Available online at www.sciencedirect.com

SciVerse ScienceDirect

Current Opinion in Plant Biology

Insights from paleogenomic and population studies into the consequences of dosage sensitive gene expression in plants

James A Birchler

Classical studies of plant phenotypes of individuals with whole or partial genome dosage changes led to the concept of genomic balance. Subsequent studies of gene expression in ploidy and aneuploidy series showed a greater number of modulations in aneuploid plants than with whole genome changes leading to the idea that gene expression processes were modulated by stoichiometric changes of interacting regulatory factors. Recent studies of genomic sequences and copy number variants in populations reveal different fates of duplicate genes depending on whole genome or segmental duplication. Following polyploidy formation, members of macromolecular complexes persist in the evolutionary lineage longer than random genes and a complementary pattern is found for segmental duplications in that there is an underrepresentation of members of macromolecular complexes. These and other studies described suggest there are negative fitness consequences when an imbalance occurs for members of macromolecular complexes including regulatory functions.

Address
Division of Biological Sciences, University of Missouri, 311 Tucker Hall, Columbia, MO 65211, United States

Corresponding author: Birchler, James A (BirchlerJ@Missouri.edu)

Current Opinion in Plant Biology 2012, **15**:544–548

This review comes from a themed issue on **Cell signaling and gene regulation**

Edited by **Jerzy Paszkowski** and **Robert A. Martienssen**

For a complete overview see the Issue and the Editorial

Available online 30th August 2012

1369-5266/$ – see front matter,

http://dx.doi.org/10.1016/j.pbi.2012.08.005

Introduction

In recent years, several studies in plants have revealed a cyclical pattern of polyploidy formation during evolution with subsequent progression toward a more diploid state [1,2]. With the loss of genes over time, there are trends of gene retention that indicate a nonrandom elimination based on gene function [1]. In particular, the classes of genes that have greater retention times tend to be involved in macromolecular complexes and highly connected in the interactome. By contrast, examination of the classes of genes that tend to be retained as segmental duplications reveal a generalized complementary pattern, namely that highly connected genes are underrepresented. In this review, this pattern of the differential fate of gene copies depending on the mode of duplication will be discussed in relation to gene expression studies in polyploidy and aneuploidy series and the implications for the evolution of patterns of gene expression.

Gene expression in whole genome and segmental dosage series

Gene expression patterns in ploidy series have been determined in many species of plants. Depending on the particular study, haploid, diploid, triploid and tetraploid individuals were examined with one, two, three and four copies of the whole genome, respectively [3–9]. With such a whole genome dosage series, the cell size typically increases proportionally with the increase in the number of genomes present [9]. The overall stature of the plants in a ploidy series will depend in large measure on the degree of heterozygosity [7,9,10]. Above the diploid level, many plant species decline in stature with increasing ploidy when maximally homozygous but increase in stature with maximal heterozygosity. A generalization across studies is that the relative expression of most genes does not change in a major way. However, with the realization of a change in the cell size with ploidy, there is an apparent proportional increase in expression per cell with the number of copies of the genome present.

There are fewer studies of segmental changes in dosage in plants. Some decades ago, molecular studies of trisomic plants were conducted to examine issues of gene dosage effects or to use molecular markers to classify aneuploid individuals [3,11–17]. In these studies there were numerous changes in gene expression that were noted. Most studies examined trisomics because of their prevalence and ease of manipulation. The most typical changes in expression found were reductions in expression relative to the normal diploid. Gene products encoded on the varied chromosome could exhibit a gene dosage effect or dosage compensation. Studies using maize allowed the examination of monosomic individuals with only one copy of selected chromosome arms [3,14]. These studies revealed that there were often negative correlations of gene expression through the one, two and three dose series, although reductions in expression were also common in the monosomics. Although not studied in depth, increasing amounts of the genome varied in dosage together caused more extensive but not necessarily additive changes in gene expression [3]. These types of dosage modulations could be reduced to the action of

single genes, which include transcription factors, signal transduction components and chromatin proteins [18]. When the results from aneuploidy and ploidy were compared, overall there were more modulations of expression from aneuploidy than for ploidy changes. This comparison led to the idea that changing the stoichiometry of the regulatory factors was responsible for the trans-acting dosage effects [3,18]. Below, we will relate this difference between whole and segmental genomic effects to the evolutionary patterns of gene retention.

Reoccurring polyploidy events in plant evolution versus copy number variants

The sequence of the baker's yeast genome revealed the presence of an ancient genome duplication that would account for numerous duplicate copies of genes [19]. Shortly thereafter, a similar phenomenon was recognized for Arabidopsis [20–24]. Based on the patterns of synonymous and non-synonymous base substitutions in the gene pairs, the whole genome duplications could be classified into alpha, beta and gamma events going back in time from the present [23]. The alpha event has a larger number of genes retained than beta, which in turn has a greater number than gamma. (Eventually two earlier additional events were documented [2].) When the function of the genes retained was examined, those involved with molecular complexes such as the ribosome, proteasome, transcription factors and members of signal transduction cascades tended to be overrepresented [22,24]. With regard to regulatory factors, the classes of genes that are preferentially retained for longer evolutionary periods are the same as those that exhibit dosage effects similar to the aneuploidy modulations [23–25]. As sequence data became available from other species, related trends were observed for independent polyploidization events and gene number reductions [1,26].

These results were compared to the spectrum of gene functions that exhibited segmental copy number duplications in several plant species [27,28,29•,30]. Some classes of genes are overrepresented as multicopy, for example, those involved with disease resistance. However, interestingly, the generalized trends of gene ontology that are retained following whole genome duplication are underrepresented in the segmental duplications. Thus, there is a complementary pattern of retained genes following whole genome versus those that persist as segmental duplications.

Gene balance hypothesis

An explanation for the complementary pattern of duplicates in WGD and CNV is that certain gene products are in a particular stoichiometry or balance, otherwise negative fitness consequences ensue [25,27,32,33••]. Deletion of certain classes of genes from WGD would be detrimental as well as the addition of them via CNV because both dosage changes would alter the stoichiometry of connected genes. The idea of 'balance' traces back to the early days of genetics which originated from the fact that changes in dosage of a single chromosome have more consequences for the phenotype than changes in ploidy [34] (Figure 1). This concept was extended to explain the greater number of modulations of gene expression in aneuploids compared to euploids in particular with the suggestion that the stoichiometry of regulatory factors involved in multi-subunit complex impacted gene expression [3,14,18]. Studies in baker's yeast demonstrated that the change in dosage of genes encoding members of macromolecular complexes negatively affected fitness with increasing connections in the interactome [35]. Biophysical considerations of protein-protein interactions support the same conclusion [36–38]. Thus, the idea is that the balance or stoichiometry of members of macromolecular complexes affects the function of the whole and that these changes have a negative fitness consequence.

The concept of balance can therefore be traced through the classical phenotypic observations, the gene

Figure 1

Current Opinion in Plant Biology

Genomic balance effects on the phenotype of maize. Plants are shown from left to right: haploid, haploid plus the long arm of chromosome 1 (1L), normal diploid and diploid plus 1L. All individuals are in the inbred background of W22. Note that the addition of 1L to a diploid genotype has a much less detrimental effect than addition to a haploid. The doubling of the whole genome from haploid to diploid compared to doubling only 1L in an otherwise haploid illustrates the negative impact of changing the stoichiometry or balance of a portion of the genome. This comparison also illustrates that a two fold dosage change in only a portion of the genome can severely affect plant stature and much more than changing the whole genome by the same magnitude. A meter stick is included for scale. Photo by Fangpu Han.

expression modulations in aneuploids and polyploids and the differential retention of genes following WGD or CNV [32,39]. Because the gene dosage parameters involved with the latter are well defined, they provide several implications for the constraints and consequences of dosage modulation of gene expression. First, for individual genes there is likely to be a reasonably linear amount of expression between the copy number of the gene and the amount of the gene product. While there are examples in which this is clearly not the case, a study in yeast indicates that the amount of gene product is reasonably well correlated with the dosage of the vast majority of genes [40]. However, when we consider the dosage of regulatory genes, they have a trans-acting dosage effect that will modulate their target loci. Indeed, this fact is probably a contributing factor to the negative fitness associated with varying the dosage of the regulator.

Secondly, the fact that many types of transcription factors and signal transduction components are retained following whole genome duplication suggests that a mere change in quantity of 25% will be selected against. This is the case because there are four copies of a gene present after the formation of a tetraploid and deletion of a single copy back to three copies is apparently detrimental. An extension of this realization is that natural variation in regulatory factors even in diploids will be constrained so that there will be very little change in magnitude of expression for the variants that can persist in a population. Indeed, studies of expression quantitative trait loci for standard target genes reveals cis-acting variation that is found less frequently but of greater magnitude than the trans-acting regulatory variation that is multigenic and of lesser magnitude (e.g. [41]) if at all detectable. This pattern of variation is consistent with the balance concept.

Another potential explanation for gene retention is that the duplicate genes have changed function such that different aspects of the progenitor are partitioned between the two new copies—the process of subfunctionalization. This hypothesis, however, cannot account for the complementary pattern of retention following segmental duplication because one would expect equal opportunity for divergence regardless of the mode of duplication [1]. Moreover, as noted above, most of the retained duplicates from WGD eventually are lost as well. Subfunctionalization involves mutational changes to subdivide gene functions [31]. Thus, for this hypothesis to account for the eventual loss of the evolutionarily temporally retained genes would require back mutation of one copy of the pair to re-establish the original functions of the progenitor copy. This scenario is highly improbable. On the contrary, the retention of regulatory factors in the evolutionary lineage for longer periods of time than random genes would provide greater opportunity for some cases of subfunctionalization to occur. Thus, the combination of balance based retention and subfunctionalization provides an extended opportunity for the evolution of new genes via WGD.

Implications of dosage effects and evolutionary data

The relationship of patterns of gene expression to the phenotype is a complex one. Nevertheless, there is clearly a connection. The facts that regulatory genes have a trans-acting dosage effect [18] and that the evolutionary results indicate selection consequences of dosage changes illustrate that the control of quantitative phenotypic traits will be affected by multiple loci that have dosage effects when the allelic variation affects the quantity of gene expression. Indeed, classical studies of quantitative traits have indicated that crosses between parents of extreme phenotypes produce progeny with an intermediate phenotype illustrating a cumulative semidominant effect [42,43]. In natural populations, the variation for phenotypic characteristics will be affected by multiple loci that are constrained in the magnitude of each as a consequence of the detrimental effects of changing the stoichiometry of the regulatory factors involved.

Differential genomic expression in polyploids

The 'paleogenomic' studies also illustrate another interesting feature of dosage effects in polyploids. As noted above, the joining of two genomes together provides four copies of every gene. When the two progenitor genomes are slightly different, the same dosage issues are involved but chromosome pairing parameters in meiosis are different such that the two genomes are fixed and cannot recombine. Interestingly, the average gene expression from these two genomes is often recognizably different [44,45,46••,47,48]. This difference in expression appears to have an impact on the fractionation of genes back toward the diploid level. The genome with lesser expression is preferentially the one with the greater number of gene losses [44,49••,50••]. This difference has been attributed to a lesser impact on the balance of gene products when genes from the more weakly expressed genome are deleted [44,49••].

Chromosomal balance is selected in newly formed polyploids

Related results come from laboratory and population studies of newly formed polyploids [51••,52••]. These often exhibit variation in chromosome number in the early generations. However, in successive generations the related chromosomes from the two progenitor genomes can shift in copy number so that the total remains the same. Thus, for the gene repertoire the balance across the genome is restored. These results illustrate from another type of data that changes in gene dosage of merely 25% can have detrimental effects and will be selected against.

What next?

The evolutionary and population genomic studies have produced a framework of how gene dosage affects fitness and the types of genes that are involved. They point to the need to understand on the molecular genetic level how changing the dosage of regulatory genes impacts their target loci and how dosage interactions among regulators of various types does or does not affect gene expression patterns. They also point to the need for further study of the biophysical parameters of protein-protein interactions among the components of macromolecular complexes in terms of how and why changes in the stoichiometry of the members have the consequences attributed to them from the evolutionary and aneuploidy studies. It is also an interesting question whether, because gametophytes would have greater imbalance with gene copy number change, gametophytically expressed genes have a greater effect on the evolutionary trajectory of WGD and CNV. Further topics in need of examination within the context of balance phenomena include the role of microRNAs that modulate transcription factor levels posttranscriptionally and the extent to which epigenetic modifications play a role in regulatory stoichiometry.

Acknowledgements

Research was supported by National Science Foundation grant DBI 0733857.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Freeling M: **Bias in plant gene content following different sorts of duplication: tandem, whole-genome, segmental, or by transposition**. *Annu Rev Plant Biol* 2009, **60**:433-453.
2. Jiao Y, Wickett NJ, Ayyampalayan S, Chanderball AS, Landherr L, Ralph PE, Tomsho JP, Hu Y, Linag H, Soltis PS *et al.*: **Ancestral polyploidy in seed plants and angiosperms**. *Nature* 2011, **473**:97-100.
3. Birchler JA, Newton KJ: **Modulation of protein levels in chromosomal dosage series of maize: the biochemical basis of aneuploid syndromes**. *Genetics* 1981, **99**:247-266.
4. Guo M, Davis D, Birchler JA: **Dosage effects on gene expression in a maize ploidy series**. *Genetics* 1996, **142**:1349-1355.
5. Martelotto LG, Ortiz JPA, Stein J, Espinoza F, Quarin CL, Pessino SC: **A comprehensive analysis of gene expression alterations in a newly synthesized *Paspalum notatum* autotetraploid**. *Plant Sci* 2005, **169**:211-220.
6. Stupar RM, Bhaskar PB, Yandell BS, Rensink WA, Hart AL, Ouyang S, Veilleux RE, Busse JS, Erhardt RJ, Buell CR, Jiang J: **Phenotypic and transcriptomic changes associated with potato autopolyploidization**. *Genetics* 2007, **176**:2055-2067.
7. Riddle NC, Jiang H, An L, Doerge RW, Birchler JA: **Gene expression analysis at the intersection of ploidy and hybridity in maize**. *Theor Appl Genet* 2010, **120**:341-353.
8. Yu Z, Haberer G, Matthes M, Rattei T, Mayer KF, Gierl A, Torres-Ruiz RA: **Impact of natural genetic variation on the transcriptome of autotetraploid *Arabidopsis thaliana***. *Proc Natl Acad Sci USA* 2010, **107**:17809-17814.
9. Yao H, Akio Kato A, Mooney B, Birchler JA: **Phenotypic and gene expression analysis of a ploidy series of maize inbred Oh43**. *Plant Mol Biol* 2011, **75**:237-251.
10. d'Erfuth I, Jolivet S, Froger N, Catrice O, Novatchkova M, Mercier R: **Turning meiosis into mitosis**. *PLoS Biol* 2009, **7**:e1000124.
11. McDaniel RG, Ramage RT: **Genetics of primary trisomic series in barley: identification by protein identification**. *Can J Genet Cytol* 1970, **12**:155-167.
12. Smith HH, Conklin ME: **Effects of gene dosage on peroxidase isozymes in *Datura stramonium* trisomics**. In *Isozymes, Vol 3. Developmental Biology.* Edited by Markert CL. New York: Academic Press; 1975:603-618.
13. Birchler JA: **A study of enzyme activities in a dosage series of the long arm of chromosome one in maize**. *Genetics* 1979, **92**:1211-1229.
14. Guo M, Birchler JA: **Trans-acting dosage effects on the expression of model gene systems in maize aneuploids**. *Science* 1994, **266**:1999-2002.
15. Makarevitch I, Phillips RL, Springer NM: **Profiling expression changes caused by a segmental aneuploidy in maize**. *BMC Genomics* 2008, **9**:7.
16. Huettel B, Kreil DP, Matzke M, Matzke AJM: **Effects of aneuploidy on genome structure, expression and interphase organization in *Arabidopsis thaliana***. *PLoS Genet* 2008, **4**:e1000226.
17. Birchler JA: **Reflections on studies of gene expression in aneuploids**. *Biochem J* 2010, **426**:119-123.
18. Birchler JA, Bhadra U, Pal Bhadra M, Auger DL: **Dosage dependent gene regulation in multicellular eukaryotes: implications for dosage compensation, aneuploid syndromes and quantitative traits**. *Dev Biol* 2001, **234**:275-288.
19. Wolfe KH, Shields DC: **Molecular evidence for an ancient duplication of the entire yeast genome**. *Nature* 1997, **387**:708-713.
20. Simillion C, Vandepoele K, Montagu MC, Zabeau M, Van de Peer Y: **The hidden duplication past of *Arabidopsis thaliana***. *Proc Natl Acad Sci USA* 2002, **99**:13627-13632.
21. Bowers JE, Chapman BA, Rong J, Paterson AH: **Unraveling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events**. *Nature* 2003, **422**:433-438.
22. Blanc G, Wolfe KH: **Functional divergence of duplicated genes formed by polyploidy during Arabidopsis evolution**. *Plant Cell* 2004, **16**:1679-1691.
23. Maere S, DeBodt S, Raes J, Casneuf T, Van Montagu M, Kuiper M, Van de Peer Y: **Modeling gene and genome duplications in eukaryotes**. *Proc Natl Acad Sci USA* 2005, **102**:5454-5459.
24. Freeling M, Thomas BC: **Gene-balanced duplications, like tetraploidy, provide predictable drive to increase morphological complexity**. *Genome Res* 2006, **16**:805-814.
25. Birchler JA, Riddle NC, Auger DL, Veitia RA: **Dosage balance in gene regulation: biological implications**. *Trends Genet* 2005, **21**:219-226.
26. Barker MS, Kane NC, Matvienko M, Kozik A, Michelmore RW, Knapp SJ, Rieseberg LH: **Multiple paleopolyploidizations during the evolution of the compositae reveal parallel patterns of duplicate gene retention after millions of years**. *Mol Biol Evol* 2008, **25**:2445-2455.
27. Freeling M, Lyons E, Pedersen B, Alam M, Ming R, Lisch D: **Many or most genes in Arabidopsis transposed after the origin of the order Brassicales**. *Genome Res* 2008, **18**:1924-1937.
28. Rosado A, Raikhel NV: **Appication of the gene dosage balance hypothesis to auxin related ribosomal mutants in Arabidopsis**. *Plant Signal Behav* 2010, **5**:450-452.
29. • Coate JE, Schlueter JA, Whaley AM, Doyle JJ: **Comparative evolution of photosynthetic genes in response to polyploidy and nonpolyploid duplication**. *Plant Physiol* 2011, **155**:2081-2095.

Provides an excellent comparative analysis of WGD and CNV for genes involved with photosynthetic complexes.

30. Rodgers-Melnick E, Shrinivasrao PM, Dharmawardhana P, Slavov GT, Crasta OR, Strauss SH, Brunner AM, DiFazio SP: **Contrasting patterns of evolution following whole genome versus tandem duplication events in Populus**. *Genome Res* 2012, **22**:95-105.

31. Lynch M, Conery JS: **The evolutionary fate and consequences of duplicate genes**. *Science* 2000, **290**:1151-1155.

32. Veitia RA, Bottani S, Birchler JA: **Cellular reactions to gene dosage imbalance: genomic, transcriptomic and proteomic effects**. *Trends Genet* 2008, **24**:390-397.

33. •• Yang L, Gaut BS: **Factors that contribute to variation in evolutionary rate among Arabidopsis genes**. *Mol Biol Evol* 2011, **28**:2359-2369.
Analysis concluding that retained duplicates following tetraploid formation are subject to purifying selection suggesting that changes in gene copy or equivalent expression of 25% have fitness consequences for the balanced interacting genes.

34. Blakeslee AF: **New Jimson weeds from old chromosomes**. *J Hered* 1934, **24**:80-108.

35. Papp B, Pal C, Hurst LD: **Dosage sensitivity and the evolution of gene families in yeast**. *Nature* 2003, **424**:194-197.

36. Veitia RA: **Exploring the etiology of haploinsufficiency**. *Bioessays* 2002, **24**:175-184.

37. Liang H, Rogale-Plazonic K, Chen J, Li WH, Fernandez A: **Protein under-wrapping causes dosage sensitivity and decreases gene duplicability**. *PLoS Genet* 2008, **4**:e11.

38. Schuster-Bockler B, Conrad D, Bateman A: **Dosage sensitivity shapes the evolution of copy-number varied regions**. *PLoS ONE* 2010, **5**:e9474.

39. Birchler JA, Veitia RA: **The gene balance hypothesis: implications for gene regulation, quantitative traits and evolution**. *New Phytol* 2010, **186**:54-62.

40. Springer M, Weissman JS, Kirshner MW: **A general lack of compensation for gene dosage in yeast**. *Mol Syst Biol* 2010, **6**:368.

41. West MAL, Kim K, Kliebenstein DJ, van Leeuwen H, Michelmore RW, Doerge RW, St. Clair DA: **Global eQTL mapping reveals the complex genetic architecture of transcript-level variation in Arabidopsis**. *Genetics* 2007, **175**:1441-1450.

42. Tanksley SD: **Mapping polygenes**. *Annu Rev Genet* 1993, **27**: 205-233.

43. Liu J, Cong B, Tanksley SD: **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* 2003, **132**:292-299.

44. Thomas BC, Pedersen B, Freeling M: **Following tetraploidy in an Arabidopsis ancestor, genes were removed preferentially from one homeolog leaving clusters enriched in dose-sensitive genes**. *Genome Res* 2006, **16**:934-946.

45. Wang J, Tian L, Lee HS, Wei NE, Jiang H, Watson B, Madlung A, Osborn TC, Doerge RW, Comai L, Chen ZJ: **Genomewide nonadditive gene regulation in Arabidopsis allotetraploids**. *Genetics* 2006, **172**:507-517.

46. •• Flagel LE, Wendel JF: **Evolutionary rate variation, genomic dominance and duplicate gene expression evolution during allotetraploid cotton speciation**. *New Phytol* 2010, **186**:184-193.
Together with [45], a detailed analysis of the phenomenon of genomic dominance in allotetraploids.

47. Buggs RJ, Chamala S, Wu W, Gao L, May GD, Schnalbe PS, Soltis PS, Barbazuk WB: **Characterization of duplicate gene evolution in the recent natural allopolyploid *Tragopogon miscellus* by next-generation sequencing and Sequenom iPLEX NassARRAY genotyping**. *Mol Ecol* 2010, **19**:132-146.

48. Chang PL, Dilkes BP, McMahon M, Comai L, Nuzhdin SV: **Homoeolog-specific retention and use in allotetraploid Arabidopsis suecica depends on parent of origin and network partners**. *Genome Biol* 2010, **11**:R125.

49. •• Schnable JC, Springer NM, Freeling M: **Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss**. *Proc Natl Acad Sci USA* 2011, **108**: 4069-4074.
This study illustrates the connection between genome dominance, gene balance and the fractionation process back toward diploidy.

50. •• Buggs RJ, Chamala S, Wu W, Tate JA, Schnable PS, Soltis DE, Soltis PS, Barbazuk WB: **Rapid, repeated, and clustered loss of duplicate genes in allopolyploid plant populations of independent origin**. *Curr Biol* 2012, **22**:248-252.
This study demonstrates the process of genome dominance and gene loss can be documented within short evolutionary time following polyploidy formation.

51. •• Xiong Z, Gaeta RT, Pires JC: **Homoeologous shuffling and chromosome compensation maintain genome balance in resynthesized allopolyploid *Brassica napus***. *Proc Natl Acad Sci USA* 2011, **108**:7908-7913.
With a newly formed tetraploid, chromosome numbers are erratic in the first generations but subsequently resolve to a balanced genome but with varying copies of the different parental chromosomes as long as the relative gene copies are similar.

52. •• Chester M, Gallagher JP, Symonds VV, Cruz da Silva AV, Mavrodiev EV, Leitch AR, Soltis PS, Soltis DE: **Extensive chromosomal variation generated in a recently formed polyploidy species, *Tragopogon miscellus* (Asteraceae)**. *Proc Natl Acad Sci USA* 2012, **109**:1176-1181.
A natural case of tetraploid formation with shifting copies of parental chromosomes but with overall maintenance of relative gene copy number.