Review Article — Open Access

Drosophila Cell Lines to Model Selection for Aneuploid States

Hangnoh Lee and Brian Oliver*

Laboratory of Cellular and Developmental Biology, National Institute of Diabetes and Kidney and Digestive Diseases, National Institutes of Health, USA

Abstract

Abnormal numbers of chromosomes, or aneuploid segments of chromosomes, are associated with multiple genetic disorders and cancers. In many chromosomal abnormalities, it is thought that genic balance of protein complexes or pathways are disrupted. In cancers and immortal cell lines, it is thought that aneuploidy confers a growth and senescence advantage. The karyotype and gene expression profiles of 19 Drosophila modENCODE cell lines highlight the evolution of advantageous gene copy numbers while maintaining genic balance. These highly aneuploid cells show coherent changes in copy number among genes encoding components of multiprotein complexes, which may reflect strong selection for genic balance. They also show copy number increases in genes that positively regulate cell cycle progression or decreases in copy number of genes that negatively regulate cell cycle progression, highlighting multiple evolutionary paths to increased growth. Some copy number changes, both increases and decreases, are recurrent. This suggests that there are some critical primary drivers of evolving the ability to grow in vitro. The small, highly rearranged genome, of Drosophila cell lines provides a powerful model system for studying numerical changes in genome, their effect, and dosage compensation against the effect.

Main Text

Non-diploid states are implicated in multiple genetic disorders including Down syndrome, Turner syndrome and Klinefelter syndrome [1]. Smaller scale DNA copy number variations (CNVs) are observed in patients with psychiatric disorders, such as autism spectrum disorder [2]. Primary effects of these genome aberrations are due to altered gene-dosage rather than production of defective proteins, highlighting the importance of quantitative gene balance in addition to gene quality. Imbalance in gene dose is believed to induce stoichiometric imbalance in protein complexes, which in turn results in partially formed complexes or proteotoxicity that leads to poor fitness of organisms [3]. Gene dose effects are not restricted to the genes showing abnormal dose and there are propagating effects in gene networks that result in misexpression throughout the genome [4].

Gene dose effects can be compensated or buffered, although this is far from uniform. In the specific cases of sex chromosomes, chromosome-wide compensating mechanisms exist. For example, X inactivation occurs in mammalian females [5] creating a functionally monosomic state, which in both females and males is countered by increased expression from the X chromosome [6]. The single X is hyperactivated in Drosophila males [7], and Painting of fourth (POF) regulates chromosome 4 genes in Drosophila enabling survival of haplo-4 individuals [8]. The 4ᵗʰ chromosome is believed to be an ancient X chromosome, suggesting that this compensation is derived. Autosomal dosage compensation is more gene specific and is at least partially achieved by interactions in gene networks [4]. Understanding the mechanisms for dosage compensation could be important for evaluating the effect of chromosomal abnormalities, and perhaps to develop therapeutic approaches. For example, introduction of the X-inactivation gene (*XIST*) to a trisomic chromosome 21 by genome editing resulted in a transcriptionally inactive chromosome 21 in induced pluripotent stem cells, derived from Down syndrome patient cells [9]. Improved cell growth and neurogenesis of the cells in this study illustrate how dosage compensation mechanism can be used to cope with chromosomal abnormalities.

Drosophila melanogaster provides an excellent model system in exploring chromosomal abnormalities. They have only 8 chromosomes: 3 pairs of autosomes (Chromosome 2, 3, and 4) and two sex chromosomes (X and Y). The 143.9 Mb genome is much smaller than that of mammals, making genome-wide copy number assays by resequencing much simpler. Flies that are monosomic or trisomic can be produced using "attached" chromosomes where two homologs share the same centromere or using mutations that increase the rate of meiotic non-disjunction [10]. Multiple panels of flies have been generated that have molecularly defined engineered deletions or duplications of segments of the genome [11-14]. For example, flies from Drosophila Deletion collection [12,13] have been used to assess the effect of gene copy loss, enabling systematic description of gene expression with reduced copy numbers [4,15].

Tissue culture cells are good models for cancer, as they exhibit many of the same properties, such as abrogated contact inhibition of cell division. Importantly, both are typically aneuploid. There has been much discussion of whether aneuploidy is a secondary effect of genome instability that gives rise to mutations in important genes or if genic balance is a driver of the cancer phenotype [16]. Regardless, the aneuploid state of a cell line is an opportunity to observe the quantitative influence of genes on the phenotype of immortal growth *in vitro*. We investigated chromosome rearrangements in Drosophila cell lines by performing karyotyping and DNA resequencing [17]. The karyotypes were highly variable both between and within cell lines. Some cell lines, such as *1182-4H* and *L1*, are quasi diploid. While, two widely used cell lines, *Kc167* and *S2-DRSC*, display highly rearranged genomes with

tetra- or quasi-tetraploidy basal states. There is great variability within cells lines. In the case of *mbn2* cells, they exhibited highly polyploidy configuration with 16 to 30 chromosomes per cell. Interestingly even though there are differences in karyotype among cells in the same culture based on cytological examination, culture to culture populations of *S2-DRSC* and *Kc167* cells show nearly identical DNA sequence copy number profiles, suggesting that the population of cells with differing karyotypes are stable [17,18]. Cell lines demonstrated varied karyotype even when they were established from the same origin. *D8* and *D9* were generated from the same lab using the same strain of flies [19], but have diploidy and quasi-tetraploidy configurations, respectively. This indicates that the particular genotype of source animal does not dictate the aneuploid state of immortalized cells from that individual. Thus, we suggest that driver mutations do not inevitably lead to a particular final karyotype [17].

Not all regions of the genome are equally likely to show a copy number change in tissue culture cells. For example, loss or gain of chromosome 4 was observed in many cell lines. The chromosome 4 has 1.3 Mb of total length with very limited genetic materials [20]. Because the chromosome is much smaller than other autosomes, adult flies with haploid chromosome 4 are viable [10], thus it is not a surprise that this state can be tolerated in tissue culture cells. Another common observation is frequent loss of Y chromosome in male cell lines. We could identify cells with only a male-like dose X chromosome in *D20-c5* (tetraploid) or *L1* (diploid) cells without the gene-poor Y chromosome, while some other cell lines like *Cl.8* or *BG3-c2* maintained their Y chromosome. Considering Drosophila sex determination is based on the number X chromosome, rather than existence of Y chromosome [21], they are male cell lines, and accordingly, displayed male-like characteristics (e.g. male specific splicing events and active X chromosome dosage compensation). However, we do not know if the cell lines were established by losing Y chromosome from originally male cells, or by losing one of the X chromosomes from originally female cells during tissue culture selection. Future study of how the cells lose one of the sex chromosome, and how the cells can mitigate the consequent change on transcriptome will be interesting questions in the field, and may provide a model for Sex chromosome-related defects, such as Turner syndrome.

The focal changes were more interesting. In 19 different cell lines, ~ 1,700 duplicated and ~ 400 deleted segments can be found. Duplication events have median of 37 Kb and deletions were with 97 kb (there was ascertainment bias, as we did not sequence deeply enough to detect small indels). While most of the numerical changes were highly cell line-specific, suggesting at least some degree of randomness in genome organization, we found 89 regions with recurrent copy number increase, and 19 with copy number decrease. Recurrent events could be due to regions of chromatin fragility that are more likely to break, selection following random breaks, or both. From total 2,411 loci where copy number continuity breaks in any of the 19 cell lines, we identified 51 "hotspots". Among them, 27 regions contain DNA repeats, such as long terminal repeats (LTR), long interspersed elements (LINE), or simple repeats, within in 1 kb range. This observation underscores the role of structural factors in boosting recurrent copy number events based on sequence-homology mediated recombination.

In addition to regions susceptible to breaking, we found evidence that altered copy number is selected. For example, we observed highly recurrent duplicated regions at a sub-telomeric region of the left arm of chromosome 3. This region covers approximately 3 Mb and is duplicated in 10 out of 19 cell lines. The most overlapping region, duplicated in 15 out of 19 cell lines, covers approximately a small 30 kb segment that contains six annotated genes including the pri-RNA of *bantam*. The *bantam* miRNA has an anti-apoptosis function by negatively regulating the pro-apoptotic function of *head involution defective* (*hid*) [22]. Thus, the observation suggests that gaining more copies of anti-apoptotic, or pro-survival, gene is favored under selection. In addition to *bantam*, the region also contains a very interesting gene, *Ultraviolet-revertible gene 1* (*Rev1*). The gene product is required for switching highly processive DNA polymerase for an error prone DNA polymerase used in repair [23]. Acquiring more copies of the gene may provide diversity (via hypermutability) on which selection could act, although the gene could be selected simply based on the proximity to *bantam* (or *vice versa*). Another noticeable example of recurrent copy number increase is *PDGF- and VEGF-receptor related* (*Pvr*) that is duplicated in 10 different cell lines. *Pvr* also has anti-apoptotic and pro-survival function [24], and the gene is highly expressed among the cell lines [25]. In conclusion, we can speculate that tissue culture cells are selected to gain more copies of genes that give advantages in cell survival and proliferation.

In addition to *bantam* and *Pvr*, we were able to observe further dramatic patterns of copy number gain or loss correlate to function of genes. Gene ontology (GO) analysis of duplicated genes in *S2-DRSC* cells and deleted genes in *Kc167* cells provide a good example of how reconfiguration of genome associates with cell cycle-related functions of genes. For *S2-DRSC* cells, the cell cycle genes that belong to high copy number group are positive regulators of the cell cycle, or proto-oncogenes that include multiple cyclins, Drosophila *Ras*, *cdc25* and others. In *Kc167* cells, we identified copy number loss of negative regulators of cell cycle, or tumor suppressors, such as Drosophila homolog of *Retinoblastoma family gene* (*Rbf*), *Brca2*, and *wee*. This observation supports an idea that tissue culture selection drives reconfiguration of genome to acquire or remove copies of cell cycle related-genes for better proliferation and survival (Table 1. See Additional File 4 in reference [17] for further detail).

While copy number changes can result in selective advantages for cells due to increased cell cycle progression and pro-survival functions, this occurs at the expense of creating genic imbalances in neighboring genes that are detrimental to those cells [3]. During generations of *in vitro* growth some of these imbalances might be corrected by secondary changes in copy number. If for example, there are copy number changes for the genes that encode components of a protein complex,

Gene Name	Cell Line	Changes in copy Number	Functions (references in the main text)
bantam	S2 & Kc	Up	Anti-apoptosis
UV-revertible gene 1 (Rev1)	S2 & Kc	Up	DNA repair
PDGF- and VEGF-receptor related (Pvr)	S2	Up	Pro-survival
Ras oncogene at 85D (Ras85)	S2	Up	Proto-oncogene
cdc25 (string)	S2	Up	Cell cycle control
Retinoblastoma family gene (Rbf)	S2	Down	Tumor Suppressor
Breast cancer 2, early onset homolog (Brca2)	S2	Down	Tumor Suppressor

Table 1: Examples of copy number changes in key survival or cell cycle-related genes observed in Drosophila *S2-DRSC* (S2) and *Kc167* (Kc) cell lines.

it is expected that one copy number change can be further stabilized by subsequent copy number changes of the genes that contribute to the same protein complex. In other words, copy change to maintain stoichiometry. We found such coherent copy number changes among genes encoding protein complex members within the Drosophila cell lines. Some of these copy number configurations may even modify the function of complexes. For example, in *S2R+* cells the genes encoding the regulatory subunit (the "lid") of the proteasome had reduced copy numbers, while genes for the 20S core, where proteins are digested, had increased copy numbers. This observation suggests that: 1) the subparts of the proteasome show coherent copy number, and 2) that proteasome activity was fine-tuned (de-regulated) to increase protein turnover and proteotoxicity in those cells. Interestingly, in aneuploid states in yeast that affect ribosome balance, increased protein turnover is induced [26,27]. Additionally, ubiquitination marks proteins for the proteasome and a loss-of-function mutation in a deubiquitinating enzyme gene improved tolerance against aneuploidy in yeast [28]. Thus this copy number configuration for the proteasome in Drosophila cells lines could have been selected due to a common effect of aneuploidy.

Another way cells could cope with disadvantageous gene dose changes is dosage compensation. In the Drosophila cell lines, genes within duplicated or deleted regions demonstrate varying degrees of compensated gene expression as measured by RNA sequencing [4,7,18,29]. In the two most widely used cell lines, *S2-DRSC* and *Kc167*, when copy number was reduced by 2-fold genes displayed 1.2X fold higher expression. However, the degree of compensation level varied from absent (1.0X) to almost perfect (1.8X) in different cell lines. Autosomal compensation is believed to be largely gene-specific [4,15], and thus overall degree of compensation may differ based on the set of genes that have altered copy number. Alternatively, it is also possible that the varying degree of compensation may be the consequence of different optimization routes against copy number changes in different cell lines. This might be achieved by selecting on copy number and/or on dosage compensation If the later is true, then it may be possible to uncover additional general systems for achieving dosage compensation. For example, it is possible that cells with a highly rearranged genome may also develop faster turnover mechanisms of gene products at the RNA and/or protein level that counteracts the gene dose effect. Investigating how cell lines adapt to growth *in vitro* will be helpful not only to understand the establishment of cell lines, but also lead to a deeper understanding of gene dose effects, gene balance, and dosage compensation.

In summary, we describe numerical abnormalities that exist in Drosophila cell line genomes, and point out how they may provide advantages to cell lines during the immortalization process. Our findings have implications for other cell-level selection progressions, which includes tumorigenesis, where extensive genomic aberration is found. Most importantly, the small but highly rearranged genome of Drosophila cell lines provides a powerful model system in studying numerical changes in genome, their effect, and dosage compensation against the effect. Study of numerical alterations of chromosomes in animal models is often difficult due to common failures in early development. It is especially the case for humans where the large genome size and redundancies in genetic pathways complicate our interpretation of copy number effects. We suggest that use of Drosophila cell line systems, with smaller genomes and streamlined pathways, will be beneficial. Future directions may include tracking generation of aneuploidy through passages while establishing cell lines. We believe that such efforts will lead better understanding of how aneuploidy progresses.

References

1. Dierssen M, Herault Y, Estivill X (2009) Aneuploidy: from a physiological mechanism of variance to Down syndrome. Physiol Rev 89: 887-920.

2. Cook EH Jr, Scherer SW (2008) Copy-number variations associated with neuropsychiatric conditions. Nature 455: 919-923.

3. Veitia RA, Potier MC2 (2015) Gene dosage imbalances: action, reaction, and models. Trends Biochem Sci 40: 309-317.

4. Malone JH, Cho DY, Mattiuzzo NR, Artieri CG, Jiang L, et al. (2012) Mediation of Drosophila autosomal dosage effects and compensation by network interactions. Genome Biol 13: r28.

5. Heard E, Disteche CM (2006) Dosage compensation in mammals: fine-tuning the expression of the X chromosome. Genes Dev 20: 1848-1867.

6. Deng X, Hiatt JB, Nguyen DK, Ercan S, Sturgill D, et al. (2011) Evidence for compensatory upregulation of expressed X-linked genes in mammals, Caenorhabditis elegans and Drosophila melanogaster. Nat Genet 43: 1179-1185.

7. Ferrari F, Alekseyenko AA, Park PJ, Kuroda MI (2014) Transcriptional control of a whole chromosome: emerging models for dosage compensation. Nat Struct Mol Biol 21: 118-125.

8. Johansson AM, Stenberg P, Bernhardsson C, Larsson J (2007) Painting of fourth and chromosome-wide regulation of the 4th chromosome in Drosophila melanogaster. EMBO J 26: 2307-2316.

9. Jiang J, Jing Y, Cost GJ, Chiang JC, Kolpa HJ, et al. (2013) Translating dosage compensation to trisomy 21. Nature 500: 296-300.

10. Ashburner M, Golic KG and Hawley RS (2005) Drosophila : A laboratory handbook (2nd edn), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

11. Parks AL, Cook KR, Belvin M, Dompe NA, Fawcett R, et al. (2004) Systematic generation of high-resolution deletion coverage of the Drosophila melanogaster genome. Nat Genet 36: 288-292.

12. Ryder E, Ashburner M, Bautista-Llacer R, Drummond J, Webster J, et al. (2007) The DrosDel deletion collection: a Drosophila genomewide chromosomal deficiency resource. Genetics 177: 615-629.

13. Ryder E, Blows F, Ashburner M, Bautista-Llacer R, Coulson D, et al. (2004) The DrosDel collection: A set of p-element insertions for generating custom chromosomal aberrations in Drosophila melanogaster. Genetics 167: 797-813.

14. Venken KJ, Popodi E, Holtzman SL, Schulze KL, Park S, et al. (2010) A molecularly defined duplication set for the X chromosome of Drosophila melanogaster. Genetics 186: 1111-1125.

15. Chen ZX, Oliver B (2015) X Chromosome and Autosome Dosage Responses in Drosophila melanogaster Heads. G3 (Bethesda) 5: 1057-1063.

16. Holland AJ, Cleveland DW (2009) Boveri revisited: chromosomal instability, aneuploidy and tumorigenesis. Nat Rev Mol Cell Biol 10: 478-487.

17. Lee H, McManus CJ, Cho DY, Eaton M, Renda F, et al. (2014) DNA copy number evolution in Drosophila cell lines. Genome Biol 15: R70.

18. Zhang Y, Malone JH, Powell SK, Periwal V, Spana E, et al. (2010) Expression in aneuploid Drosophila S2 cells. PLoS Biol 8: e1000320.

19. Ui K, Ueda R, Miyake T (1987) Cell lines from imaginal discs of Drosophila melanogaster. In Vitro Cell Dev Biol 23: 707-711.

20. Hoskins RA, Carlson JW, Wan KH, Park S, Mendez I, et al. (2015) The Release 6 reference sequence of the Drosophila melanogaster genome. Genome Res 25: 445-458.

21. Salz HK, Erickson JW (2010) Sex determination in Drosophila: The view from the top. Fly (Austin) 4: 60-70.

22. Brennecke J, Hipfner DR, Stark A, Russell RB, Cohen SM (2003) bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in Drosophila. Cell 113: 25-36.

23. Kane DP, Shusterman M, Rong Y and McVey M (2012) Competition between replicative and translesion polymerases during homologous recombination repair in Drosophila. PLoS Genet 8: e1002659.

24. Brückner K, Kockel L, Duchek P, Luque CM, Rørth P, et al. (2004) The PDGF/VEGF receptor controls blood cell survival in Drosophila. Dev Cell 7: 73-84.

25. Cherbas L, Willingham A, Zhang D, Yang L, Zou Y, et al. (2011) The transcriptional diversity of 25 Drosophila cell lines. Genome Res 21: 301-314.

26. Tsay YF, Thompson JR, Rotenberg MO, Larkin JC and Woolford JL, Jr. (1988) Ribosomal protein synthesis is not regulated at the translational level in Saccharomyces cerevisiae: Balanced accumulation of ribosomal proteins L16 and rp59 is mediated by turnover of excess protein. Genes Dev 2: 664-76.

27. Dephoure N, Hwang S, O'Sullivan C, Dodgson SE, Gygi SP, et al. (2014) Quantitative proteomic analysis reveals posttranslational responses to aneuploidy in yeast. Elife 3: e03023.

28. Torres EM, Dephoure N, Panneerselvam A, Tucker CM, Whittaker CA, et al. (2010) Identification of aneuploidy-tolerating mutations. Cell 143: 71-83.

29. Stenberg P, Lundberg LE, Johansson AM, Rydén P, Svensson MJ, et al. (2009) Buffering of segmental and chromosomal aneuploidies in Drosophila melanogaster. PLoS Genet 5: e1000465.