Chapter 2

The Gene Balance Hypothesis: Dosage Effects in Plants

James A. Birchler and Reiner A. Veitia

Abstract

The concept of genomic balance traces to the early days of genetics. In recent years, studies of gene expression have found parallels to the classical phenotypic studies in that aneuploid changes have greater effects than whole genome changes. This has an explanation in terms of potential stoichiometric imbalances of the gene products encoded in the aneuploid regions. Studies of transcriptional factor mutations indicated that they tend to be haplo-insufficient as heterozygotes. Molecular evolution studies found that genes encoding members of macromolecular complexes were preferentially retained following polyploidy and underrepresented in copy number variants. In this review chapter, we synthesize these observations under the rubric of the Gene Balance Hypothesis.

Key words Aneuploidy, Ploidy, Copy number variants, Quantitative traits, Gene expression, Dosage compensation

1 Introduction to the Gene Balance Hypothesis

The Gene Balance Hypothesis posits that varying the stoichiometry of members of multi-subunit complexes will affect the function of the whole complex as a result of the topology, kinetics, and mode of assembly [1–5]. This principle applies to any type of macromolecular complex but perhaps its most critical implications are in the area of gene regulation, which is mediated in large part by oligomeric complexes. Because varying the stoichiometry of subunits has an effect, this will be manifested in a dosage response when the encoding gene is varied in copy number. Thus, gene regulatory systems tend to be dosage-dependent and thus will impact quantitative characteristics. The general idea of balance traces to the early days of genetics [6–8] but more recently a synthesis pulling together data from quantitative genetics, biophysics, molecular evolution, and studies of gene expression has been formulated. In this review article, we summarize the evidence for this synthesis and note some implications.

Charles Spillane and Peter C. McKeown (eds.), *Plant Epigenetics and Epigenomics: Methods and Protocols*,
Methods in Molecular Biology, vol. 1112, DOI 10.1007/978-1-62703-773-0_2,

One of the lines of evidence for the dosage balance concept is the classical observation that aneuploidy is generally more severe than ploidy changes. This concept was first formulated by Blakeslee and colleagues using the flowering plant, Datura [6, 7]. Trisomics were isolated for each of the 12 chromosomes. Each exhibited a characteristic phenotype. In comparison, a whole genome ploidy series was produced by chromosome doubling. The phenotypic changes in this case were not as dramatic as for the individual chromosome copy number modulations. This relationship has been found in many other plant and animal species over the subsequent decades [9, 10].

More recently, studies of gene expression modulation in aneuploidy and ploidy series showed that a greater number of modulations were found with aneuploidy than ploidy-level variations, in parallel to the phenotypic effects. There are two major types of modulations in aneuploids. One involves positive correlations with the varied chromosome that act in *trans* across the genome. The other type of *trans*-acting effect found was an inverse correlation of gene expression with the dosage of the varied chromosome [12]. These effects were found on the enzyme activity [11], protein [12], and messenger RNA levels [13]. In the latter study, the modulations caused by chromosomal dosage were within the direct and inverse correlative levels in the diploid embryo as well as in the triploid endosperm. In other words, the magnitude of genomic imbalance at the respective ploidy levels determined the magnitude of the effects. Changes of whole ploidy show fewer effects [14].

For the genes on the varied chromosome, it is generally assumed that a structural gene dosage effect will be produced with a change of chromosomal dosage. This is indeed the case for many gene products, but many cases of dosage compensation were also observed [11, 15, 16]. Dosage compensation is the phenomenon that the same amount of gene product is produced regardless of the chromosomal dosage. Examples of the *alcohol dehydrogenase 1* [11] and the *PRO* [12] genes located on the long arm of chromosome 1 exhibited the same amount of gene product in a 1-to-3 dosage series of this chromosome arm. In the case of *adh1*, the basis of the compensation was shown to be that an inverse dosage effect was operating on the locus in question which counteracted the structural gene dosage effect that might otherwise occur [12, 16]. Division of the long arm of chromosome 1 revealed a region that produces an inverse dosage effect upon *adh1* and that varying the dosage of a small region around *adh1* itself produced a gene dosage effect [16].

2 Gene Balance and Aneuploidy

The basis of the aneuploid effect was shown to be able to be reduced to the action of single genes [17, 18]. The leaky *white-apricot* allele of the *white* eye color gene in Drosophila was

used as a reporter to identify modifiers that would increase or decrease the amount of pigment when the new mutation was heterozygous. This situation would mimic a "monosomic" condition but on the single gene level. From over 2 decades of screening, 47 such modifiers were identified [18]. The majority of them acted negatively. Such a large number of modifiers are likely to result from the fact that many processes operate through regulatory hierarchies and/or through oligomeric regulatory factors. Each modifier would affect overlapping sets of genes.

This type of result has parallels in the genetics of quantitative traits. Quantitative trait loci are usually additive and multigenic [19] as are aneuploid syndromes [13]. Furthermore, they are controlled by many genes usually of small effect that are additive [20–25]. In other words, there is a dosage effect of the controlling alleles. Thus, there are similarities among the control of quantitative traits, the impact of multiple aneuploidies on the phenotype and the multigenic set of modifiers identified for a single phenotype [18].

Indeed, of the quantitative trait loci whose molecular nature has been elucidated, they are typically some type of regulatory factor. The first QTL cloned and molecularly characterized was *fw2.2*, which controls fruit weight in tomato [20]. When a transgenic dosage series was produced for this gene, a negative dosage effect on fruit weight was realized [26]. Among the collection of modifiers of the *white* eye color gene, those whose molecular nature is known consist of transcription factors, signal transduction components, and chromatin-modifying factors [18].

Another line of evidence in support of the Gene Balance Hypothesis is that haplo-insufficient genes in yeast and humans are enriched for proteins within complexes [27–30]. While these genes include the spectrum of those involved in macromolecular complexes, they include transcription factors and signal transduction components. The concept of balance was examined by overexpression of the same genes, which was found to be detrimental also [29]. However, co-over-expression was capable of correcting the fitness defects of interactors [29].

Further evidence comes from studies of molecular evolution. Throughout the plant kingdom [31–37], but also in yeast [38] and the animal kingdom [39, 40], there have been cycles of whole genome duplication (polyploidization) following by fractionation (diploidization). As genes are lost in the latter process, there is not a random distribution of the functional classes of genes that are retained [34–36]. Indeed, there is a preferential retention of genes whose products are involved with macromolecular complexes [34–36]. Included among these are transcription factors and signal transduction components. The implication is that if the stoichiometry of these gene products is important, deletion of one member of a duplicate pair might act like an aneuploid effect and be selected against, thus resulting in retention over longer periods of evolutionary time than other classes of genes.

The reciprocal result is found for segmental duplications and copy number variants. In this case there is an underrepresentation of genes whose products are involved in oligomeric complexes [34–36, 41–44]. Instead, genes encoding products that provide a selective advantage via greater quantity without balance defects are preferentially represented in partial genome duplications. This principle is reinforced by the realization that proteins that are increasingly under-wrapped (a measure of the reliance of a protein on binding partnerships to maintain structural integrity) are less likely to be correlated with gene duplicability [45]. Indeed, an inverse relationship between the extent of protein under-wrapping and gene family size has been demonstrated. Thus, gene duplication is unlikely to be tolerated if the structure of the corresponding protein requires substantial protein–protein stabilizing interactions unless the latter are co-duplicated or co-retained. Moreover, copy number polymorphisms in Drosophila [43] and humans [46] for genes with network centrality are significantly underrepresented.

Lastly, there are constraints on the tolerated variation of regulatory genes. In *Paramecium tetraurelia*, which has experienced three detectable whole genome duplication events as revealed by the genome sequence, there is evidence of purifying selection, based on Ka/Ks ratios, on the coding sequence of both members of a retained duplicate pair implying that dosage is important [39]. Because the conserved duplicate genes are likely to have kept the ancestral function, neofunctionalization cannot explain their retention. Instead, this result might be explained if mutations that upset the stoichiometric balance are selected against leaving the sequence signature of purifying selection. A similar conclusion can be drawn from an illuminating mutation accumulation experiment in *C. elegans* [47]. Mutations were allowed to accumulate and then patterns of gene expression were measured. Considerable variation for changes in the expression of individual target genes was revealed but there was conservation of the global patterns of gene expression suggesting that purifying selection was occurring for changes in the quantities of regulatory factors [47].

In a similar vein, studies of *cis* and *trans* variation in gene expression in general find that *cis* variation is typically of greater magnitude, although less pleiotropic, than *trans* variation but for any one modulation of a gene product there is a greater number of these more subtle changes [48–60]. This type of result would occur if target genes were not constrained for the type of *cis* regulatory variation that could be tolerated (probably within limits) but that the multiple regulatory genes have a constraint on the magnitude of variation that can be tolerated and maintained in populations.

3 Implications of the Gene Balance Hypothesis

The Gene Balance Hypothesis suggests that new mutations in regulatory genes of various types will likely produce a semidominant dosage effect to some degree and to have a (subtle) phenotypic effect. The consequence of this is that new mutations will be available for selection, be that either purifying or adaptive. Mutations that are completely recessive are not available for selection. They may become lost in a population or alternatively, only in a small population would drift and inbreeding make them homozygous and thus responsive to selective forces. The implication is that there is a greater availability for adaptive selection for regulatory genes than for others that do not exhibit dosage stoichiometries.

While new dosage-sensitive mutations would be readily available for selection, it is likely that this property of regulatory genes would also work to maintain the status quo in regulatory processes due to purifying selection against detrimental mutations perturbing the stoichiometric balance. It is generally considered that purifying selection is more common than adaptive selection, but once adaptations occur, purifying selection would maintain them.

Another principle suggested by the results described above is that regulatory changes would have an impact on evolution in subtle increments but that many genes can contribute to any one trajectory. The evidence, noted above, from the study of modifiers of the *white* eye color gene and from quantitative trait multigenic control, illustrates that many genes can impact a single phenotypic characteristic. The data from retention of duplicate genes encoding macromolecular complexes following ancient polyploidization events and their underrepresentation in copy number variants suggests that the magnitude of tolerable dosage effect is narrow and well below a twofold range. Thus, the standing variation in regulatory processes is likely to be quite subtle but would be contributed by many genes. Thus, the control of quantitative traits will be determined by many genes each with a small effect.

Future studies involved with the Gene Balance Hypothesis might focus on the effect of stoichiometric changes of individual subunits of macromolecular complexes and how these changes alter the function of the whole complex. Some possibilities might be that the kinetics of assembly lead to unproductive partial complexes [3] or that targeted degradation of unused subunits may alleviate or, on the contrary, enhance dosage effects [5]. Another question involves how new balances are achieved during evolution. As noted above, *cis* variation will accumulate in target genes and eventually will be in conflict with the *trans* regulatory system if critical target genes change their expression. The evolutionary evidence from preferential retention following polyploidization suggests that there is resistance to altered balance but ultimately

this would change and elucidating the processes by which this occurs would be illuminating. Further, microRNAs are known to impact gene expression in a dosage-sensitive manner and so they are likely to play a role in gene balance mechanisms but basically nothing is known about this possibility at present. Lastly, it is of interest to decipher whether issues of regulatory gene balance play any role in speciation [3]. If new balances are indeed achieved in separate evolutionary lineages, then their combination in hybrids might prevent gene flow by causing reduced fitness at some level.

Acknowledgements

Research in our labs is supported by National Institutes of Health grant RO1GM068042-05 and National Science Foundation grant DBI 0733857 Plant Genome.

References

1. Birchler JA, Riddle NC, Auger DL, Veitia RA (2005) Dosage balance in gene regulation: biological implications. Trends Genet 21: 219–226
2. Birchler JA, Veitia RA (2007) The gene balance hypothesis: from classical genetics to modern genomics. Plant Cell 19:395–402
3. Birchler JA, Veitia RA (2010) The gene balance hypothesis: implications for gene regulation, quantitative traits and evolution. New Phytol 186:54–62
4. Birchler JA, Yao H, Chudalayandi S (2007) Biological consequences of dosage dependent gene regulatory systems. Biochim Biophys Acta 1769:422–428
5. Veitia RA, Bottani S, Birchler JA (2008) Cellular reactions to gene dosage imbalance: genomic, transcriptomic and proteomic effects. Trends Genet 24:390–397
6. Blakeslee AF, Belling J, Farnham ME (1920) Chromosomal duplication and Mendelian phenomena in Datura mutants. Science 52: 388–390
7. Blakeslee AF (1934) New Jimson weeds from old chromosomes. J Hered 24:80–108
8. Bridges CB (1925) Sex in relation to chromosomes and genes. Am Nat 59:127–137
9. Lindsley DL, Sandler L, Baker BS, Carpenter AT, Denell RE, Hall JC, Jacobs PA, Miklos GL, Davis BK, Gethman RC et al (1972) Segmental aneuploidy and the genetic gross structure of the Drosophila genome. Genetics 71:157–184
10. Bond DJ, Chandley AC (1983) Aneuploidy. Oxford University Press, Oxford
11. Birchler JA (1979) A study of enzyme activities in a dosage series of the long arm of chromosome one in maize. Genetics 92:1211–1229
12. Birchler JA, Newton KJ (1981) Modulation of protein levels in chromosomal dosage series of maize: the biochemical basis of aneuploid syndromes. Genetics 99:247–266
13. Guo M, Birchler JA (1994) Trans-acting dosage effects on the expression of model gene systems in maize aneuploids. Science 266:1999–2002
14. Guo M, Davis D, Birchler JA (1996) Dosage effects on gene expression in a maize ploidy series. Genetics 142:1349–1355
15. Birchler JA, Hiebert JC, Paigen K (1990) Analysis of autosomal dosage compensation involving the alcohol dehydrogenase locus in *Drosophila melanogaster*. Genetics 124:677–686
16. Birchler JA (1981) The genetic basis of dosage compensation of alcohol dehydrogenase-1 in maize. Genetics 97:625–637
17. Rabinow L, Nguyen-Huynh AT, Birchler JA (1991) A trans-acting regulatory gene that inversely affects the expression of the white, brown and scarlet loci in *Drosophila melanogaster*. Genetics 129:463–480
18. Birchler JA, Bhadra U, Pal Bhadra M, Auger DL (2001) Dosage dependent gene regulation in multicellular eukaryotes: implications for dosage compensation, aneuploid syndromes and quantitative traits. Dev Biol 234:275–288
19. Tanksley SD (1993) Mapping polygenes. Annu Rev Genet 27:205–233
20. Frary A, Nesbitt RC, Frary A, Grandillo S, van der Knaap E, Cong B, Liu J, Meller J, Elber R,

Alpert KB, Tanksley SD (2000) *fw2.2*: a quantitative trait locus key to the evolution of tomato fruit size. Science 289:85–88
21. Cong B, Liu J, Tanksley SD (2002) Natural alleles at a tomato fruit size quantitative trait locus differ by heterochronic regulatory mutations. Proc Natl Acad Sci U S A 99: 13606–13611
22. Cong B, Barrero LS, Tanksley SD (2008) Regulatory change in YABBY-like transcription factor led to evolution of extreme fruit size during tomato domestication. Nat Genet 40:800–804
23. Liu J, Van Eck J, Cong B, Tanksley SD (2002) A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. Proc Natl Acad Sci U S A 99:13302–13306
24. Burke JM, Tang S, Knapp SJ, Rieseberg LH (2002) Genetic analysis of sunflower domestication. Genetics 161:1257–1267
25. Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ et al (2009) The genetic architecture of maize flowering time. Science 325:714–718
26. Liu J, Cong B, Tanksley SD (2003) Generation and analysis of an artificial gene dosage series in tomato to study the mechanisms by which the cloned quantitative trait locus *fw2.2* controls fruit size. Plant Physiol 132:292–299
27. Veitia RA (2002) Exploring the etiology of haploinsufficiency. Bioessays 24:175–184
28. Seidman JG, Seidman C (2002) Transcription factor haploinsufficiency: when half a loaf is not enough. J Clin Invest 109:451–455
29. Papp B, Pal C, Hurst LD (2003) Dosage sensitivity and the evolution of gene families in yeast. Nature 424:194–197
30. Kondrashov FA, Koonin EV (2004) A common framework for understanding the origin of genetic dominance and evolutionary fates of gene duplications. Trends Genet 20:287–290
31. Simillion C, Vandepoele K, Montagu MC, Zabeau M, Van de Peer Y (2002) The hidden duplication past of *Arabidopsis thaliana*. Proc Natl Acad Sci U S A 99:13627–13632
32. Bowers JE, Chapman BA, Rong J, Paterson AH (2003) Unraveling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. Nature 422: 433–438
33. Chapman BA, Bowers JE, Feltus FA, Paterson AH (2006) Buffering of crucial functions by paleologous duplicated genes may contribute cyclicality to angiosperm genome duplication. Proc Natl Acad Sci U S A 103:2730–2735
34. Maere S, DeBodt S, Raes J, Casneuf T, Van Montagu M, Kuiper M, Van de Peer Y (2005) Modeling gene and genome duplications in eukaryotes. Proc Natl Acad Sci U S A 102: 5454–5459
35. Blanc G, Wolfe KH (2004) Functional divergence of duplicated genes formed by polyploidy during Arabidopsis evolution. Plant Cell 16:1679–1691
36. Freeling M, Thomas BC (2006) Gene-balanced duplications, like tetraploidy, provide predictable drive to increase morphological complexity. Genome Res 16:805–814
37. Barker MS, Kane NC, Matvienko M, Kozik A, Michelmore RW, Knapp SJ, Rieseberg LH (2008) Multiple paleopolyploidizations during the evolution of the compositae reveal parallel patterns of duplicate gene retention after millions of years. Mol Biol Evol 25:2445–2455
38. Wolfe KH, Shields DC (1997) Molecular evidence for an ancient duplication of the entire yeast genome. Nature 387:708–713
39. Aury J-M, Jaillon O, Duret L, Jubin C, Porcel BM, Segurens B, Daubin V, Anthouard V, Aiach N, Arnaiz O et al (2006) Global trends of whole-genome duplications revealed by the ciliate *Paramecium tetraurelia*. Nature 444: 171–178
40. Blomme T, Vandepoele K, De Bodt S, Simillion C, Maere S, Van de Peer Y (2006) The gain and loss of genes during 600 million years of vertebrate evolution. Genome Biol 7:R43
41. Freeling M, Lyons E, Pedersen B, Alam M, Ming R, Lisch D (2008) Many or most genes in Arabidopsis transposed after the origin of the order Brassicales. Genome Res 18:1924–1937
42. Davis JC, Petrov DA (2005) Do disparate mechanisms of duplication add similar genes to the genome? Trends Genet 21:548–551
43. Dopman EB, Hartl DL (2007) A portrait of copy-number polymorphism in *Drosophila melanogaster*. Proc Natl Acad Sci U S A 104: 19920–19925
44. Hakes L, Pinney JW, Lovell SC, Oliver SG, Robertson DL (2007) All duplicates are not equal: the difference between small-scale and genome duplications. Genome Biol 8:R209
45. Liang H, Rogale-Plazonic K, Chen J, Li WH, Fernandez A (2008) Protein under-wrapping causes dosage sensitivity and decreases gene duplicability. PLoS Genet 4:e11
46. Schuster-Bockler B, Conrad D, Bateman A (2010) Dosage sensitivity shapes the evolution of copy-number varied regions. PLoS One 5:e9474
47. Denver DR, Morris K, Streelman JT, Kim SK, Lynch M, Thomas WK (2005) The transcriptional consequences of mutation and natural selection in *Caenorhabditis elegans*. Nat Genet 37:544–548
48. Schadt EE, Monks SA, Drake TA, Lusis AJ, Che N, Colinayo V, Ruff TG, Milligan SB, Lamb JR, Cavet G, Linsley PS, Mao M, Stoughton RB, Friend SH (2003) Genetics of gene expression surveyed in maize, mouse and man. Nature 422:297–302

49. Yvert G, Brem RB, Whittle J, Akey JM, Foss E, Smith EN, Mackelprang R, Kruglyak L (2003) Trans-acting regulatory variation in *Saccharomyces cerevisiae* and the role of transcripton factors. Nat Genet 35:57–64
50. Wittkopp PJ, Haerum BK, Clark AG (2004) Evolutionary changes in cis and trans gene regulation. Nature 430:85–88
51. Wayne ML, Pan Y-J, Nuzhdin SV, McIntyre LM (2004) Additivity and trans-acting effects on gene expression in male *Drosophila simulans*. Genetics 168:1413–1420
52. Morley M, Molony CM, Weber TM, Devlin JL, Ewens KG, Spielman RS, Cheung VG (2004) Genetic analysis of genome-wide variation in human gene expression. Nature 430:743–747
53. Hughes KA, Ayroles JF, Reedy MM, Drnevich JM, Rowe KC, Ruedi EA, Caceres CE, Paige KN (2006) Segregating variation in the transcriptome: cis regulation and additivity of effects. Genetics 173:1347–1355
54. West MAL, Kim K, Kliebenstein DJ, van Leeuwen H, Michelmore RW, Doerge RW, St. Clair DA (2007) Global eQTL mapping reveals the complex genetic architecture of transcript-level variation in Arabidopsis. Genetics 175:1441–1450
55. Wang D, Sung H-M, Wang T-Y, Huang C-J, Yang P, Chang T, Wang Y-C, Tseng D-L, Wu J-P, Lee T-C, Shih M-C, Li W-H (2007) Expression evolution in yeast genes of single-input modules is mainly due to changes in trans-acting factors. Genome Res 17:1161–1169
56. Petretto E, Mangion J, Dickens NJ, Cook SSA, Kumaran MK, Lu M, Fischer J, Maatz H, Kren V, Pravenec M, Hubner N, Aitman TJ (2006) Heritability and tissue specificity of expression quantitative trait loci. PLoS Genet 2:e172
57. Grieve IC, Dickens NJ, Pravenec M, Kren V, Hubner N, Cook SA, Ailtman TJ, Petretto E, Mangion J (2008) Genome-wide co-expression analysis in multiple tissues. PLoS One 3:e4033
58. Lemos B, Araripe LO, Fontanilla P, Hartl DL (2008) Dominance and the evolutionary accumulation of cis- and trans-effects on gene expression. Proc Natl Acad Sci U S A 105:1813–1822
59. L'Hote D, Serres C, Veitia RA, Montagutelli X, Oulmouden A, Vaiman D (2008) Gene expression regulation in the context of mouse interspecific mosaic genomes. Genome Biol 9:R133
60. Tirosh I, Reikhav S, Levy AA, Barkai N (2009) A yeast hybrid provides insight into the evolution of gene expression regulation. Science 324:659–662

http://www.springer.com/978-1-62703-772-3

Plant Epigenetics and Epigenomics
Methods and Protocols
Spillane, C.; McKeown, P.C. (Eds.)
2014, X, 245 p. 39 illus., 12 illus. in color., Hardcover
ISBN: 978-1-62703-772-3
A product of Humana Press