Puente P, Azab F, Azab AK. 2015. The role of hypoxia in cancer progression, angiogenesis, metastasis, and resistance to therapy. *Hypoxia* **3**: 83.

Pavelka N, Rancati G, Zhu J, Bradford WD, Saraf A, Florens L, Sanderson BW, Hattem GL, Li R. 2010. Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. *Nature* **468**: 321–325.

Petrucelli N, Daly MB, Pal T. 1993. *BRCA1- and BRCA2-Associated Hereditary Breast and Ovarian Cancer*. University of Washington, Seattle

Roper J, Tammela T, Cetinbas NM, Akkad A, Roghanian A, Rickelt S, Almeqdadi M, Wu K, Oberli MA, Sánchez-Rivera FJ, et al. 2017. In vivo genome editing and organoid transplantation models of colorectal cancer and metastasis. *Nat Biotechnol* **35**: 569–576.

Ross-Innes CS, Becq J, Warren A, Cheetham RK, Northen H, O'Donovan M, Malhotra S, di Pietro M, Ivakhno S, He M, et al. 2015. Whole-genome sequencing provides new insights into the clonal architecture of Barrett's esophagus and esophageal adenocarcinoma. *Nat Genet* **47**: 1038–1046.

Rozhok AI, Salstrom JL, DeGregori J. 2016. Stochastic modeling reveals an evolutionary mechanism underlying elevated rates of childhood leukemia. *Proc Natl Acad Sci U S A* **113**: 1050–1055.

Rutledge SD, Douglas TA, Nicholson JM, Vila-Casadesús M, Kantzler CL, Wangsa D, Barroso-Vilares M, Kale SD, Logarinho E, Cimini D. 2016. Selective advantage of trisomic human cells cultured in non-standard conditions. *Sci Rep* **6**: 1–12.

Santaguida S, Richardson A, Iyer DR, M'Saad O, Zasadil L, Knouse KA, Wong YL, Rhind N, Desai A, Amon A. 2017. Chromosome Mis-segregation Generates Cell-Cycle-Arrested Cells with Complex Karyotypes that Are Eliminated by the Immune System. *Dev Cell* **41**: 638-651.e5.

Sharp AJ, Locke DP, McGrath SD, Cheng Z, Bailey JA, Vallente RU, Pertz LM, Clark RA, Schwartz S, Segraves R, et al. 2005. Segmental duplications and copy-number variation in the human genome. *Am J Hum Genet* **77**: 78–88.

Sheltzer JM, Blank HM, Pfau SJ, Tange Y, George BM, Humpton TJ, Brito IL, Hiraoka Y, Niwa O, Amon A. 2011. Aneuploidy Drives Genomic Instability in Yeast. *Science (80-)* **333**: 1026–1030.

Silk AD, Zasadil LM, Holland AJ, Vitre B, Cleveland DW, Weaver BA. 2013. Chromosome missegregation rate predicts whether aneuploidy will promote or suppress tumors. *Proc Natl Acad Sci U S A* **110**.

Soto M, García-Santisteban I, Krenning L, Medema RH, Raaijmakers JA. 2018. Chromosomes trapped in micronuclei are liable to segregation errors. *J Cell Sci* **131**.

Storchova Z, Kuffer C. 2008. The consequences of tetraploidy and aneuploidy. *J Cell Sci* **121**: 3859–3866.

Sun S, Osterman MD, Li M. 2019. Tissue specificity of DNA damage response and tumorigenesis. *Cancer Biol Med* **16**: 396–414.

Teixeira VH, Pipinikas CP, Pennycuick A, Lee-Six H, Chandrasekharan D, Beane J, Morris TJ, Karpathakis A, Feber A, Breeze CE, et al. 2019. Deciphering the genomic, epigenomic, and transcriptomic landscapes of pre-invasive lung cancer lesions. *Nat Med* **25**: 517–525.

Thompson SL, Compton DA. 2008. Examining the link between chromosomal instability and aneuploidy in human cells. *J Cell Biol* **180**: 665–672.

Torres EM, Sokolsky T, Tucker CM, Chan LY, Boselli M, Dunham MJ, Amon A. 2007. Effects of aneuploidy on cellular physiology and cell division in haploid yeast. *Science (80-)* **317**: 916–924.

Vogelstein B, Kinzler KW. 2004. Cancer genes and the pathways they control. *Nat Med* **10**: 789–799.

Vousden KH, Lane DP. 2007. p53 in health and disease. *Nat Rev Mol Cell Biol* **8**: 275–283.

Wang RW, Vigano' S, Ben-David U, Amon A, Santaguida S. 2020. Aneuploid cells activate NF-κB to promote their immune clearance by NK cells. *bioRxiv* 2020.06.25.172239.

Weinberg F, Ramnath N, Nagrath D. 2019. Reactive oxygen species in the tumor microenvironment: An overview. *Cancers (Basel)* **11**.

Williams BR, Prabhu VR, Hunter KE, Glazier CM, Whittaker CA, Housman DE, Amon A. 2008. Aneuploidy Affects Proliferation and Spontaneous Immortalization in Mammalian Cells. *Science (80-)* **322**: 703–709.

Yona AH, Manor YS, Herbst RH, Romano GH, Mitchell A, Kupiec M, Pilpel Y, Dahan O. 2012. Chromosomal duplication is a transient evolutionary solution to stress. *Proc Natl Acad Sci U S A* **109**: 21010–21015.

Zimmermann A. 2004. Regulation of liver regeneration. *Nephrol Dial Transplant* **19**.