Single cell sequencing reveals low levels of aneuploidy across mammalian tissues

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
Single cell sequencing reveals low levels of aneuploidy across mammalian tissues

Kristin A. Knouse,a,b Jie Wu,c Charles A. Whittakerc, and Angelika Amona,d,1

1Koch Institute for Integrative Cancer Research, Department of Biology, and 2Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA 02139; 3Division of Health Sciences and Technology, Harvard Medical School, Boston, MA 02115; and 4The Barbara K. Ostrom (1978) Bioinformatics and Computing Facility in the Swanson Biotechnology Center, Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139

Contributed by Angelika Amon, August 11, 2014 (sent for review July 24, 2014)

Whole-chromosome copy number alterations, also known as aneuploidy, are associated with adverse consequences in most cells and organisms. However, high frequencies of aneuploidy have been reported to occur naturally in the mammalian liver and brain, fueling speculation that aneuploidy provides a selective advantage in these organs. To explore this paradox, we used single cell sequencing to obtain a genome-wide, high-resolution assessment of chromosome copy number alterations in mouse and human tissues. We find that aneuploidy occurs much less frequently in the liver and brain than previously reported and is no more prevalent in these tissues than in skin. Our results highlight the rarity of chromosome copy number alterations across mammalian tissues and argue against a positive role for aneuploidy in organ function. Cancer is therefore the only known example, in mammals, of altering karyotype for functional adaptation.

DNA copy number alterations, deviations from the diploid DNA content, can vary in length from a few base pairs to entire genomes. When the entire genome is duplicated, cells are referred to as polyploid. Aneuploidy refers to a change in the copy number of individual chromosomes. The changes in gene copy number caused by these karyotype alterations typically change expression of the affected genes (1–3). Therefore, copy number alterations can have profound effects on cellular and organismal physiology. This is especially true for aneuploidy, in which the relative dosage of many genes is altered.

In multicellular organisms, aneuploidy can be present either in all cells, termed “constitutional aneuploidy,” or only in select cells, called “somatic aneuploidy.” Constitutional aneuploidy typically has adverse consequences for the organism. In humans a third copy of chromosome 21 causes Down syndrome, and a third copy of any other autosome almost always results in embryonic or childhood lethality (4). Supporting the adverse effects of aneuploidy at the organismal level, chromosome copy number changes are also associated with impaired fitness at the cellular level. Aneuploid yeast and mammalian cells exhibit slower proliferation and proteotoxic stress (2, 3, 5, 6).

Somatic aneuploidy also has adverse consequences for the organism. In humans, mutations in BUB1B cause the disease mosaic variegated aneuploidy (MVA) (7). BUB1B encodes BubR1, a component of the spindle assembly checkpoint required for accurate chromosome segregation (8). Cells of MVA patients are frequently aneuploid for one or more chromosomes, and MVA patients exhibit growth retardation, developmental defects, and childhood cancers. Mice homozygous for a hypomorphic mutation of BUB1B (BUB1B<sup>1/2</sup>) show similar phenotypes (9). Cancer is another example of a disease characterized by aneuploidy. In this case, however, changes in gene copy number are thought to increase the fitness of tumor cells relative to untransformed cells because of selection for oncogene gains and tumor suppressor gene losses (10).

In light of the many reports describing significant adverse effects of aneuploidy on cellular and organismal fitness, it is surprising that two organs, the liver and brain, are reported to harbor high levels of aneuploidy. Hepatocytes, the primary cell type of the liver, are diploid in neonates. However, hepatocytes become polyploid during early postnatal development by initiating mitosis but failing cytokinesis (11, 12). Importantly, a recent reexamination of hepatocyte ploidy by FISH reported aneuploidy in more than 50% of hepatocytes (13, 14). Duncan et al. (13) suggested that polyploid hepatocytes underwent aberrant mitoses with high levels of chromosome mis-segregation, thereby generating aneuploidy among polyploid cells. The authors further speculated that aneuploidy endows hepatocytes with phenotypic variability and adaptability upon exposure to various noxious agents and metabolic stresses (15).

Rehen et al. (16) were first to report elevated aneuploidy in the mammalian brain. Using spectral karyotyping (SKY), they reported that 33% of embryonic mouse neural progenitor cells harbored loss or gain of one or more chromosomes. Subsequent studies using FISH indicated that approximately 20% of adult mouse and human brain cells were aneuploid (16–20). These aneuploid cells were believed to arise from chromosome mis-segregation events in neural progenitor cells and were shown to integrate into brain circuitry (21, 22). These high levels of aneuploidy were proposed to provide the brain with its notable phenotypic diversity while simultaneously predisposing the organ to neurodegeneration (23–26).

The reports of aneuploidy in the liver and brain suggest that, in these tissues, aneuploidy may not compromise organ function but instead provide beneficial phenotypic plasticity. Indeed a positive role for aneuploidy is well documented in experimental evolution studies in microorganisms (27, 28). The reports of high levels of aneuploidy in the brain and liver thus raised the exciting possibility that these organs somehow avoid the adverse effects at the cellular and organismal levels.

Significance

Aneuploidy refers to the gain or loss of individual chromosomes within a cell. Typically, aneuploidy is associated with detrimental consequences at both the cellular and organismal levels. However, reports of high levels of aneuploidy in the brain and liver suggested that aneuploidy might play a positive role in these organs. Here we use single cell sequencing to determine the prevalence of aneuploidy in somatic tissues. We find that aneuploidy is a rare occurrence in the liver and brain and is no more prevalent in these tissues than in skin. Our results demonstrate high karyotypic stability in somatic tissues, arguing against a role for aneuploidy in organ function and reinforcing its adverse effects at the cellular and organismal levels.

Author contributions: K.A.K. and A.A. designed research; K.A.K. performed research; K.A.K., J.W., and C.A.W. analyzed data; and K.A.K. and A.A. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Data deposition: The sequences reported in this paper have been deposited in the National Center for Biotechnology Information Sequence Read Archive (accession no. SRP041670).

1To whom correspondence should be addressed. Email: angelika@mit.edu. This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1415287111/-/DCSupplemental.
consequences of aneuploidy and instead use these changes to their benefit, all while avoiding oncogenic transformation. Here, we use single cell sequencing to provide a genome-wide, high-resolution assessment of aneuploidy in mammalian tissues. Unlike previous reports, we find that the prevalence of aneuploidy in liver and brain is low and comparable to the frequency in skin. Our findings argue against a positive role for somatic chromosome copy number alterations in tissue function and instead reinforce their adverse consequences for cellular and organismal fitness.

**Results**

**Detecting Somatic Aneuploidy by Single Cell Sequencing.** The studies reporting aneuploidy in the brain and liver used FISH and SKY to quantify aneuploidy. Both methods are prone to artifacts. In FISH, probe stretching, probe clustering, failed hybridization, or off-target hybridization can cause signals to be lost or gained inappropriately. Even a low frequency of artifacts for a single chromosome can lead to a gross overestimation of aneuploidy when extrapolated across all chromosomes. SKY can also overestimate aneuploidy because the spreading procedure can...
move chromosomes between nuclei. We therefore used single cell sequencing as an alternative method to assess somatic aneuploidy in brain and liver. We dissociated tissues from mice and humans, amplified the genomic DNA of single cells, sequenced the amplified DNA, and inferred copy number from sequencing read depth (Fig. 1A). When possible, we isolated tissues from males because monosomy of the X chromosome provides an internal control for identifying copy number alterations. We avoided permeabilization and fixation of target cells to minimize interference with whole-genome amplification. In the liver, we isolated nuclei instead of whole cells to facilitate detection of copy number changes in a polyploid setting.

For each cell, on average, 8.5 × 10^6 reads aligned to the reference genome, corresponding to 0.1×X coverage. We quantified reads in 500-kb windows and used a hidden Markov model to generate segments of defined copy number relative to a euploid reference (29). We did not analyze the Y chromosome for copy number alterations because of its low amount of unique sequence.

Approximately 10% of sequenced cells exhibited aberrantly high intrachromosomal sequence depth variability (Fig. S1A, Top and Upper Middle), as might arise from incomplete fragmentation of genomic DNA or biased amplification of fragments during whole-genome amplification. To identify these cells and exclude them from the analysis, we measured the SDs of corrected read copies (log2-based) within sliding windows (30 adjacent 500-kb bins) across a single chromosome, averaged the SDs for each chromosome, computed the mean average SD across the three most variable autosomes, and excluded cells for which this variability score (VS) exceeded 0.34 (SI Materials and Methods and Fig. S1 B and C). In total, we excluded 10.1% of sequenced cells (Fig. S1D). This metric does not bias against aneuploidy, because variability is measured within each chromosome. Consistent with this, the average VS of all aneuploid cells we identified was equivalent to the average VS of all euploid cells (Fig. S1A, Lower Middle and Bottom, and S1C).

We validated several aspects of our approach to ensure that this method reliably detects somatic aneuploidy. First we confirmed that loss or gain of a chromosome caused a change in relative read depth that exceeded background fluctuations in read depth. We sequenced single brain cells from trisomy 16 mouse embryos (3) and were indeed able to detect an additional copy of chromosome 16 in all cells we measured the SDs of corrected read copies ([log2-based] within sliding windows (30 adjacent 500-kb bins) across a single chromosome, averaged the SDs for each chromosome, computed the mean average SD across the three most variable autosomes, and excluded cells for which this variability score (VS) exceeded 0.34 (SI Materials and Methods and Fig. S1B, Right). We next determined whether the method was able to detect aneuploidies of unknown composition. We sequenced single brain cells from an adult male BUB1B<sup>H14I</sup> mouse, because the BUB1B mutation is reported to cause aneuploidy in one-third of cells (9). Indeed, 8 of the 21 brain cells we analyzed [38.1%, 95% confidence interval (CI) 18.1%–61.6%] were aneuploid for one or more chromosomes (Fig. 1C and D).

To determine whether we were able to detect aneuploidy in polyploid cells, we sequenced hepatocyte nuclei from an adult male BUB1B<sup>H14I</sup> mouse. Here, 3 of the 16 hepatocyte nuclei we analyzed (18.8%, 95% CI 4%–45.6%) harbored one or more chromosomes with read depths indicating loss or gain in a tetraploid nucleus (Fig. 1C and E). We conclude that our method can detect single as well as multiple chromosome gains or losses in diploid and polyploid cells.

Last, we tested whether our approach selected against aneuploid cells because of differences in viability during tissue dissociation or operator bias during microaspiration. We mixed equal volumes of trisomy 16 and euploid brain, dissociated the tissue together, and sequenced single cells from this mixture. We found that five of seven cells (71.4%, 95% CI 29%–96.3%) were trisomic for chromosome 16, indicating that our procedure does not discriminate against aneuploid cells.

Together, these experiments show that single cell sequencing can detect simple and complex aneuploidies in an unbiased manner. Furthermore, by compiling the relative copy numbers (log2 ratios) for all aneuploid chromosomes and X chromosome monosomes, we were able to establish cutoffs for detecting copy number variants in all subsequently sequenced diploid and polyploid cells (SI Materials and Methods).

Prevalence of Aneuploidy in Skin. There are no reports of elevated aneuploidy in skin. Therefore, we sequenced single cells from skin to establish a baseline level of somatic aneuploidy. We isolated keratinocytes by dissociating epidermis and selecting uncornified cells. In the mouse we found one aneuploid cell (trisomy 12) among 37 cells analyzed, indicating that 2.7% of mouse keratinocytes are aneuploid (95% CI 0.1%–14.2%; Figs. S2 and S3 and Table S1). To examine aneuploidy in human skin, we isolated keratinocytes from epidermis obtained upon autopsy of a 48-y-old woman and a 68-y-old man. Of the 53 keratinocytes analyzed, none were aneuploid (Fig. S2 and Table S1). We conclude that aneuploidy in human skin is low (95% CI 0–6.7%; Table S1).

Prevalence of Aneuploidy in Brain. In the mammalian brain, aneuploidy is reported to be highest in neural progenitor cells, with a prevalence of 33% (16). To investigate aneuploidy in embryonic neural progenitor cells, we isolated brains from mouse embryos expressing the neural progenitor marker nestin-GFP. Of the 36 cells analyzed, none were aneuploid (Fig. S2 and Table S1). To assess the degree of aneuploidy in the adult mouse brain, we dissociated gray matter from cerebral cortex and enriched for neurons and glia by density gradient sedimentation. Single cell gene expression analysis revealed that the cell suspension obtained by this method contained 75% neurons, 12.5% glia, and 12.5% unknown cell types (Fig. S4A). Of the 43 cells analyzed, we identified one aneuploid cell (trisomy 15; Figs. S2 and S4B). Thus, only 2.3% of mouse brain cells are aneuploid (95% CI 0.1%–12.3%; Table S1).

To specifically examine neurons, we immunostained adult mouse brain cells with an antibody against the neuronal nuclear marker NeuN and enriched for NeuN-positive cells by FACS. Of 18 cells sequenced, we had to exclude 9 cells because of high variability in read depth between adjacent genomic windows (Fig. S1 B–D). Of the 9 cells analyzed, none were aneuploid (Fig. S2). We presume that the high VS in many of the neurons isolated by this method was due to permeabilization, fixation, and/or immunostaining interfering with whole-genome amplification. As an alternate approach, we purified DRD2-expressing medium spiny neurons of the basal ganglion from an adult mouse carrying a DRD2-GFP transgene. We sequenced 10 cells, all of which could be analyzed and none of which were aneuploid (Fig. S2). We conclude that the sequence variability observed in the neurons purified by NeuN immunostaining was due to the immunostaining procedure interfering with sample preparation. Importantly, our analysis of 19 neurons did not reveal a single aneuploid cell (95% CI 0–17.6%; Fig. S2 and Table S1). This result is not significantly different from that for the mixed brain cell population. In total, the prevalence of aneuploidy in the mouse brain is 1% (95% CI 0–5.6%; Fig. 2A).

To extend these studies to humans, we dissociated gray matter of frontal lobe obtained during autopsy of four individuals: a 52-y-old man, 68-y-old man, 70-y-old man, and a 48-y-old woman. None of these individuals had a history of neurologic disease or evidence of brain pathology upon autopsy. Of the 89 cells analyzed, all but two cells were euploid (Fig. 2C and Fig. S2). Of the aneuploid cells, one harbored a monosomy for chromosome 22 (Fig. 2D). The other cell had a read depth on chromosome 18 consistent with trisomy in a tetraploid cell (Fig. S4C). Because cell suspensions were extremely dilute, we believe that this indeed represents a tetraploid cell with chromosome loss rather than two cells isolated together, one of which was aneuploid. We conclude that the prevalence of aneuploidy in the adult human brain is 2.2% (95% CI 0.3%–7.9%; Fig. 2B and Table S1), significantly less than the more than 20% aneuploidy reported by prior studies (16–20).

Knouse et al.
Aneuploidy is more difficult to detect in polyploid cells because the adjusted read depth on the aneuploid chromosome is smaller. As described above, we were able to detect aneuploidy in tetraploid hepatocyte nuclei from BUB1B(H14) mice (Fig. 1E). Importantly, in all presumably diploid cells we analyzed from other tissues, only one brain cell harbored a chromosome with an adjusted read depth consistent with a single chromosome gain or loss in a tetraploid cell. Thus, the chromosome changes we observed in BUB1B(H14) hepatocyte nuclei likely represent actual aneuploidy rather than background fluctuations in read depth. Detection of aneuploidy in octaploid nuclei is more challenging because the adjusted read depth from a single chromosome loss or gain is further reduced. Indeed, 14% of all diploid cells analyzed harbored one or more chromosomes with an average read depth that would be consistent with a chromosome gain or loss in an octaploid cell. Thus, chromosome gains or losses in an octaploid cell cannot be distinguished from background fluctuations in read depth. We therefore did not consider fluctuations of this magnitude as aneuploidy. This did not significantly affect the outcome of our analysis because octaploid hepatocyte nuclei are rare (Fig. 3B and Fig. S5A).

Having established cut-offs for detecting aneuploidy in a polyploid setting, we isolated nuclei from mouse liver. To enrich for hepatocytes, we preferentially picked nuclei with diameters consistent with tetraploidy (Fig. S5A). Of the 66 cells analyzed, we did not identify a single aneuploid nucleus (Fig. 3C, Fig. S2, and Table S1). We next sequenced commercially available human hepatocytes prepared from a 46-y-old man and a 51-y-old man. Neither of these individuals had liver disease; however, the second individual had an extensive drug and alcohol history. The purity of these hepatocyte preparations exceeds 70%. Of the 39 hepatocytes analyzed from the first donor, all cells except one were euploid (Fig. 3E and Fig. S2). Although we could not assign a karyotype to the other cell owing to extensive fluctuation of segments, it met our quality control criteria and was considered aneuploid (Fig. 3F, Upper). Of the 61 nuclei analyzed from the second donor, all but three cells were euploid (Fig. 3E and Fig. S2). All three aneuploid nuclei were tetraploid with gains of one or two chromosomes (Fig. 3F, Lower, and Fig. S5B). In total, the prevalence of aneuploidy in human hepatocytes is 4% (95% CI 1.1%–9.9%; Fig. 3D). Thus, single cell sequencing indicates that less than 5% of hepatocytes are aneuploid (Table S1).

Discussion
Here we provide the first assessment (to our knowledge) of chromosome copy number alterations in multiple tissues at genome-wide, single cell resolution. Our sequencing results suggest that the prevalence of aneuploidy in the mammalian brain and liver is less than 5% (Table S1). This observation is in agreement with a recent single-neuron sequencing study that, despite identifying a high frequency of subchromosome copy number variants, did not observe high levels of whole-chromosome aneuploidy in neurons (30). Prior reports used SKY and FISH to assess the prevalence of aneuploidy and found it to exceed 50% and 20% in the liver and brain, respectively (13, 14, 16–20). We attribute this difference to drawbacks associated with using FISH to detect somatic aneuploidy. Indeed, using dual-color, single-chromosome FISH we too observed low levels of aneuploidy for a single chromosome that would, when extrapolated across all chromosomes, indicate that over 40% of hepatocytes and 10% of brain cells are aneuploid (Fig. 3B). However, we also observed similar rates of aneuploidy in skin. That we had to exclude up to one-third of cells from analysis because of discrepancies in signal number between two probes targeting the same chromosome illustrates the limitations of FISH. The rare cases of signal gain or loss for both probes could just as likely represent two simultaneous hybridization artifacts. FISH artifacts that are interpreted as aneuploidy even when occurring at low frequency can lead to gross

![Fig. 2. Prevalence of aneuploidy in brain. (A and B) Prevalence of euploid and aneuploid cells in mouse (A) and human (B) brain. (C) Segmentation plots of euploid human brain cells. (D) Segmentation plot of an aneuploid human brain cell harboring a monosomy for chromosome 22.](http://www.pnas.org/cgi/doi/10.1073/pnas.1415287111)
overestimation of aneuploidy when used to infer the prevalence of aneuploidy across the genome. We conclude that single cell sequencing is a superior method for characterizing somatic karyotypic changes.

Whether the rare cells harboring copy number alterations that we did identify in normal tissues represent incipient transformation remains to be determined. Furthermore, it is important to note that we only analyzed skin, brain, and liver tissue that had no evidence of disease. It is possible that aneuploidy is elevated in disease states other than cancer. Indeed, it is interesting to note that proteotoxic stress is a hallmark of both aneuploidy and neurodegenerative disease (6, 31).

The Implications of Low Aneuploidy for Hepatocyte Proliferation.

The low prevalence of aneuploidy in normal tissues means that either the events producing such aberrations occur rarely or the cells harboring such changes are eliminated. Both possibilities are likely to occur. The wide array of detrimental phenotypes associated with constitutional aneuploidy in vitro and in vivo suggests that aneuploid somatic cells may indeed be selected against when arising amid a euploid population. On the other hand, our sequencing of BUB1B<sup>+/−</sup> mice indicates that cells can survive despite multiple copy number alterations, although their survival in these mice may be due to the presence of fewer euploid cells and thus reduced selective pressure (9).

The low level of aneuploidy has particular implications for our understanding of liver biology. Hepatocytes are not the only example of polyploidy in mammals. Trophoblast giant cells, which constitute the placenta, and megakaryocytes, the precursors of platelets, are also polyploid. However, in these two cases, and indeed in cases in other organisms, polyploidy is associated with a terminally differentiated state and cell cycle arrest (32). Hepatocytes are an exception to this rule. Upon liver damage, hepatocytes reenter the cell cycle and proliferate extensively (33). This is a nontrivial task because polyploid cells, in addition to having multiple copies of their genome, have multiple centrosomes. When a cell with multiple centrosomes enters mitosis, a multipolar spindle forms that, if uncorrected, randomly segregates chromosomes to produce highly aneuploid cells. Studies of cancer cells have shown that multiple centrosomes can cluster to
form bipolar spindles (34). However, the process of centrosome clustering produces many merotelic kinetochore-microtubule attachments, where a kinetochore attaches to microtubules emanating from both spindle poles. Chromosomes with merotelic attachments are frequently mis-segregated (35). It is difficult to imagine that the chromosome segregation error rate increases significantly, and the observed frequency of aneuploidy expected to arise from these mitotic errors is remarkably low. This implies that merotelic attachments have either evolved mechanisms for detecting and correcting merotelic attachments or use a unique mechanism for segregating chromosomes in the presence of multiple centrosomes. Determining how merotelic attachments prevent or correct merotelic attachments will be an important future question.

The Implications of Low Levels of Aneuploidy for Neural Diversity. The brain is notable for its remarkable diversity, both at the cellular and functional level. There are more than 10,000 different types of neurons, and even neurons of the same type can differ in their expression of synapse components and their thresholds for excitation and firing (23). Diversity at the cellular level presumably translates into diversity at the functional level. There is remarkable interindividual heterogeneity in cognition, intellect, and behavior. For example, even isogenic mouse brain regions exhibit remarkable diversity in stress responses (36). The report of high levels of aneuploidy in the brain led to extensive speculation that random aneuploidies endowed the brain with cellular and thus functional diversity (23). However, our study shows that somatic aneuploidy is much less common than previously reported. Given the low prevalence of aneuploidy in normal tissues, it is difficult to imagine that such changes play a positive role in organ function. We infer that normal mammalian tissues do not use aneuploidy as a means of phenotypic diversification or adaptation.

Materials and Methods

Tissue Sources. Frozen human liver was obtained from the Brigham and Women’s Hospital Tissue and Blood Repository. Human hepatocytes were purchased from Life Technologies. Fresh human frontal lobe and epidermis were obtained upon autopsy at Massachusetts General Hospital.

Single Cell Sequencing and Copy Number Analysis. Tissues were dissociated into single cells. Single cells and nuclei were isolated by microsuspension, and genomic DNA was amplified using the GenomePlex Single Cell Whole Genome Amplification Kit (Sigma). Amplified DNA was purified, barcoded, pooled, and sequenced on an Illumina HiSeq2000. Sequencing reads were aligned using BWA (0.6.1). HMMcopy (0.1.1) was used to estimate gene copy number in 500-kb bins (29). Cells with high variability in copy number across bins were excluded from analysis.

ACKNOWLEDGMENTS. We thank the Massachusetts Institute of Technology BioMicro Center for sequencing; the Koch Institute Swanson Biotechnology Center for flow cytometry and histology; the Whitehead Institute Keck Microscopy Facility for imaging; Jan van Deursen, David Scadden, and David Housman for providing mouse strains; Matthew Frosch, Daniel Mordes, and Inma Cobos Sillero for providing autopsy samples; Zachary Crook for technical assistance; Mark Daly for helpful discussions; and David Housman, Frank Solomon, Mark Daly, and members of the A.A. laboratory for their critical reading of the manuscript. This work was supported by National Institutes of Health (NIH) Physical Sciences Oncology Center Grant 5-USA CA143874, a senior scholar award from the Ellison Medical Foundation, Koch Institute Support Grant P30-CA14051 from the National Cancer Institute, and by grants from the Howard Hughes Medical Institute and the Kathy and Curt Marble Cancer Research Fund. The Neuropathology Core of the Massachusetts Alzheimer Disease Research Center is supported by NIH Grant P50 AG005134. K.A.K. is supported by National Institute of General Medical Sciences Training Grant T32GM007753.