# THE EFFECTS OF ANEUPLOIDY ON GENE EXPRESSION IN A DOSAGE SERIES OF MAIZE CHROMOSOME ARM 1L

A Dissertation

Presented to

The Faculty of the Graduate School

At the University of Missouri

In Partial Fulfillment

Of the Requirements for the Degree

Doctor of Philosophy

By

Adam Johnson

Dr. James Birchler, Dissertation Supervisor

May, 2017

The undersigned, appointed by the dean of the Graduate School,

have examined the Dissertation entitled

THE EFFECTS OF ANEUPLOIDY ON GENE EXPRESSION IN A DOSAGE SERIES OF MAIZE CHROMOSOME ARM 1L

Presented by Adam Johnson

A candidate for the degree of

Doctorate of Philosophy

And hereby certify that, in their opinion, it is worthy of acceptance.

_____

Dr. James Birchler

_____

Dr. Jianlin Cheng

_____

Dr. Kathleen Newton

_____

Dr. J. Chris Pires

# ACKNOWLEDGMENTS

To complete the work described in this dissertation, I relied on the help, guidance, and collaboration of many people. Here is a partial list of the colleagues, friends, and family who contributed to making this possible:

Dr. James Birchler, for providing patience and sound reasoning to a guy who really needed it.

All past and present members of the Birchler lab in every capacity, including Patrice Albert, Weihong Chen, Jon Cody, Ryan Donohue, Dr. Ryan Douglas, Dr. Robert Gaeta, Dr. Zhi Gao, Nat Graham, Spencer Lambdin, Morgan McCaw, Dr. Abhijit Sanyal, Dr. Xiaowen Shi, Dr. Lin Sun, Nathan Swyers, Dr. Weiwu Xie, Dr. Hua Yang, and Dr. Changzeng Zhao.

Dr. Jianlin Cheng and his lab, including Jie Hou, Md Soliman Islam, and Dr. Jilong Li.

Dr. Birchler, Dr. Cheng, Dr. Kathleen Newton, and Dr. Chris Pires for serving as my committee.

Debbie Allen and the Life Sciences Fellowship, Vicki Bryan and the Interdisciplinary Plant Group, and Nila Emerich and the Division of Biological Sciences, for the support – financial, administrative, community, you name it.

Dr. Birchler, Dr. David Braun, and Dr. Amanda Durbak for providing me the opportunity to teach under their guidance.

My parents, John and Sally, and my brother and sister, Ryan and Rebecca.

My wife and best friend, Linh Ngo, who knows why I saved this part for last.

# Table of Contents

# List of Illustrations

## Figures

**Tables**

# Chapter 1 – Introduction

## Aneuploidy and gene dosage – history of the concept, occurrence in nature, and application to medicine

Aneuploidy is the condition of having a different copy number of a chromosome or chromosomal region compared to the copy number of the remainder of the genome. It was first described in *Drosophila* (BRIDGES 1916) and *Datura* (BLAKESLEE *et al.* 1920) in the 1920s. With the previous recognition that genes map to specific loci on chromosomes, it was deduced that aneuploidy results in a change of gene dosage. Aneuploidy can produce strong phenotypic effects, and from the earliest studies it is known that aneuploidies for different chromosomes may be distinct from each other (BLAKESLEE 1934). Understanding of how gene dosage and aneuploidy produce phenotypic effects is incomplete, and is an area of ongoing investigation (BIRCHLER 2013).

Aneuploidy was first described in the plant species *Datura stramonium*, aka Jimson weed. Aneuploidy has no known agricultural value, while being a significant human health issue. Reliable karyotyping of human cells did not become technically feasible until the 1950s. However, plant material could produce high-quality chromosome samples for microscopy, using the techniques of the time. Plant research contributed significantly to the fundamentals of chromosome structure and function, by work of such giants of cytogenetics as Barbara McClintock and George Beadle. Overlapping with the early development of plant karyotyping techniques was the heyday of the use of *Datura stramonium* as a model organism. Alfred Blakeslee made the first descriptions of

aneuploidy using *Datura* in 1921, and was responsible for creating the basic vocabulary of aneuploidy, which is still used to this day. Under this system, individuals with a balanced set of chromosomes (such as a diploid, in which every chromosome exists as part of a homologous pair in each cell) are euploid, and those with different numbers of chromosomes in one or more homologous sets are aneuploid. The condition of carrying an extra chromosome compared to the balanced set is called hyperploidy, while having a missing chromosome is called hypoploidy. Referring to the varied chromosome, to have a set of 3 is called a trisomy, of 4 is a tetrasomy, of only 1 is a monosomy, and of 2 (for example, in an otherwise-haploid individual) is a disomy.

Autosomal aneuploidy in adult humans is mostly limited to trisomy of chromosome 21, known as Down syndrome (LANA-ELOLA *et al.* 2011). There are other, more rare trisomies possible in humans, including 18 (Edwards syndrome) (CEREDA and CAREY 2012) and 13 (Patau syndrome) (CAREY 2012), though the vast majority of individuals born with these conditions die in infancy. Mouse models of Down syndrome have been produced, which are trisomic for part of chromosome 16, orthologous to human chromosome 21 (YAMAKAWA 2012). However, neither mice nor humans can tolerate other large-scale ploidy changes, which are usually lethal in early embryonic development (HASSOLD and JACOBS 1984). Down syndrome is a human developmental condition caused by the presence of an extra copy of chromosome 21, though this was not known until long after the initial descriptions of aneuploidy in plants. It occurs approximately once for every 700 live births. The condition was first systematically described by John Langdon Down in 1862, based on his observations of individuals legally under his care at a British state institution. Down noted the most distinct physical

effects of the condition, including intellectual disability, eye shape (resulting from a prominent nasal bridge), and a characteristic palmar crease. Down syndrome is now known to cause an increased risk of heart disease, leukemia, and susceptibility to pneumonia. Penetrance of the features of Down syndrome is highly variable, and not all characteristic symptoms are universal. Jerome Lejeune, one of the first researchers to successfully produce human karyotypes, identified the genetic cause of Down syndrome in 1959. In the intervening years, health outcomes for those diagnosed with Down syndrome have improved dramatically. The development of antibiotics increased life expectancy to 25 years in 1985 (from an estimate of 10 years in 1910). Modern techniques in heart surgery have brought life expectancy to an estimate of 60 years today.

Regarding dosage sensitivity, sex chromosomes form a category of their own. They offer one platform for the study of dosage compensation, which is the adjustment of gene products to a normal level despite an altered copy number of the gene itself. In species with sex chromosomes, males and females have different karyotypes for them, meaning one sex is inherently aneuploid. Human males are "aneuploid" (differing from an otherwise diploid state) for the X chromosome, but adaptive dosage compensation ensures that X-linked gene expression is similar in males and females. Sex chromosome aneuploidies (in the sense that they differ from the XX/XY dichotomy) also occur in humans, and with higher frequency than aneuploidy of the autosomes. Abnormal chromosome counts of either the X or the much smaller Y chromosome can result in a variety of diagnosed conditions (HALL *et al.* 2006). XXY is a type of Klinefelter syndrome (defined by the presence of at least one extra X in males), and is the most common human sex chromosome aneuploidy (VISOOTSAK and GRAHAM 2006). It occurs

possibly as often as one in 500 live births, and the primary consistent feature is sterility. Otherwise, symptoms are highly variable, and are not as profound as seen in autosomal aneuploidies. Turner syndrome is another example, occurring in females missing all or part of an X chromosome (HJERRILD *et al.* 2008). Up to one in 2000 live births are affected, and as with Klinefelter syndrome, the medical impact is less severe than for monosomy of any autosome – in fact, autosomal monosomies are not seen at all in live-born individuals. As with Klinefelter syndrome, Turner syndrome results in sterility and an otherwise highly variable set of physical attributes.

The widespread occurrence of aneuploidy in cancer cells has been reported from the early 20th century (RICKE *et al.* 2008). Many contradictory claims have been made about causes and effects, but there is some consensus about the progression of events. Aneuploidy becomes progressively worse as chromosome stability in the cancerous cells decreases (CHAN 2011). In advanced cancers, massive aneuploidy can be found in most cells, often with completely different chromosome sets in different cells (PFAU and AMON 2012). The impact of this on further tumor development is speculative, but some interesting assumptions can be made from an evolutionary point of view. A multicellular organism apparently needs to maintain most of its genomic content in a balanced state to survive, but this is not true for an individual cell. It only needs to satisfy its own structural and metabolic needs, independent of what would benefit the tissue or organism. Just how much of a balanced genome that requires is unknown, but the fact that tumor cells may survive even with different copy numbers of multiple chromosomes suggests it is very little. It may even be that certain combinations of chromosomes in certain dosages provide selective advantages to the cell (PAVELKA *et al.* 2010).

While aneuploidy implies a gene dosage change for all the genes on a particular chromosome, changes at smaller scales are also known to produce major impacts on organisms. For example, dosage sensitivity of single genes may be found in the form of haploinsufficiency. This is the source of dominant loss-of-function genetic conditions: to have zero copies is presumably lethal, but one is tolerated. If the disabled copy of the sensitive gene is passed on, the offspring will also have only one working gene. This is rare enough to merit special attention in the literature, because the majority of described genetic conditions are recessive, meaning both copies have to be non-functional to produce a measurable phenotype (VEITIA and BIRCHLER 2010). In other words, most genes are not dosage sensitive. X-linked disorders are good examples of the apparent majority of genes: males with a non-functioning gene display a phenotype, but females with even one functioning copy do not. There is no phenotypic difference between females with one or two functioning copies, or males with one. Autosomal dominant conditions include Ehlers-Danlos syndrome and Marfan syndrome, both connective tissue disorders (CALLEWAERT *et al.* 2008).

When studied from an evolutionary point of view, gene dosage work is often focused on whole genome duplications (WGDs). WGDs may result in instant reproductive isolation, making them important to speciation (SEMON and WOLFE 2007). WGDs do not result in a permanently doubled (polyploid) genome size, but instead are followed by a gradual whittling down of genes back to diploidy (EDGER and PIRES 2009). Certain gene types are preferentially retained in duplicate, and others preferentially returned to single-copy status with unexplained consistency (EDGER and PIRES 2009). Presuming an individual must maintain some minimum level of fitness just to survive

amongst its neighbors, it can only lose genes and survive when there is little fitness cost for doing so. This means that altering the dosage of that gene relative to the rest must have a low impact – it must be dosage insensitive. Predictive *in silico* models of the yeast metabolic network have become valuable tools in studies of gene loss following WGDs. These models match the known outcomes of WGDs and subsequent gene loss in *Saccharomyces cerevisiae* to an impressive degree (CONANT and WOLFE 2007).

Polyploidy has occurred numerous times in the course of evolutionary history, including within our own lineage. Whole-chromosome aneuploid events, in contrast, are not detectable. This disparity is consistent with Alfred Blakeslee's Datura studies, from which he defined the fundamental concepts of aneuploidy. His comparison of the phenotypic effects of aneuploidy and polyploidy showed that an imbalanced genome, despite having a smaller absolute change of genome size, produced more detrimental phenotypic effects than a balanced genome, even in a polyploid state. This observation has been noted in a number of other species, including Arabidopsis, maize and Drosophila. Copy number variants for single genes, or small segments of chromosomes containing a few genes, are common in evolutionary history, as are whole genome duplications.

**Gene expression studies – with methods including cytogenetics and discussion of results**

Aneuploid conditions can result from improper segregation (nondisjunction) of the chromosomes in either meiosis or mitosis. Nondisjunction in meiosis results in a completely aneuploid individual, since the fusion of the mostly haploid gamete with a

normal haploid gamete forms the template for all of the new individual's (mostly) diploid somatic cells. Nondisjunction is responsible for the most common human aneuploid syndromes (NICOLAIDIS and PETERSEN 1998). Mitotic nondisjunction results in a mosaic of regular and aneuploid cells, depending on the cell lineage of the progenitor aneuploid cell. Aneuploid syndromes due to incorrect mitotic segregation are uncommon compared to meiotic nondisjunction, though mitotic mis-segregation is notably the direct cause of the aneuploidy seen in cancer cells (KING 2008). The rate of nondisjunction is uncertain in both mitotic and meiotic cells, though *in vitro* studies suggest rates as high as one per 100 divisions (GORDON *et al.* 2012). Theories about the cytological causes of nondisjunction are numerous, including improper numbers of centrosomes and extra-strong binding of kinetochores to microtubules (CHAN 2011).

Nondisjunction of B chromosomes in *Zea mays* has proven to be very useful to the study of aneuploidy and dosage effects. Maize B chromosomes are supernumerary non-vital chromosomes consisting of a centromere with a single long arm; the short arm is vanishingly small (RANDOLPH 1941). They are selfish elements: in male flowers B chromosomes tend to nondisjoin during the second pollen mitosis (ROMAN 1947), and those gametes that have two copies of the B have more reproductive success than those with zero (JONES *et al.* 2008). In this way, B chromosomes harness nondisjunction as a drive mechanism. B chromosomes occur naturally in some maize lines, and do not cause any noticeable impairment to the plant unless their number approaches that of the A chromosomes (meaning the regularly numbered chromosomes) (RANDOLPH 1941). An A chromosome segment can be translocated onto a B centromere (BIRCHLER *et al.* 1990a). Such a B-A chromosome can itself participate in crossing over with homologous A

chromosome regions, and researchers have taken advantage of this to add marker genes to the B-A (BIRCHLER and ALFENITO 1993). If a marker for kernel color has been added, for example, this allows easy identification of seeds that have the B-A chromosome. A dominantly marked embryo will have been generated by the joining of the egg and a sperm with two B-A chromosomes; the endosperm, resulting from fertilization of the polar nuclei by the counterpart sperm, will be missing the dominant marker (BIRCHLER *et al.* 1990a). A line of maize with B-A chromosomes may be aneuploid for that A chromosome region, and that allows for the creation of a dosage series for any chromosome so desired. Fifteen unique B-A chromosomes with dominant kernel color markers have been developed and introgressed into the maize W22 line (AUGER and BIRCHLER 2002). These stocks can be used to study the effects of both hypoploidy (missing chromosomes) and hyperploidy (extra chromosomes) for most of the maize genome.

Use of B-A chromosomes relies on alleles of the *r* gene, which in a specific dominant form, *R-scm2*, provides a dark purple pigment to the maize kernel, both embryo and endosperm. For use in creating a dosage series for a given chromosome arm, *R-scm2* is present with other recessive anthocyanin gene alleles. If the B-A chromosome has been backcrossed into a tester line with a recessive marker gene, the independent genotypes of the embryo and endosperm then allow us to distinguish which of them has inherited any copies of the B-A chromosome. Given a male parent with at least one B-A chromosome and its translocation partner (an A-B chromosome), the complete genomic content of the offspring can be derived. If the kernel includes a colored embryo and a colorless endosperm, then the embryo is hyperploid and the endosperm hypoploid. By way of

explanation, meiosis in the male parent requires balance for continued viability – the microspore will be euploid, containing either the A chromosome or B-A plus A-B. Nondisjunction of B chromosomes may occur during the second pollen mitosis, potentially resulting in gametes with 0 or 2 copies of the translocated arm of A. If either the embryo or endosperm has received a B-A (and its dominant anthocyanin allele), then both parts of the kernel have received the A-B. It follows, then, that in a kernel with color in the embryo and not the endosperm, the embryo has received 2 B-As, 1 A-B, and one normal A chromosome (from the female parent), for a total A arm dosage of three. Meanwhile, the normally triploid endosperm in the same kernel has inherited 0 B-As, 1 A-B, and two normal chromosome 1s (from the female parent), for a total arm dosage of 2. Conversely, if an endosperm has color while the embryo does not, then the endosperm has an extra copy of the translocated arm and the embryo a missing copy. If both parts of the kernel are colored, then both are euploid. This system can be used for the production of a dosage series, by selection of hyperploid kernels (colored embryos, colorless endosperms) for use as parents, initially backcrossing them with the recessive tester of inbred line W22. Then, hyperploids can be self- or sibling-crossed to produce a dosage series among the offspring. These offspring may include a wide variety of genotypes, with dosages of the A arm ranging from 1 to 4, and the correlation of color to dosage is lost in this generation. Alternatively, a hyperploid parent may be crossed with a recessive tester to produce a dosage series with a dosage range of 1 to 3. Each individual can be genotyped from kernel phenotype, and cytologically as a secondary verification. Haploids and disomic haploids may be produced by crossing a haploid-inducer parent such as RWS to a hyperploid (ROEBER *et al.* 2005).

Fluorescence microscopy is the primary cytological method applied to B chromosomes in maize. Fluorescence In Situ Hybridization (FISH) is performed using oligonucleotide fluorescent probes. The probes bind to complementary strands in the genomic DNA of cells, which are already crosslinked to a slide. Fluorescent signals from the hybridization sites are overlaid with DAPI signals for DNA, producing a karyotype that indicates the number and loci of a sequence of interest. The most common sequence of interest in maize with B-A translocations is the B-repeat sequence. This is a ~1,000-nucleotide repeat that occurs uniquely on the B chromosome, with the predominant signal at the centromere (ALFENITO and BIRCHLER 1993). A second, weaker signal occurs at the distal tip of the single B chromosome arm. When karyotyping an individual with B chromosomes but no translocations with any A chromosome, both the strong and the weak B-repeat signals can be seen together, demonstrating the number of the cytologically distinct B chromosomes. When B-A translocations are present, the strong and weak signals are separated. The strong B centromere signal shows the presence of the B-A chromosome, while the weak signal shows an A-B chromosome.

The effects of aneuploidy on gene expression levels can be widespread. A B-A chromosome dosage series in maize showed extensive differences in the quantities of selected proteins depending on how many doses of chromosome arm 1L were present (BIRCHLER 1979). The gene loci of the affected proteins were not on 1L as a direct dosage relationship would suggest, but instead were scattered around the whole genome. This has come to be defined as a *trans*-effect, as opposed to a *cis*-effect which affects genes of the varied region itself. Even more unexpected, many of the *trans*-affected genes showed an inverse relationship to the dosage of 1L (BIRCHLER 1979). Given three copies

of 1L, affected proteins were found at levels lower than in a normal diploid, and given one copy their levels were increased. Another important finding of this study was that a protein encoded on the varied portion of 1L was not affected by the dosage of its own locus – it was dosage compensated (BIRCHLER 1979).

One possible explanation for this is that certain regulatory genes on 1L are down-regulating genes on the other chromosomes, and the more these 1L genes are expressed, the less their targets would be. However, if the dosage of 1L regulators has a negative downstream effect on the genes of other chromosomes, it may also have the same effect on some of its own genes. Furthermore, the altered copy number of a given 1L gene may counter the effect of those genes that down-regulate it, with the result being dosage compensation. This theoretical explanation for dosage compensation of *cis* genes was demonstrated in the same maize dosage series, but the aneuploid chromosome arm 1L was varied in small segments, not all together (BIRCHLER 1981). When the whole arm is varied, the easily measured *Adh* gene is dosage compensated (BIRCHLER 1979). When the *Adh* structural gene locus (as distinguished from the loci that regulate its production) is varied by itself, a direct correlation is seen with its protein product (BIRCHLER 1981). *Adh* has several regulators nearby on 1L, and varying any one of them causes an inverse effect on the production of ADH protein (BIRCHLER 1981). As a fraction of normal levels, the gene's structural locus dosage multiplied by the inverse of its regulatory loci dosage approximates one.

The work on *Adh* established that there is an inverse dosage effect of 1L on other loci in the genome, and that its similar effects on itself were countered by the same altered dosage (BIRCHLER 1981). It could not be concluded, however, whether this is a

general effect of chromosome variance or something unique to that region of chromosome 1. This question was addressed using other B-A chromosome lines, which in total allowed for analysis of dosage series for approximately one third of the maize genome (BIRCHLER and NEWTON 1981). The results show that the inverse effect described for 1L also occurs for most of the other tested chromosome segments. This response was initially found with protein levels; mRNA transcripts were later shown to be similarly impacted (GUO and BIRCHLER 1994). This suggests that the inverse effect is a general outcome of chromosome variation, and it has been proposed to be responsible for many aspects of aneuploidy in general.

Sex chromosomes are essentially a built-in, permanent dosage variation. Many studies of sex chromosome balance and dosage compensation have been done in *Drosophila* (BIRCHLER *et al.* 2011). As in mammals, *Drosophila* females have two X chromosomes, while males have one (DISTECHE 2012). Rather than X-inactivation, the typical mechanism of dosage compensation in mammals, the single X in a male is expressed at a similar level to both X chromosomes combined in the female (DISTECHE 2012). The gene content of the Y chromosome found in males is vanishingly small, and may be ignored in dosage studies looking at global effects (BIRCHLER 2013). X can be varied against a diploid background, for example in metafemales, which have 3 copies of X. Alternately, the ploidy of the autosomes can be changed, as in a triploid which might have 1, 2, or 3 copies of X (1 X chromosome in a triploid is defined as a metamale, and 2 X chromosomes is an intersex) (SUN and BIRCHLER 2009). Easily measured phenotypes such as eye color can be connected to the expression of particular genes or chromosome regions (SUN *et al.* 2013). Studies focused exclusively on the autosomes have found that

the inverse *trans*-effect and dosage compensation in the varied region are both commonplace (BIRCHLER *et al.* 1990b). In this instance, *Drosophila* autosomes can be compared to the unvaried chromosomes of maize, together showing that gene balance effects are widespread in diploid organisms.

One category of non-genic material to be considered is transposable elements (also called transposons). In maize, transposons form a large majority of the total genome (BENNETZEN 2000). They tend to be epigenetically silenced, but the mechanism by which the cell targets them is unknown (BIRCHLER *et al.* 2000). An inverse correlation of transposon dosage and expression in diploids has been noted (MEYERS *et al.* 2001). Transposable elements with low copy numbers are expressed at a higher level than those with high copy numbers. Certain transposons can be found in maize with copy numbers in the hundreds of thousands; their typical expression levels are very low (FESCHOTTE *et al.* 2002).

**The gene balance hypothesis – proposed mechanisms of aneuploid effects and interpretation of gene expression studies**

The model of genes as having alleles which are either dominant or recessive is based on an all-or-nothing assumption regarding gene functions. Presuming a gene exists to produce a protein, it either does or does not, and variation is essentially of the presence-or-absence kind. Enzymes were the first type of protein to be functionally described, and the kinetics of this class of molecules with their substrates makes this interpretation quite logical. For a single-protein enzyme, there is a steep upward curve on the graph of quantity in relation to functionality. Direct dosage effects have been assumed

to be the cause of aneuploid phenotypes, and this must be correct at the most fundamental level. However, the complex nature of the disruptions caused by the extra products of extra chromosomes is not conveyed by the statement that certain proteins are overabundant.

From numerous studies, a theoretical model of dosage sensitivity has been developed, called the gene balance hypothesis (BIRCHLER and VEITIA 2007). It suggests that when gene products are part of a larger molecular complex, a change of subunit stoichiometry affects the function of the whole. That is to say, the quantities of each gene create a particular stoichiometry for a larger system, and upsetting that stoichiometry will upset the system (BIRCHLER and VEITIA 2010). These molecular complexes could be a variety of types, including (but not limited to) metabolic pathways, cell signaling pathways, and multi-subunit protein complexes such as transcription factors. For a multi-subunit protein complex, a protein may be defined as a "bridge" if it must bind with at least two other proteins for the complex to function. If the amount of the bridge protein in increased beyond the dosage of its binding partners, then a shortage of the complete set may result, due to the formation of incomplete complexes. The stoichiometry of this can be described with simple fractions. For example, increasing the gene dose from 2 to 3 will cause a protein to be expressed at 3/2 of the normal level. The amount of whatever lies downstream, such as a completed complex, will fall to its inverse, which is 2/3 (BIRCHLER and VEITIA 2012). This effect is illustrated in Figure 1. *In vivo*, these mathematically simple effects may be buffered by a variety of processes, including epigenetic modifications and varying rates of synthesis and degradation of transcripts and their encoded proteins (VEITIA *et al.* 2013).

The sex chromosome regulatory mechanism of Drosophila appears to have co-opted the phenomenon of dosage compensation in order to maintain tolerable expression levels of the autosomes, regardless of how many X chromosomes are present (BHADRA *et al.* 2000). Presuming a *trans* effect, and without a special mechanism to correct it, the 50% decrease in the dosage of X-linked regulatory genes in males would cause a 200% expression of their X and autosomal targets. This is exactly what happens when the regulatory apparatus is knocked out, specifically the key complex called male specific lethal (MSL) (BHADRA *et al.* 1999). For some time after MSL was identified as being part of this process, it was assumed that it simply doubled expression of the male X, to match the combined output of the two X chromosomes in females (LUCCHESI *et al.* 2005). In fact, newer evidence suggests that MSL works on the other side of the inverse: it sequesters an up-regulating histone acetylase onto the X, to keep autosome expression stable by blocking the inverse *trans* effect (BHADRA *et al.* 2000). The extra histone acetylase is also prevented from altering X-expression by MSL (BHADRA *et al.* 2000). A specific means to increase X-expression turns out to be unnecessary, perhaps because it is subject to the same dosage compensation mechanism as described in the gene balance hypothesis.

Despite the distant relationship of the approaches, metabolic modeling has reinforced some of the same conclusions as the gene balance hypothesis. The genes most likely to be kept in duplicate are highly connected in the metabolic network (ZHU *et al.* 2007). This is comparable to the hypothesis that in molecular systems, stoichiometry of the component parts is a vital matter (BIRCHLER and VEITIA 2010).

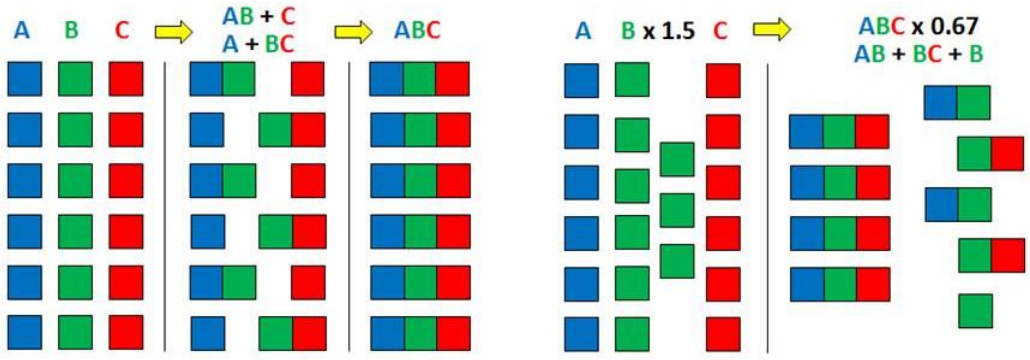


**Figure 1 – Kinetics of a multi-subunit complex**

In a hypothetical regulatory molecular complex A-B-C, B acts as a bridge between A and C. Due to assembly kinetics, an increase in the quantity of B results in a decrease in the total amount of A-B-C, which is reflected in target gene expression.

# Chapter 2 – Modulation of gene expression in aneuploids involving the long arm of chromosome 1 of maize

## Introduction

Using a set of genetic tools uniquely available in maize, a dosage series for a maize chromosome arm was produced in order to analyze the effects of genomic imbalance on the expression of genes. Maize plants were produced that were variable in dosage for the long arm of chromosome 1, referred to as 1L. Two separate sets of plants were raised to compare 6 distinct ploidy situations: diploid with 1, 2, 3, or 4 copies of 1L; and haploid with 1 or 2 copies of 1L. RNA sequencing was used to determine gene expression in leaf tissue from plants at each dosage level. This provides a view of the effects of aneuploidy on the entire transcriptome, by comparing differing levels of aneuploidy for the same chromosome arm against euploid controls.

## Methods

To determine RNA expression of all genes, leaf tissue was collected from plants at each dosage level. The first set included diploid plants with 2, 3, and 4 doses of chromosome arm 1L, and haploid plants with 1 and 2 doses. Leaf tissue was harvested at 45-46 days post-germination, so as to avoid differential expression of genes based on varying stages of development. A second set included diploid plants with 1, 2, and 3 doses, and leaf tissue was harvested at 55-57 days post-germination. Plants were greenhouse-grown to minimize environmental variation, and leaf tissue was collected at the same time of day (late mornings) to minimize circadian rhythm variation. The section

of leaf tissue collected was the region approximately between 1.5 centimeters and 15 centimeters above the auricle. Only the leaf blade was kept, and the midrib was discarded. The tissue was collected from the lowest adult-phase leaves, typically leaves 5 and 6 in the W22 inbred line. RNA was extracted using a TRIzol-based protocol (RIO *et al.* 2010). RNA sequencing was performed by the University of Missouri DNA Core, using Illumina technology (BENTLEY *et al.* 2008).

Different numbers of biological replicates were available for each dosage, in some cases due to mischaracterization of dosage before tissue collection, in others due to insufficient quality of RNA extracted. Four biological replicates were submitted for each dosage level in each of the two sets. The number kept, which is to say used for analysis and comparisons, was as follows: for the first set, Diploid – 4, Trisomic – 5, Tetrasomic – 3, Haploid – 3, Disomic – 2; and for the second set, Monosomic – 4, Diploid – 3, Trisomic – 3. Average expression values for each gene at each dosage were derived from these, and are used in all comparisons to other average expression values.

The two sets of plants for analysis were phenotyped using different methods, each of which was later verified by checking the relative consistency of cis gene expression in leaf tissue. In the first set, dosage of chromosome 1L was confirmed by FISH karyotyping, an example of which is shown in Figure 2. In the second set, dosages were limited to 1, 2, and 3 copies of 1L in a diploid background, which allowed for phenotyping of kernels according to the presence of a dominant color gene in the embryo or endosperm. Though differing methods of dosage verification do not interfere with comparability of results from the two sets, there were some distinctive environmental factors that suggest biological replicates from the two sets are ideally compared with

others of the same set, rather than the two sets against each other. Regarding adult plant phenotype, diploids were typically taller and more vigorous than trisomics, but not with great consistency. Tetrasomics were less vigorous than trisomics, and monosomics the least healthy in a diploid background. The slow development of the monosomics was the primary motivating factor in waiting longer than 45 days to collect leaf tissue, as had been done in the first set of plants. Haploids are shorter and less healthy than diploids; disomic haploids are severely stunted.

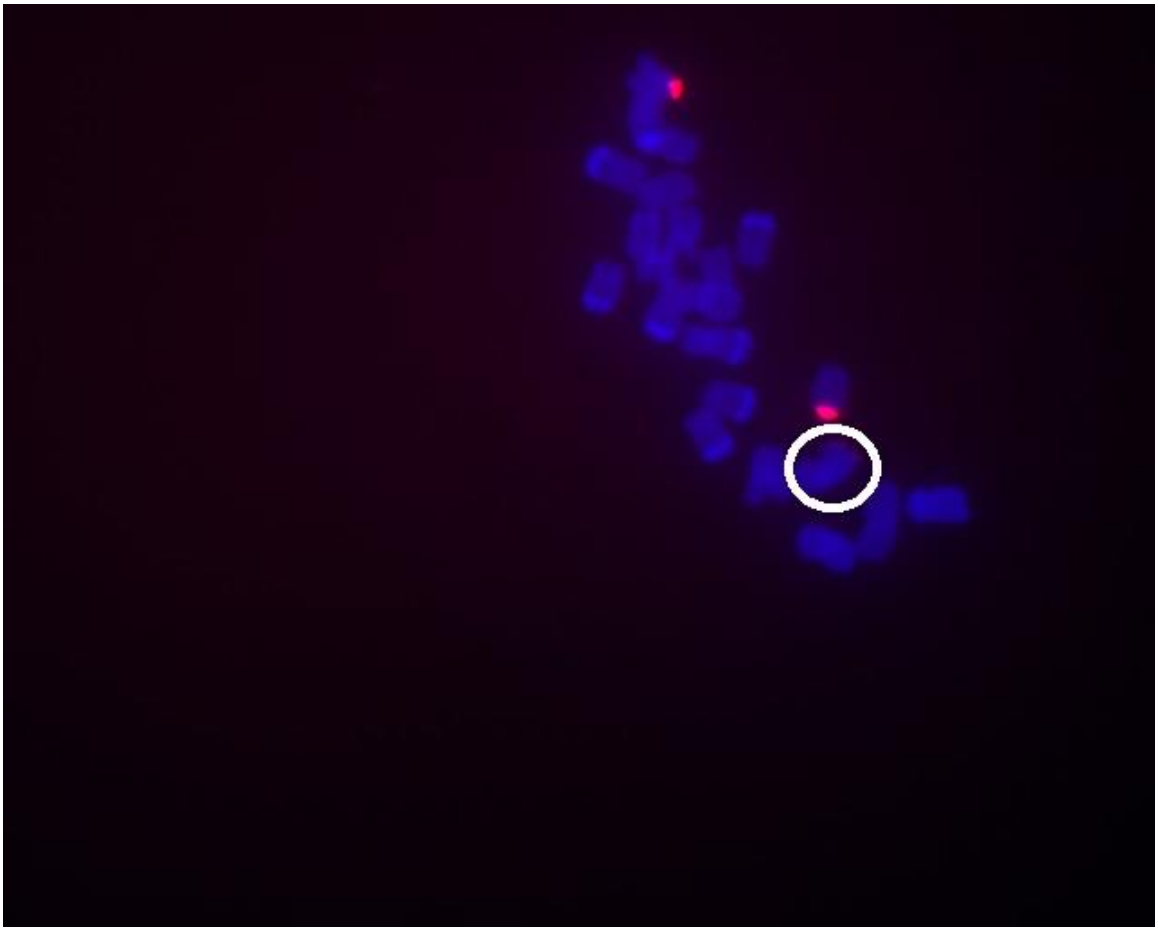

**Figure 2 – FISH image of trisomic chromosome spread**

Fluorescence *In Situ* Hybridization (FISH) was used to verify the karyotype of the first dosage series. The red fluorescent probe was complementary to the B-repeat sequence. The two brightest red signals represent the B chromosome centromere, in this case carrying a translocated chromosome arm 1L. A fainter signal, circled in white, occurs at the distal tip of the B chromosome arm, in this case translocated to chromosome 1. The two B-A chromosomes and the lone un-translocated chromosome 1 contribute a total of three doses of chromosome arm 1L.

## Results

The results of RNA sequencing for a collection of plants at various chromosome dosage levels were compared to the set of genes with non-zero expression in the collection of diploid controls. Genes with expression in the diploid but not the compared plants were removed from the set; this number was variable in the different dosage groups, and is one contributing factor to the different sizes of compared gene sets. For each gene in the set, RNA expression in the dosage-varied plant was divided by that of the diploid (or other euploid), producing a single value for each gene, with a ratio of 1.00 representing no difference between the two sets. In Figures 1 and 2 (ratio distribution plots), genes to the left of 1.00 on the x-axis have a decrease of expression compared to a diploid, and genes to the right have in increase. Figure 3 is comprised of cis gene ratios (genes located on chromosome arm 1L), and Figure 4 is comprised of trans gene ratios (genes located elsewhere in the genome). The two figures show a large difference of characteristic effects.

In all graphs comparing aneuploids to euploids, large-scale movement away from 1.00 and not corresponding to a normal distribution can be identified. Starting at the low end of the spectrum, all ratio distributions contain a spike in the bin ranging from zero to 0.05. This bin represents genes which are expressed at detectable levels in both diploid and aneuploid, but show at least a 20-fold decrease of expression in the aneuploid compared to the diploid. Another spike occurs at the high end of the spectrum, indicated a greater-than-6-fold increase in expression in the aneuploid. These sets may include genes which have been essentially switched off or on by the aneuploid condition, or they may be genes with very low read counts, leading to statistical noise. Figures S3 and S4

present a breakdown of ratios by expression level as measured by read count. The genes in the outlying spikes are almost entirely found in the low read count subset, which lends support to the latter interpretation.

A general trend observed in Figure 3 is that in cis, gene expression changes in correlation with gene dosage. A large peak can be seen near the guide corresponding to dosage increases, 1.50 in trisomic/diploid and 2.00 in tetrasomic/diploid and disomic/haploid. In the monosomic/diploid comparison, the peak appears over 0.50, corresponding to a dosage decrease. The peaks are broad, and tend to fall to the left of the guides in each hyperploid, and to the right in the hypoploid. This indicates the direct dosage effect in cis is partly compensated for different genes to different degrees. Gene expression ratios in trans, shown in Figure 4, do not resemble those seen in cis. Instead, in hyperploids the peaks move below 1.00, towards the inverse guides, which are 0.67 in trisomic/diploid and 0.50 in tetrasomic/diploid and disomic/haploid. In the monosomic/diploid, the peak moves above 1.00. These peaks are fairly broad, spilling over past 1.00 into the range of increasing gene expression. While inverse changes in gene expression predominate, a substantial number of genes are not affected, or are seen to change in the same direction as gene dosage. Among genes that are inversely affected, the typical range lies between the inverse ratio and 1.00.

Table 1 shows the results of a t-test for significant expression changes in each of the six comparisons shown in Figures 3 and 4. The haploid/diploid comparison has no meaningful distinction between cis and trans genes, and both gene sets show similar proportions of affected genes, with a slight favoring of increased expression compared with decreased. In cis, all five aneuploid/euploid comparisons show a large number of

genes with significant expression changes, and the vast majority are directly proportional to gene dosage. In the first dosage series set, the number of affected genes is greater in the tetrasomic/diploid and disomic/haploid than in the trisomic/diploid, correlating with their higher level of aneuploidy. In the same dosage series in trans, the trisomic/diploid comparison has slightly more genes affected than the euploid comparison, but the decreased-expression category is now favored over increased. In both the tetrasomic/diploid and disomic/haploid comparisons in trans, the number of affected genes is substantially larger than the trisomic/diploid, and the favoring of decreased expression is maintained. In the second dosage series in trans, the trisomic/diploid and monosomic/diploid comparisons both show more genes with a significant increase than a decrease. The favoring of increased expression is slightly larger in the hypoploid than in the hyperploid, in contrast to their dosage relationship. Figures 5 and 6 display the results of the t-tests in volcano plots. Table 2 displays the outcome of Kolmogorov-Smirnov (KS) tests comparing ratio sets of different aneuploid conditions; ratio sets for each dosage are shown to be significantly different from each other.

**Discussion**

Based on RNA sequencing results for the 1L dosage series, a number of general conclusions may be drawn. The one with perhaps the highest confidence is the importance of gene dosage in cis for gene expression at the RNA level. At each dosage level, genes with a structural locus of chromosome arm 1L were much more likely to be upregulated in a hyperploid as compared to a diploid, and downregulated in a hypoploid. The relative expression of most genes does not change in the case of polyploids.

Moderated or buffered effects are frequent, as shown by the ratio distribution plots. The broad peak, which accounts for the great majority of expressed genes, spans a range from dosage compensation (which implies no change from diploid) to direct dosage effect (implying a proportional change matching the gene dosage), with most falling in between. Few genes are directly affected beyond the ratio of their dosage change, and likewise few genes are reduced below their euploid level. The source of such widespread buffering may be the kinetics described in the gene balance hypothesis, with the direct effects of gene dosage and the inverse effects of genomic imbalance counteracting each other, resulting in a variety of intermediate effects.

A second phenomenon which can be seen in the aneuploid comparisons to euploid is the prevalence of inverse effects in trans. This directional change of gene expression, opposite the dosage of the varied chromosome, is less prominent in haploid comparisons to diploid than in aneuploid comparisons to euploid. In aneuploids, the inverse effect can be seen both holistically (in the ratio distribution plots) and statistically (in the volcano plots). To a greater degree than the cis direct dosage effect, the inverse trans effect is often a partial one, and direct trans effects also occur (though they are also typically partial). In most cases, a smaller portion of trans genes than cis genes can be statistically confirmed as having altered expression. According to the ratio distribution plots, fully inverse (or direct) effects in trans are relatively rare compared with partial effects. Because the effects are skewed from a ratio of 1.00 and are not random, it is possible that the RNA sequencing procedure will make the effects appear less extreme than reality. Though molecular kinetics may account for observed effects in both cis and trans, their partial buffering hints at great complexity in the relationships among genes with

regulatory effects, such as differential rates of synthesis and degradation of both mRNA and protein and how that affects the regulatory interactions. Expression ratios more extreme than inverse are rare.

Based upon a comparison of various aneuploid conditions, the effects of dosage appear to be progressive, in the sense that increasing levels of aneuploidy correlate with increasing alteration of gene expression. This applies both to the number of genes affected and the typical degree of change. Both disomic haploids and tetrasomic diploids have a greater level of imbalance (the varied chromosome occurs at 2x the copy number of the rest of the genome) than trisomic diploids (1.5x), and the number of genes with significant expression changes is higher in the disomic and tetrasomic than the trisomic. Based on the ratio distribution plots, the range of expression changes in cis moves from near 1.50 in the trisomic set to near 2.00 in the tetrasomic and disomic sets, proportional with gene dosage. In trans, the broad peak comprising the majority of genes moves further to the left in the tetrasomic and disomic compared with the trisomic, illustrating the greater inverse effect by increasing levels of genomic imbalance. A further hypothesis suggested by these data is that the inverse effect is not a general suppression of transcription due to aneuploidy-induced stress. This is most directly demonstrated by the presence of inverse effects in trans in hypoploid plants, which in that condition equates to an increase in expression. Furthermore, if a general decrease of expression were present, the direct trans effects seen in the trisomic set would be diminished relative to the inverse effect in the more extreme hyperploids. In fact, the number of direct trans effects also increases with increasing imbalance, implying that increasing disruption of the regulatory

network causes more and larger effects on individual genes rather than the entire set of genes as a whole.

Previous studies of gene imbalance in maize similarly found inverse trans effects, though these were conducted with kernel tissue rather than leaf tissue. The similarity of the phenomena observed suggests that despite the developmental differences between the two tissue types and stages, a common mechanism is responsible for the effects seen in both. Phenotypically, aneuploids (including maize aneuploids of 1L) are affected more strongly than polyploids, and modulations of gene expression observed here correlate with that observation. These changes in gene expression may be hypothesized to be the cause of the phenotypic effects seen in aneuploidy. If that is so, then aneuploid syndromes are conditioned by the altered stoichiometry of regulatory genes and their products.

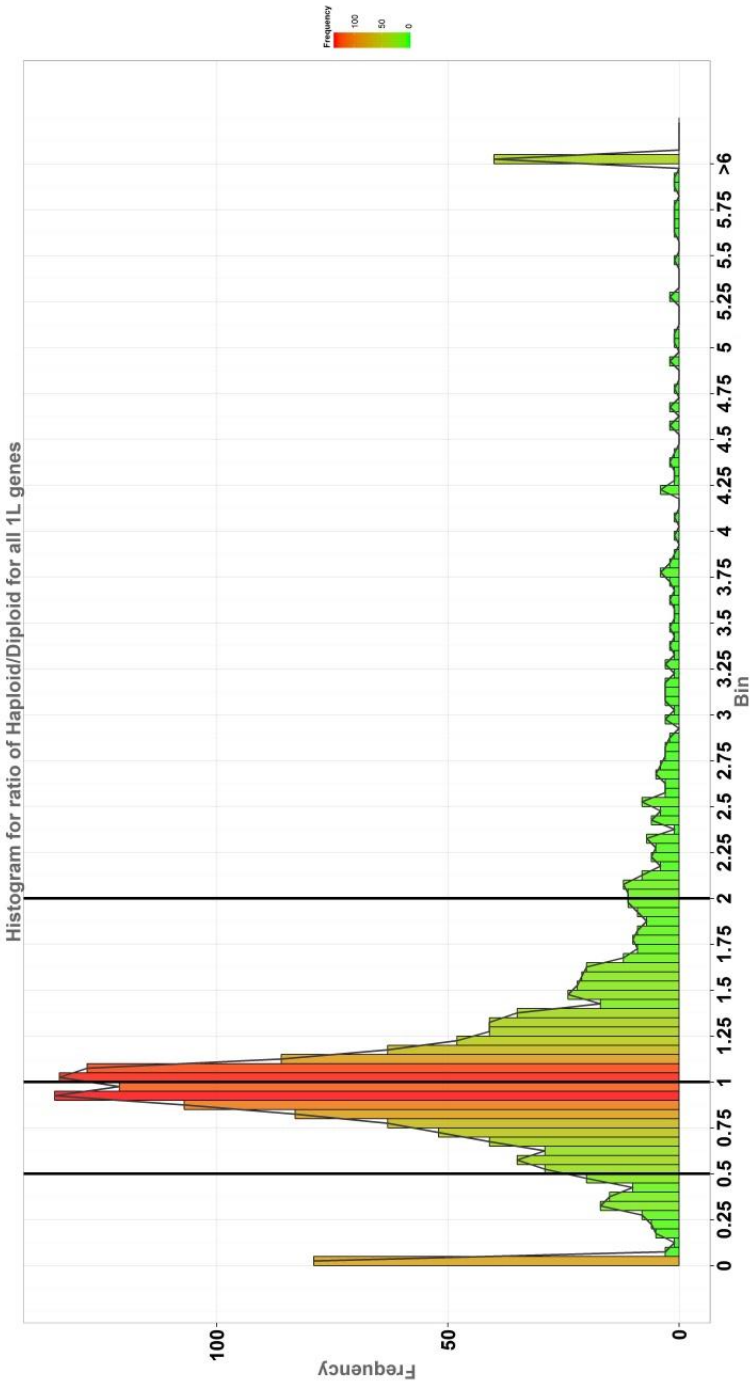


**Figure 3a – Ratio distribution plot for cis genes, haploid/diploid**

For every gene, the ratio of expression in an aneuploid plant compared to a euploid plant was determined. The x-axis indicates the expression ratio, and the y-axis indicates the number of genes in a given range. The three vertical guide bars indicate an inverse

relationship to dosage (0.50), a value unchanged from euploid (1.00), and a direct relationship to dosage (2.00). In plots of trisomic/diploid ratios, guide bars are placed at 0.67, 1.00, and 1.50. All expressed cis genes were considered. If a gene has the same average expression in the aneuploid condition as in the euploid condition, then its ratio will be 1.00. If expression decreases in an aneuploid, its contribution to the plot will fall to the left of 1.00; if expression increases, its position on the plot will be to the right of 1.00. This set of plots includes genes whose structural locus is on chromosome arm 1L, which is varied for dosage in these experiments. Figure 3a is a control comparison of haploid with diploid. Ratios are predominantly focused over 1.00, which suggests there are no systematic changes of relative (per transcriptome) gene expression in a comparison of two euploids. A rightward trend can be observed in hyperploids (Figures 3b, 3c, 3d, and 3f), and a leftward trend in the hypoploid (Figure 3e), indicating a direct (but often partial) correlation between gene dosage and RNA expression in cis.

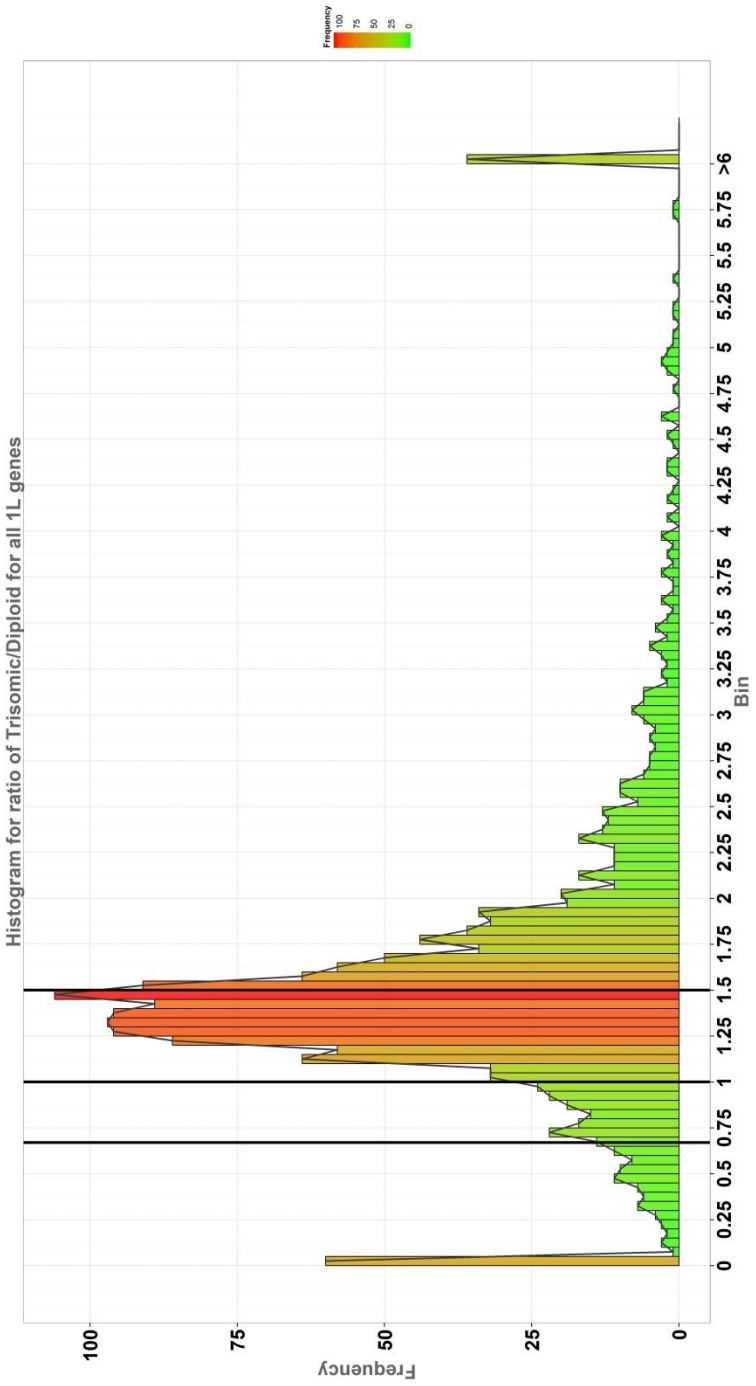


**Figure 3b – Ratio distribution plot for cis genes, trisomic/diploid**

For explanation see Figure 3a.

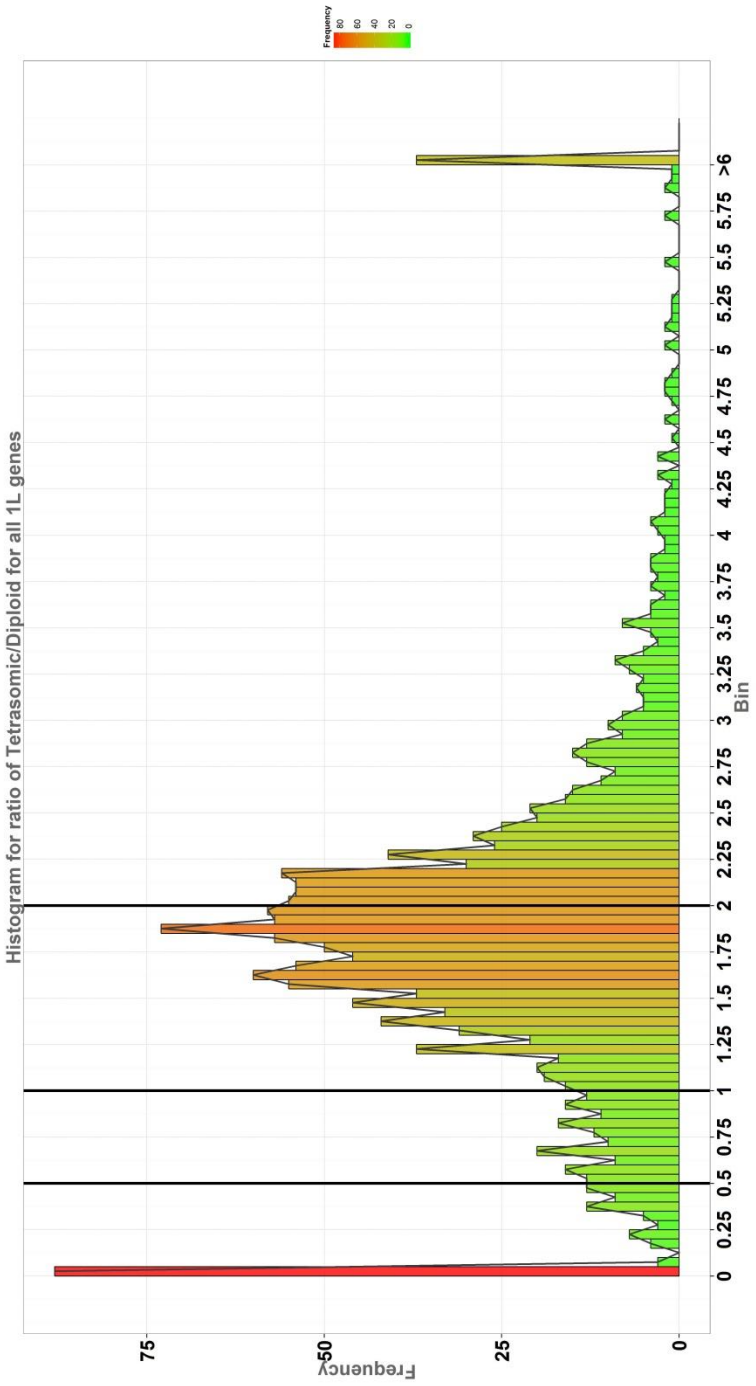


**Figure 3c – Ratio distribution plot for cis genes, tetrasomic/diploid**

For explanation see Figure 3a.

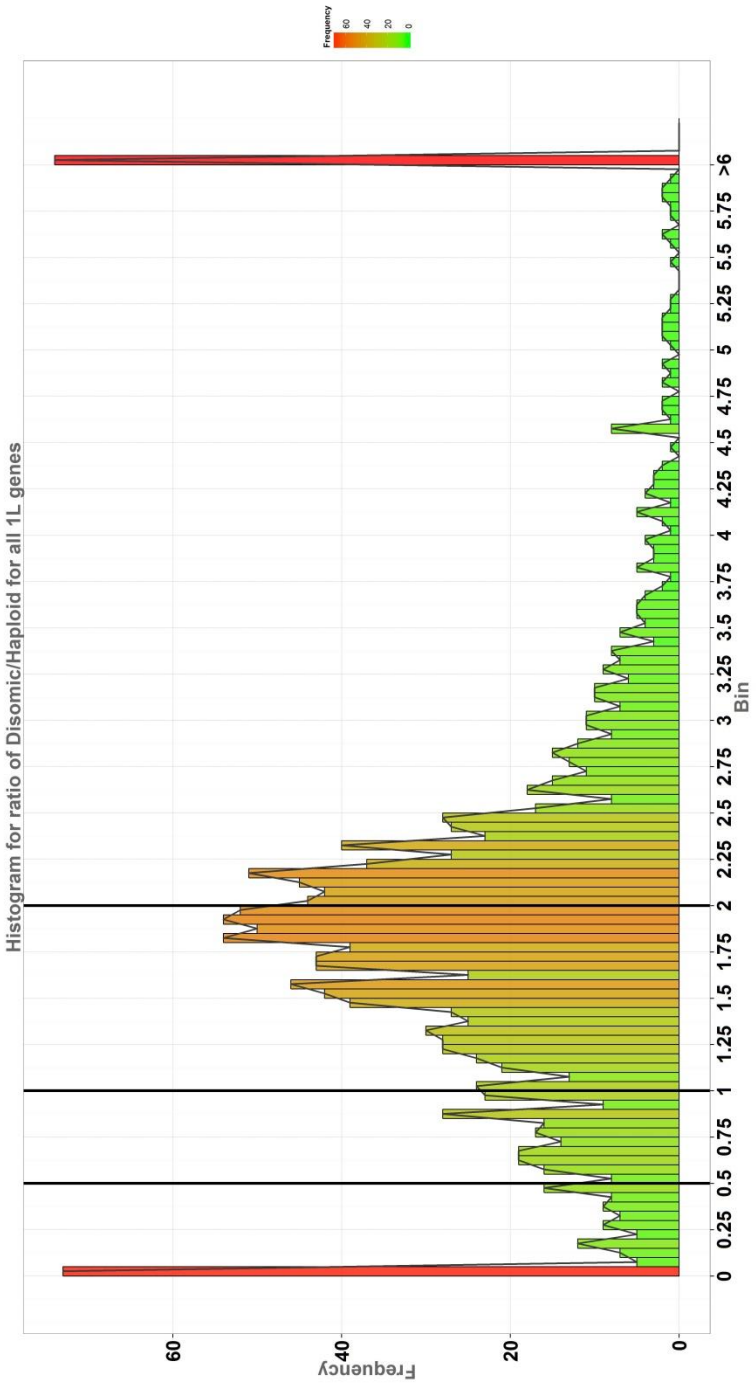


**Figure 3d – Ratio distribution plot for cis genes, disomic/haploid**

For explanation see Figure 3a.

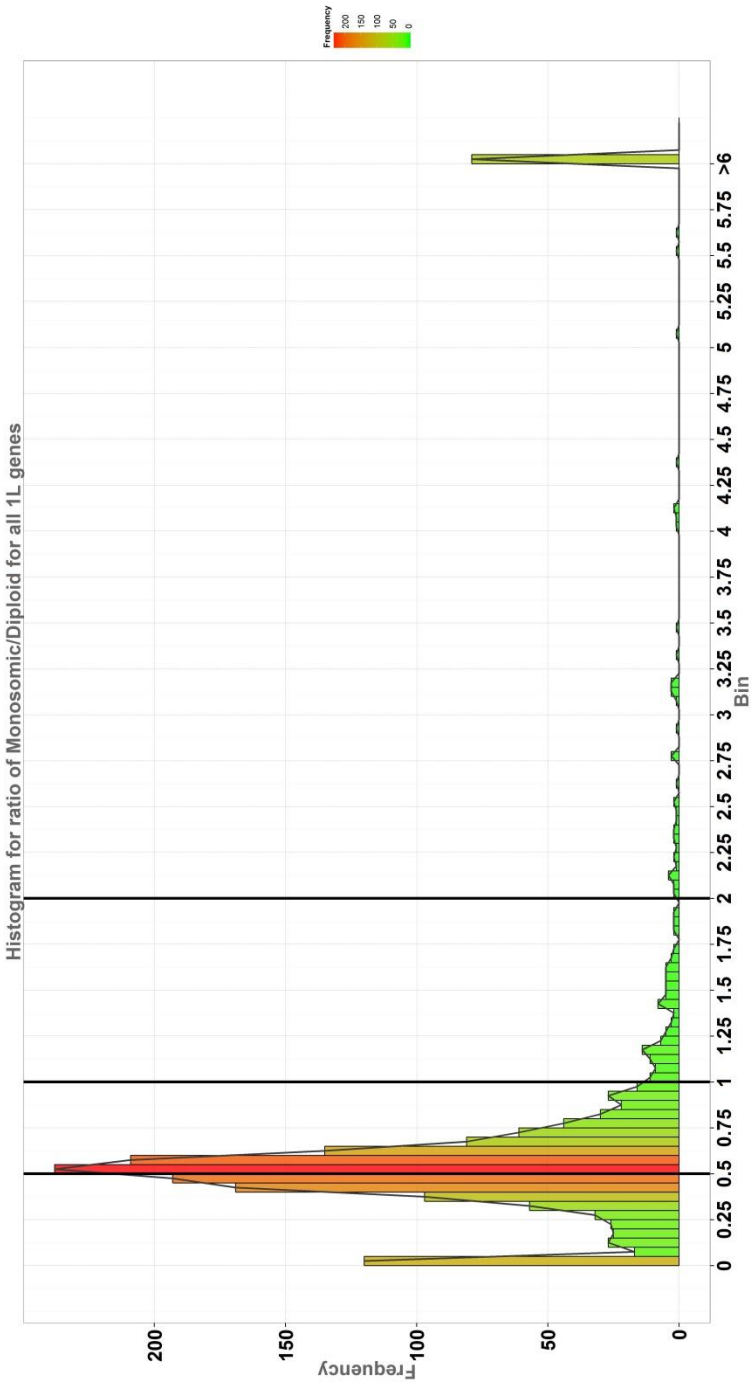


**Figure 3e – Ratio distribution plot for cis genes, monosomic/diploid, second dosage series**

For explanation see Figure 3a.

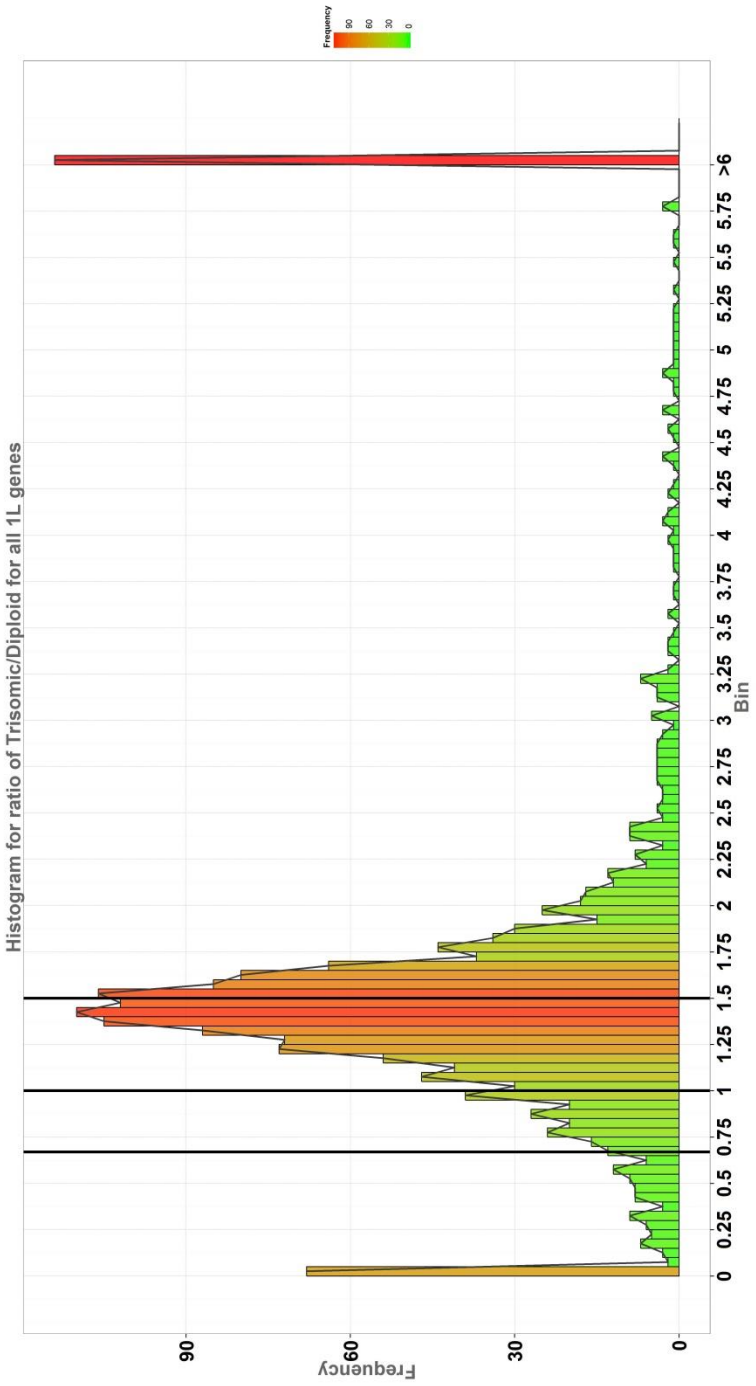


**Figure 3f – Ratio distribution plot for cis genes, trisomic/diploid, second dosage series**

For explanation see Figure 3a.

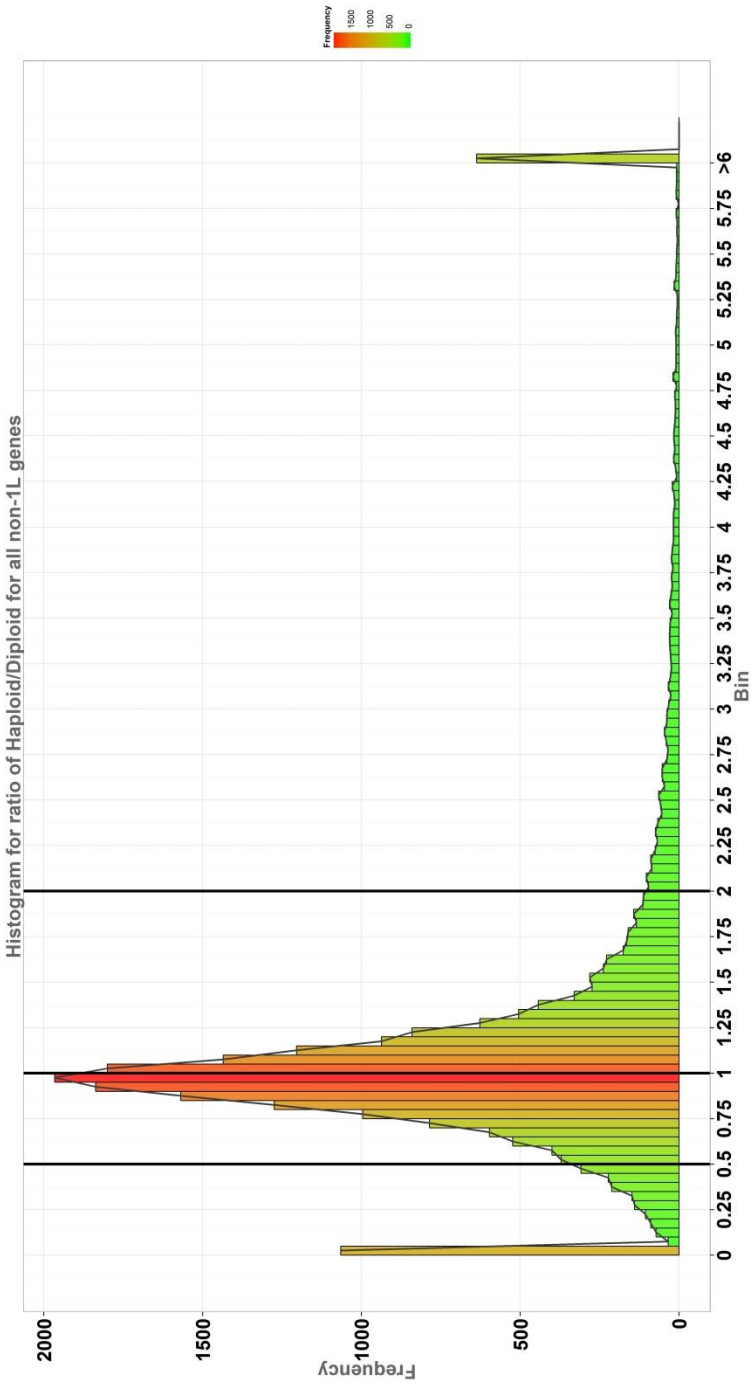


**Figure 4a – Ratio distribution plot for trans genes, haploid/diploid**

This set of plots includes genes whose structural locus is not on chromosome arm 1L, which indicates the dosage of these genes remains constant in these experiments. As in

Figure 3a, Figure 4a compares gene expression in haploids with diploids, and no systematic changes of relative gene expression are found. A “shoulder” can be observed on the left side of the central peak in the other three figures from the first dosage series (Figures 4b, 4c, and 4d), indicating that some genes show an inverse correlation between 1L gene dosage and non-1L RNA expression.

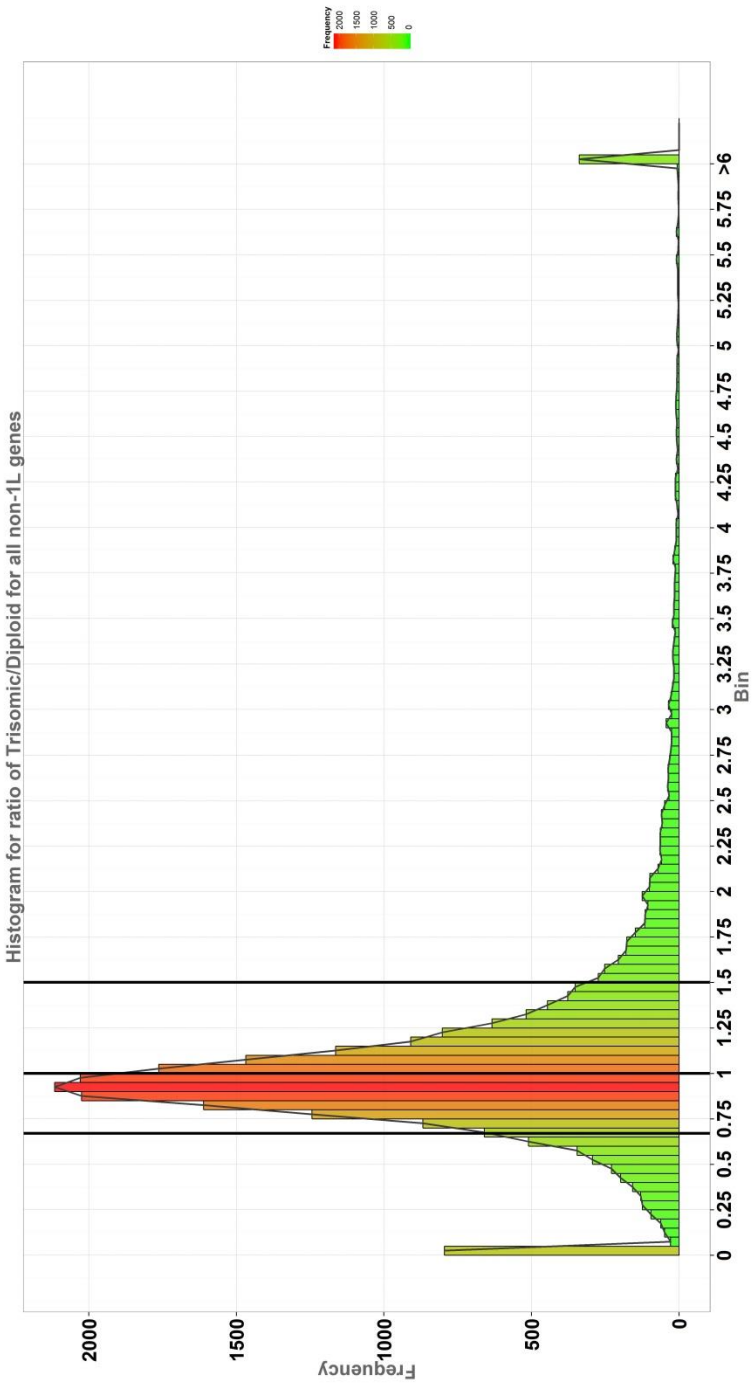


**Figure 4b – Ratio distribution plot for trans genes, trisomic/diploid**

For explanation see Figure 4a.

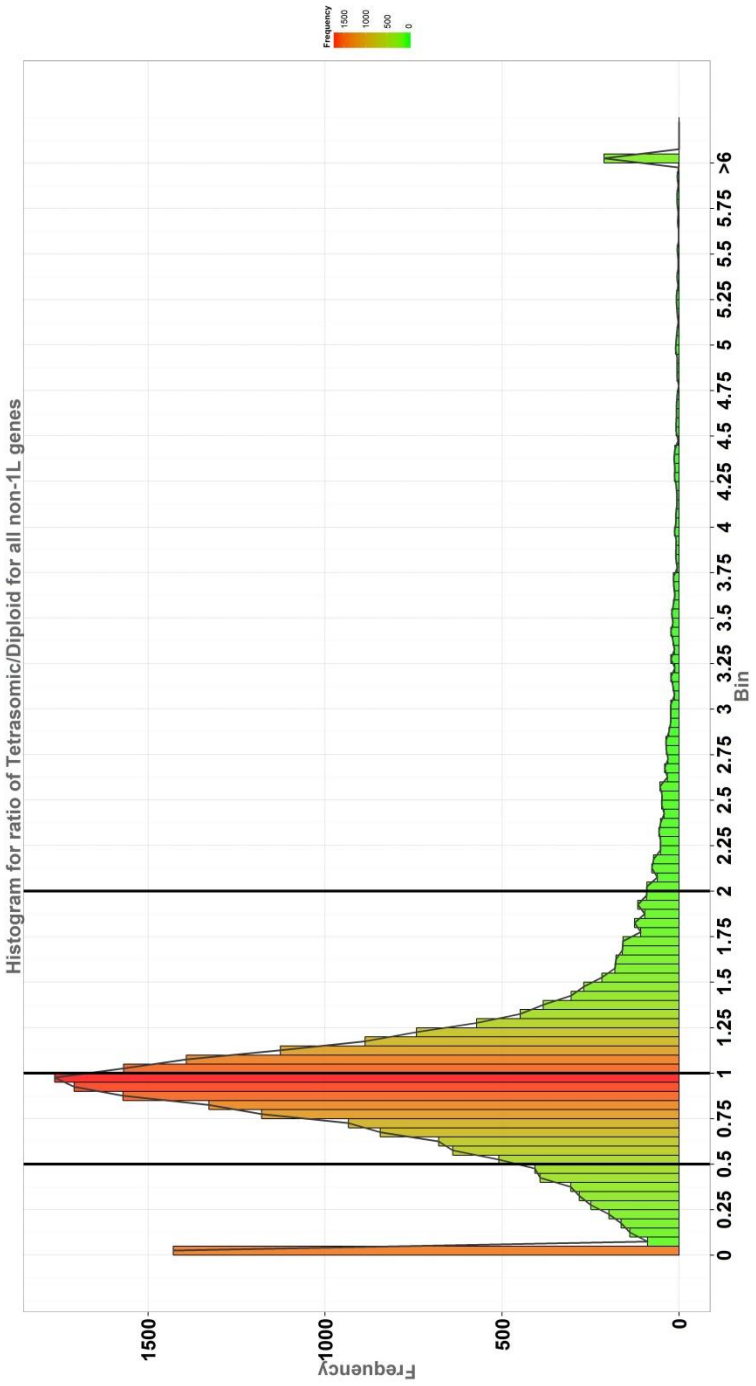


**Figure 4c – Ratio distribution plot for trans genes, tetrasomic/diploid**

For explanation see Figure 4a.

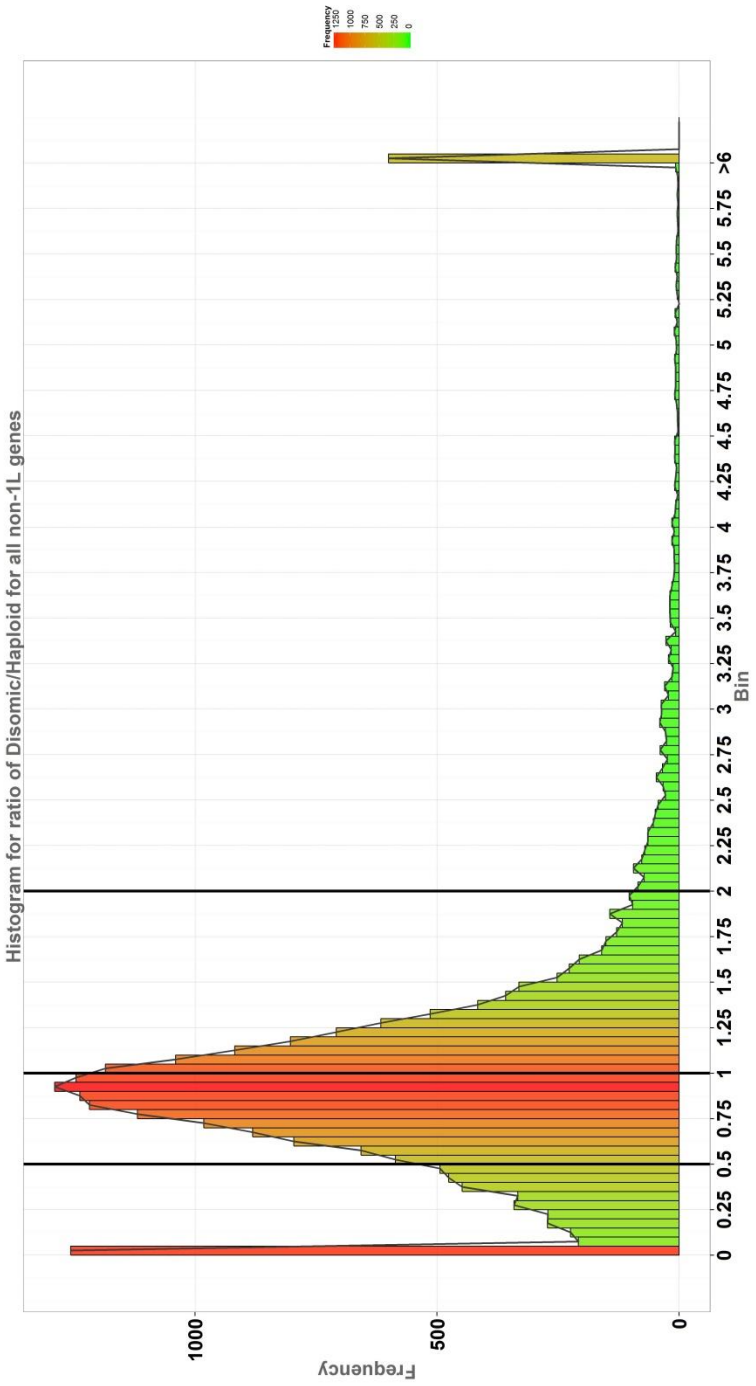


**Figure 4d – Ratio distribution plot for trans genes, disomic/haploid**

For explanation see Figure 4a.

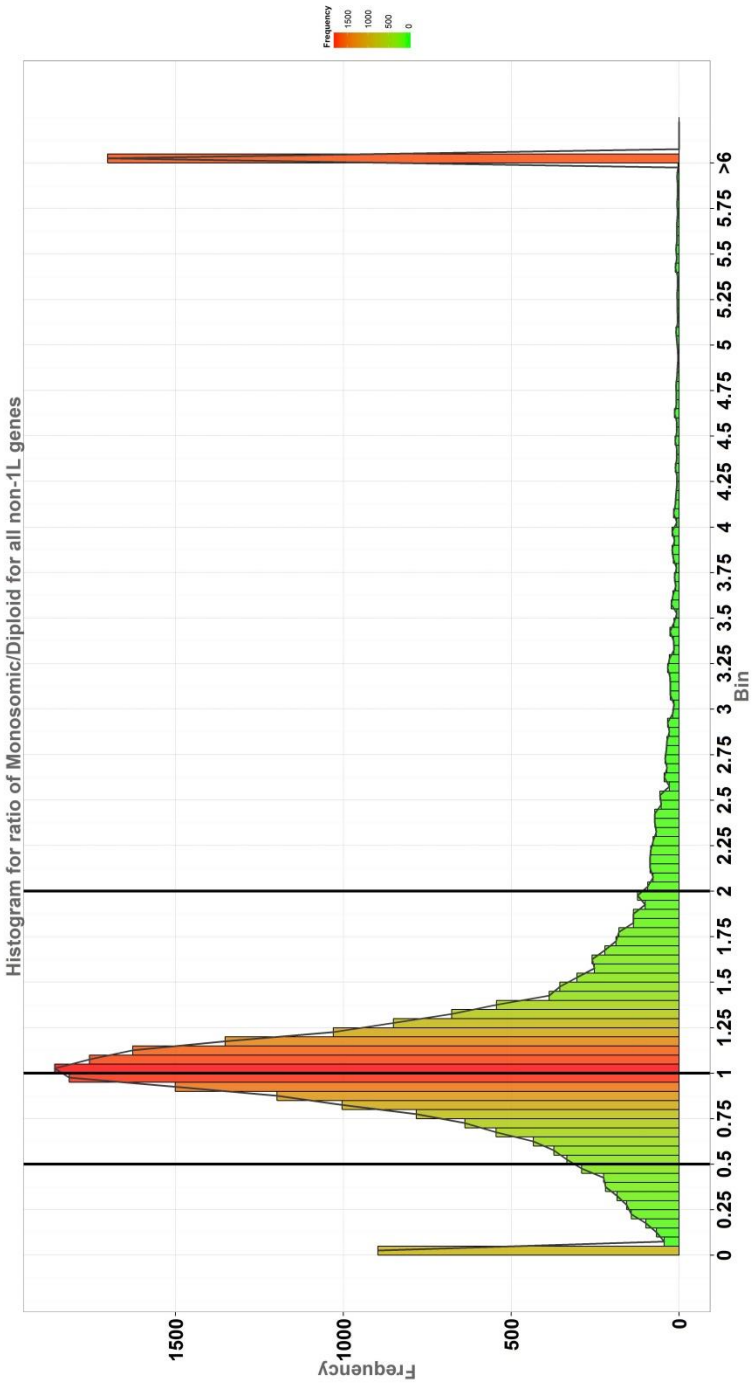


**Figure 4e – Ratio distribution plot for trans genes, monosomic/diploid, second dosage series**

For explanation see Figure 4a.

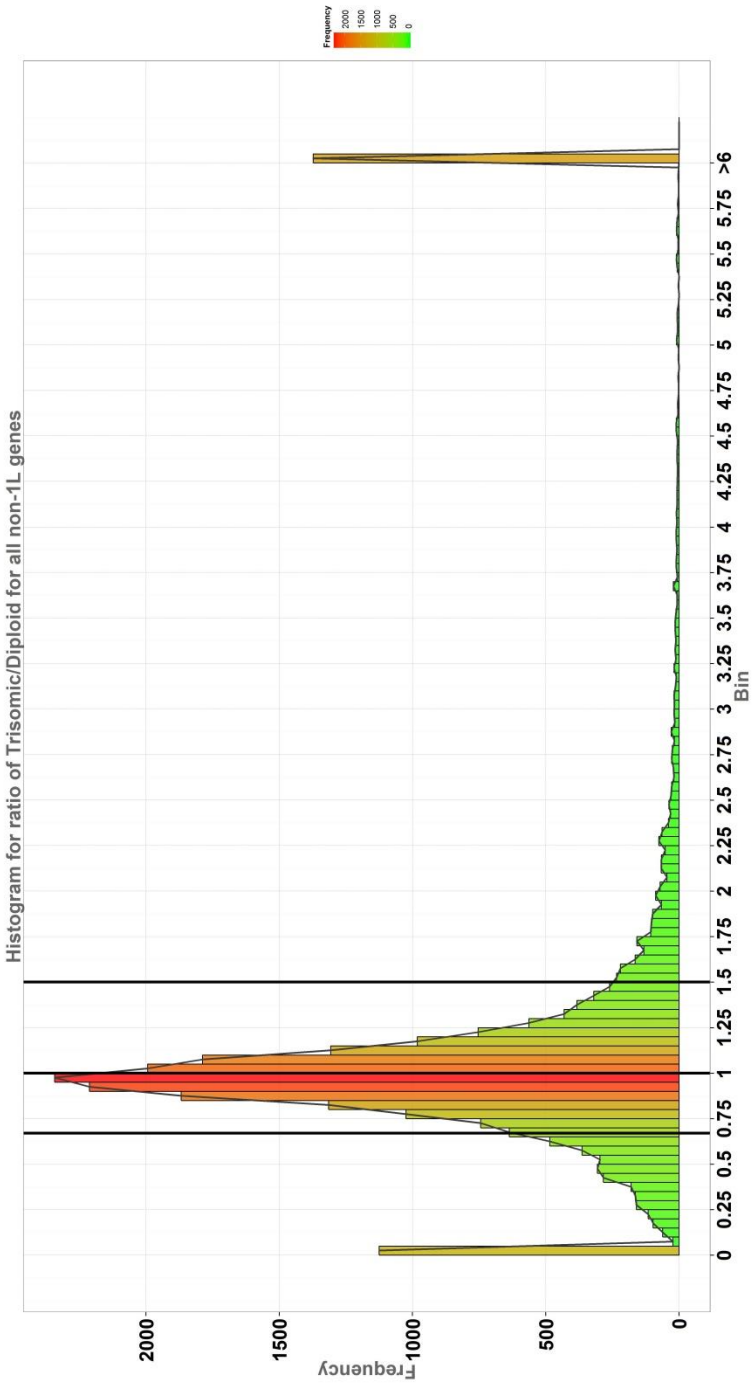


**Figure 4f – Ratio distribution plot for trans genes, trisomic/diploid, second dosage series**

For explanation see Figure 4a.

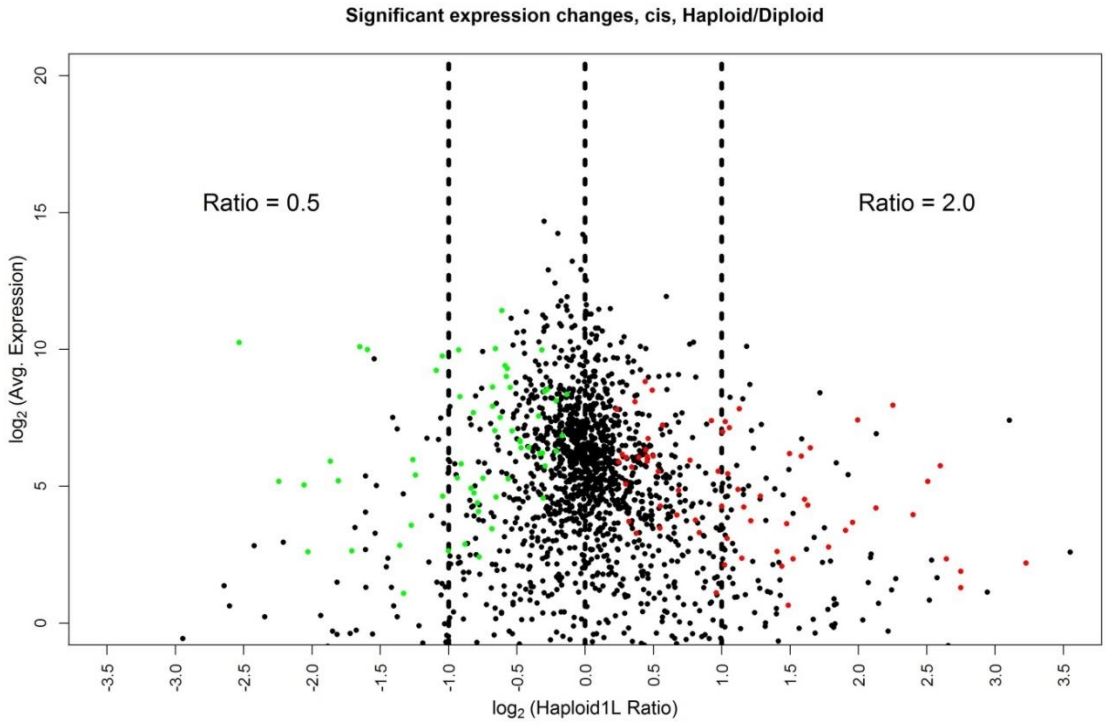


**Figure 5a – Volcano plot for cis genes, haploid/diploid**

The six plots are in the same order as Figures 3 and 4. The x-axis is expression ratio (log2 scale), with the central guide representing a ratio of 1 (no change). The y-axis is FPKM value (log2 scale) in the diploid control. The value given as "Ratio" in the plot indicates the ratio of the nearby vertical guide bar. The guide bar in the middle represents a ratio of 1.00. Green points on the left of the centerline represent genes with a statistically significant decrease in gene expression from euploid to aneuploid. Red points on the right represent an increase in gene expression. The control euploid comparison (a) shows few genes significantly changed, with similar numbers increasing and decreasing. In aneuploids, the prevalence of direct effects in cis is apparent.

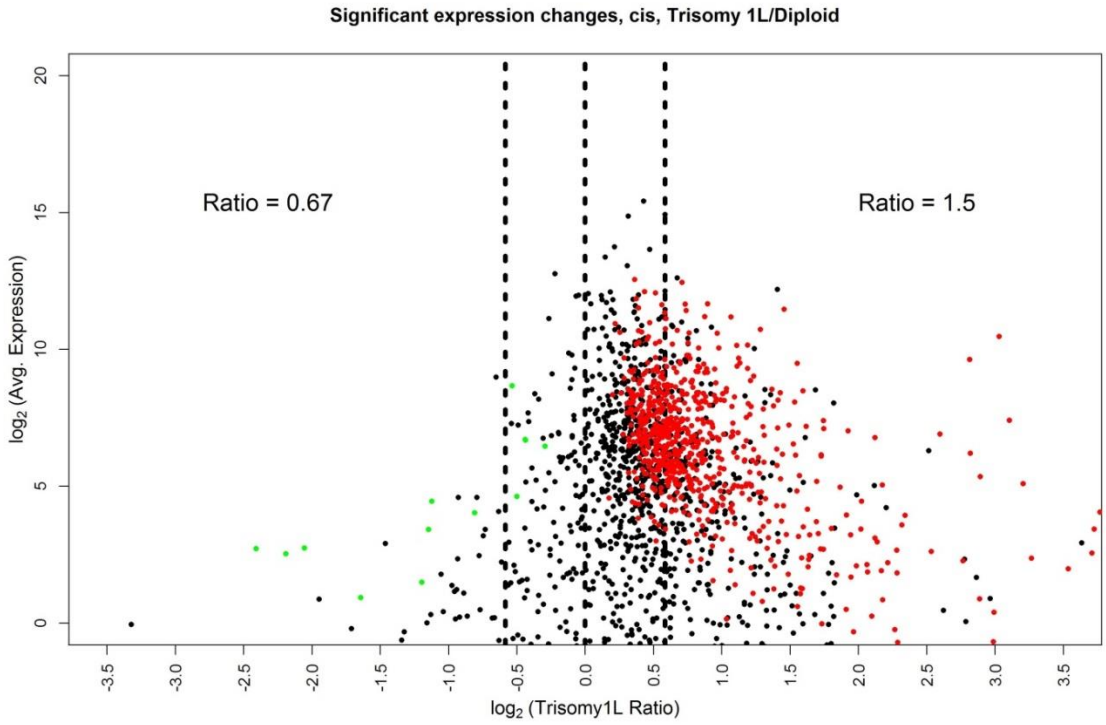


**Figure 5b – Volcano plot for cis genes, trisomic/diploid**

For explanation see Figure 5a.

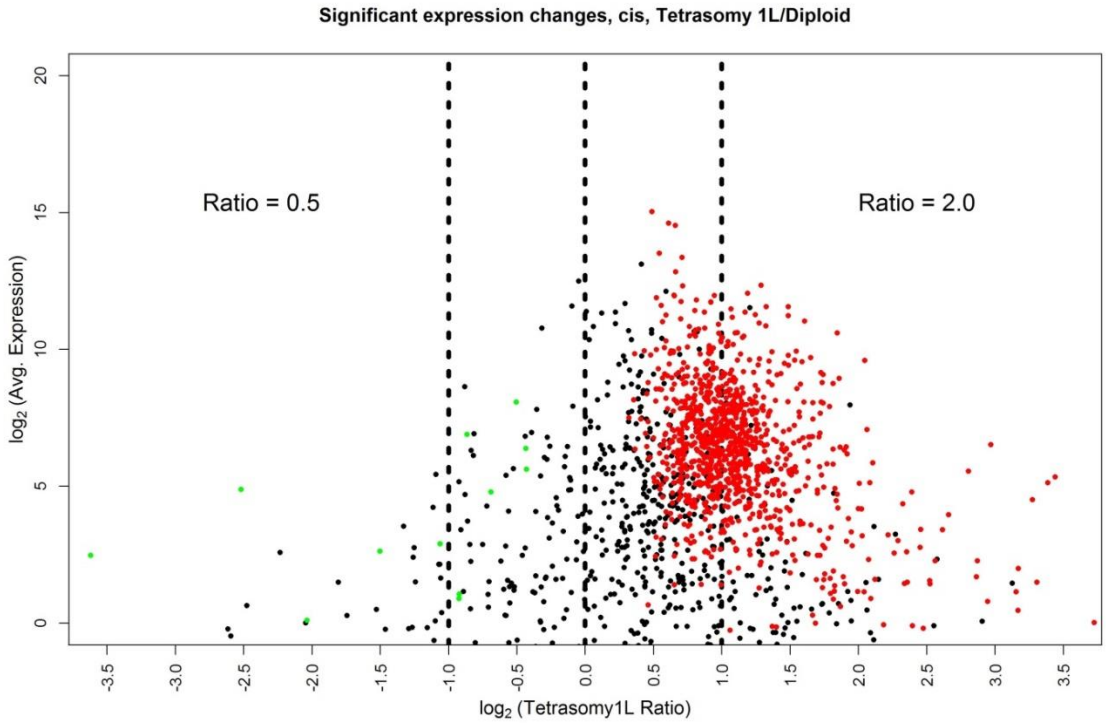


**Figure 5c – Volcano plot for cis genes, tetrasomic/diploid**

For explanation see Figure 5a.

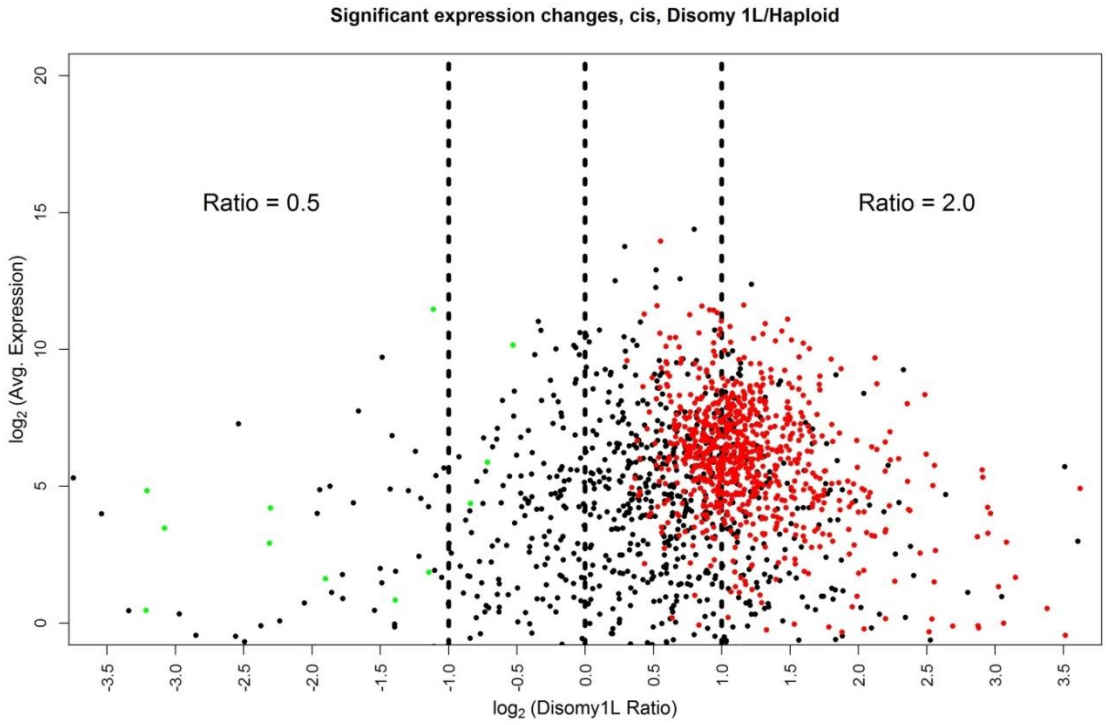


**Figure 5d – Volcano plot for cis genes, disomic/haploid**

For explanation see Figure 5a.

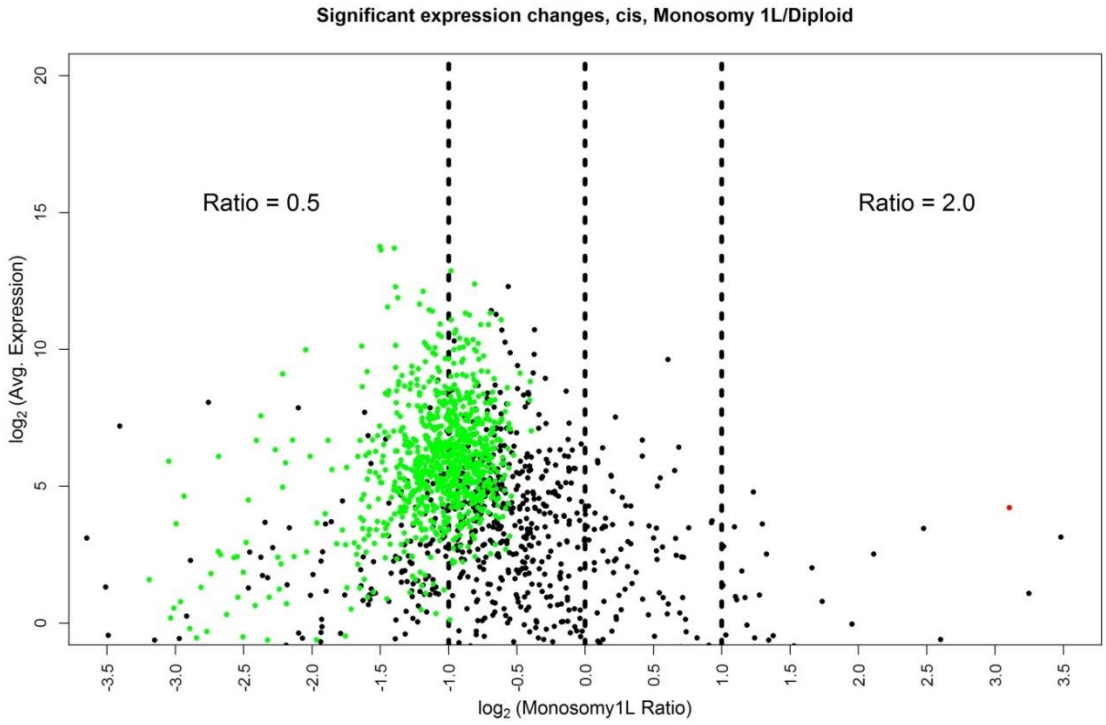


**Figure 5e – Volcano plot for cis genes, monosomic/diploid, second dosage series**

For explanation see Figure 5a.

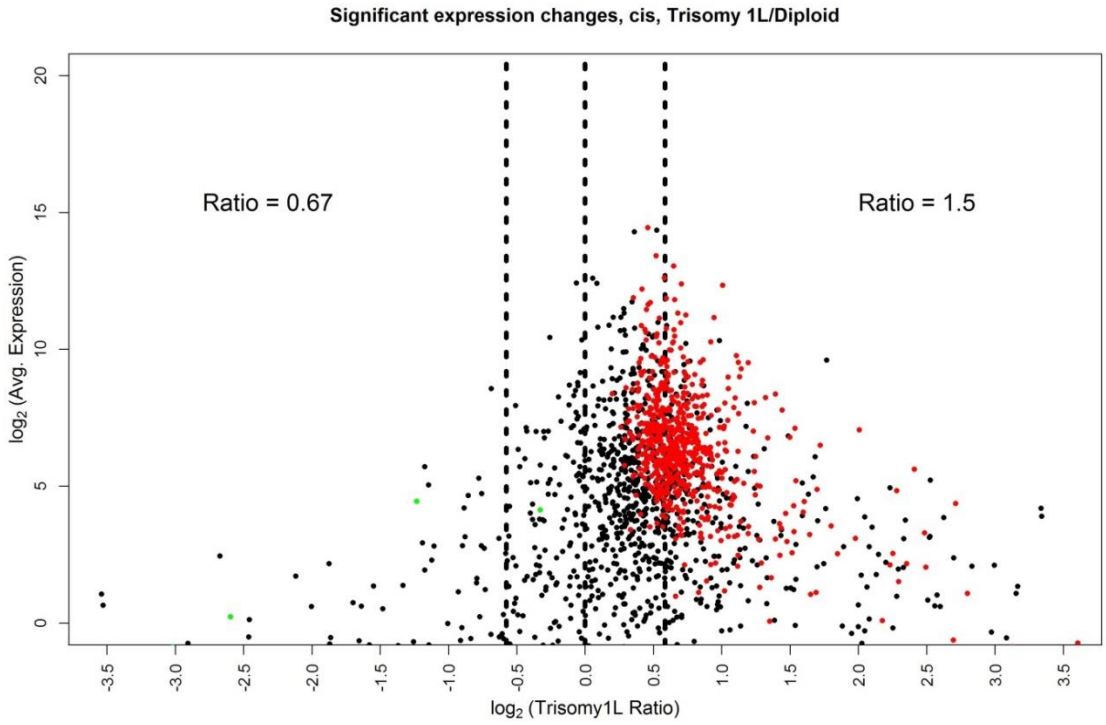


**Figure 5f – Volcano plot for cis genes, trisomic/diploid, second dosage series**

For explanation see Figure 5a.

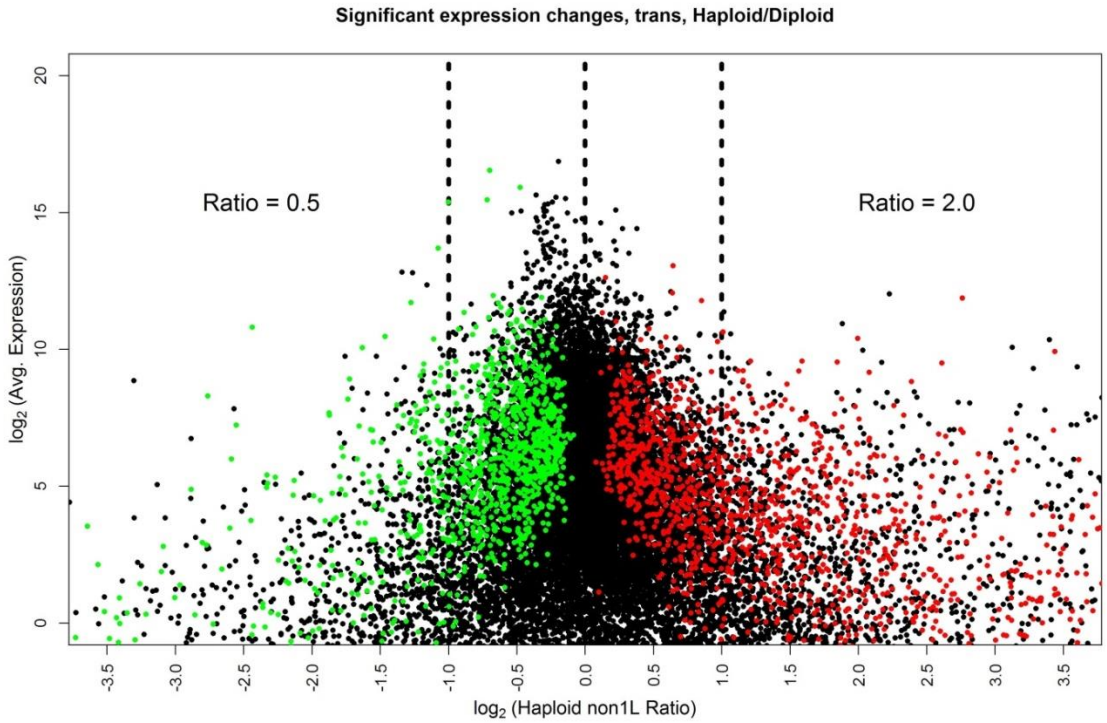


**Figure 6a – Volcano plot for trans genes, haploid/diploid**

Unlike in Figure 5, in aneuploids, the relationship of dosage to expression is more often inverse than direct, shown by the prevalence of significant down-regulated genes (green). Figures 6e and 6f are from the same dosage series, and in both cases the prevalent effect between up- and down-regulation is not apparent. Table 1 indicates both slightly favor up-regulation, though the trisomic/diploid comparison can be seen to have more down-regulated trans genes than the monosomic/diploid comparison.

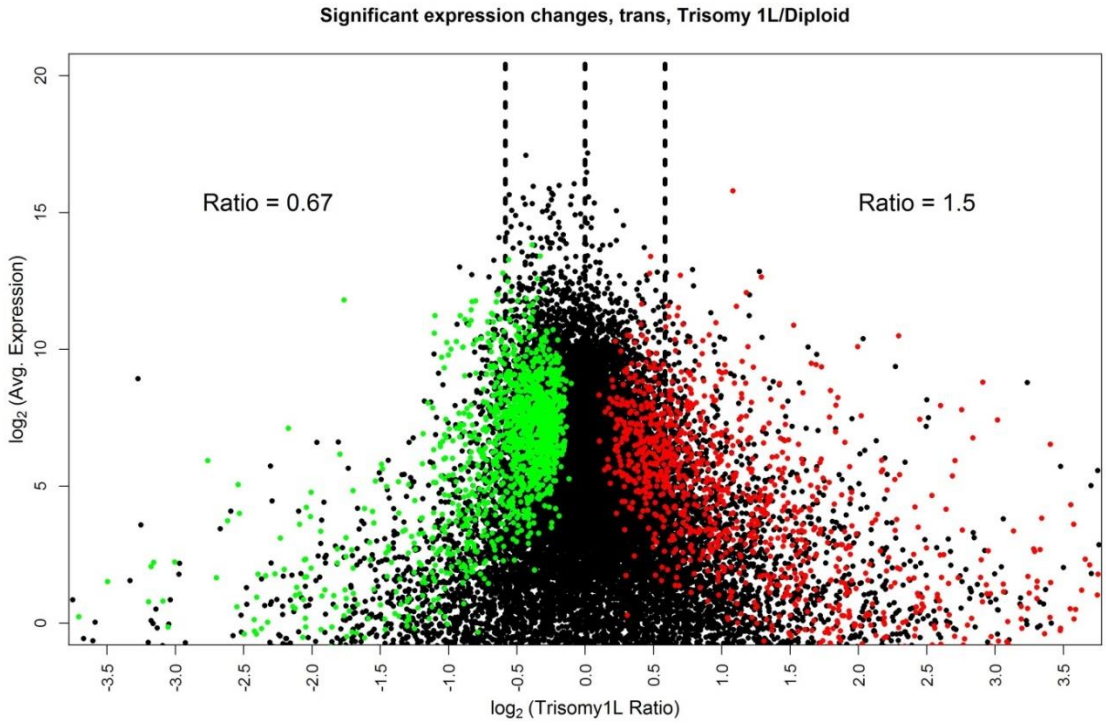


**Figure 6b – Volcano plot for trans genes, trisomic/diploid**

For explanation see Figure 6a.

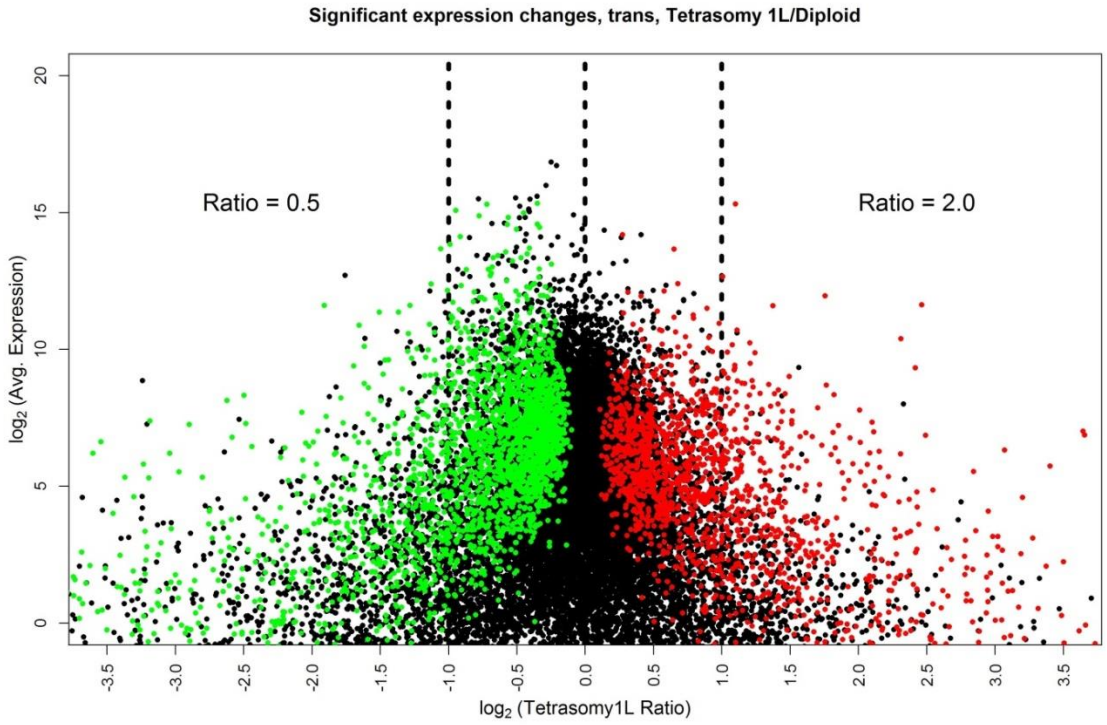


**Figure 6c – Volcano plot for trans genes, tetrasomic/diploid**

For explanation see Figure 6a.

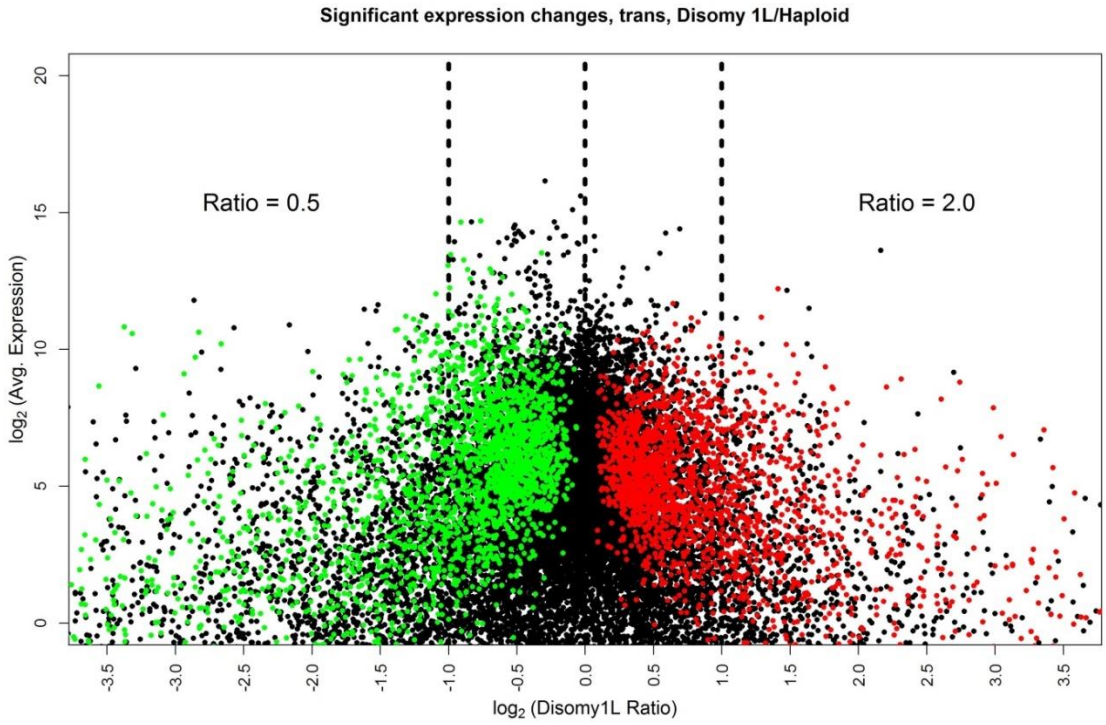


**Figure 6d – Volcano plot for trans genes, disomic/haploid**

For explanation see Figure 6a.

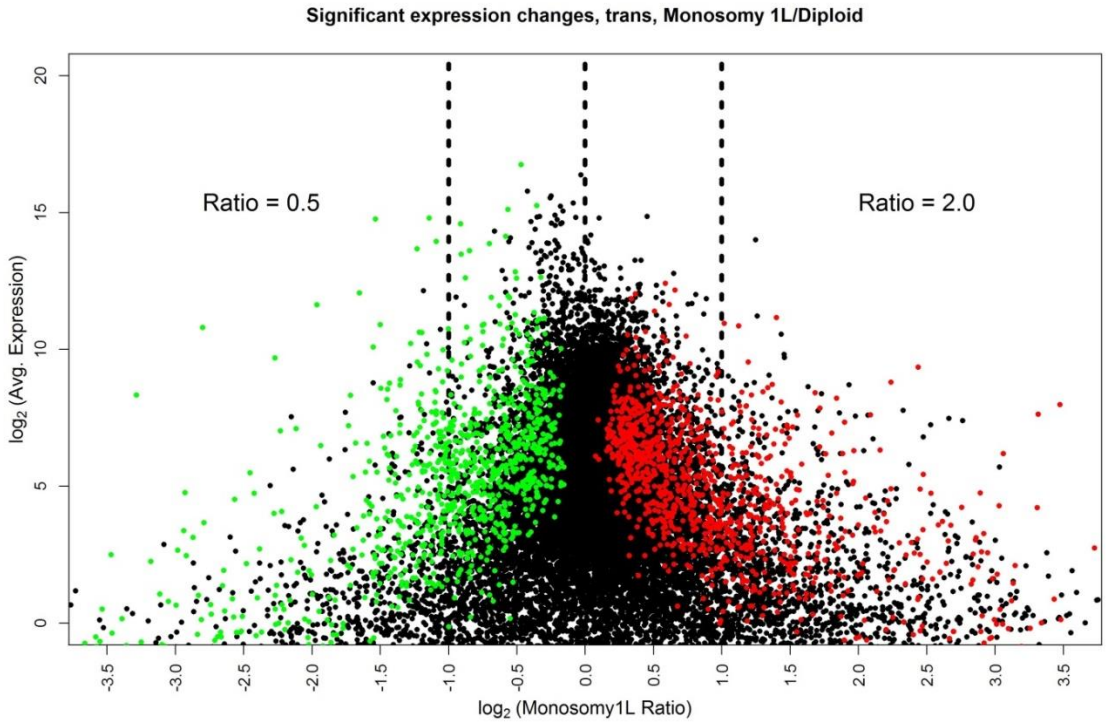


**Figure 6e – Volcano plot for trans genes, monosomic/diploid, second dosage series**

For explanation see Figure 6a.

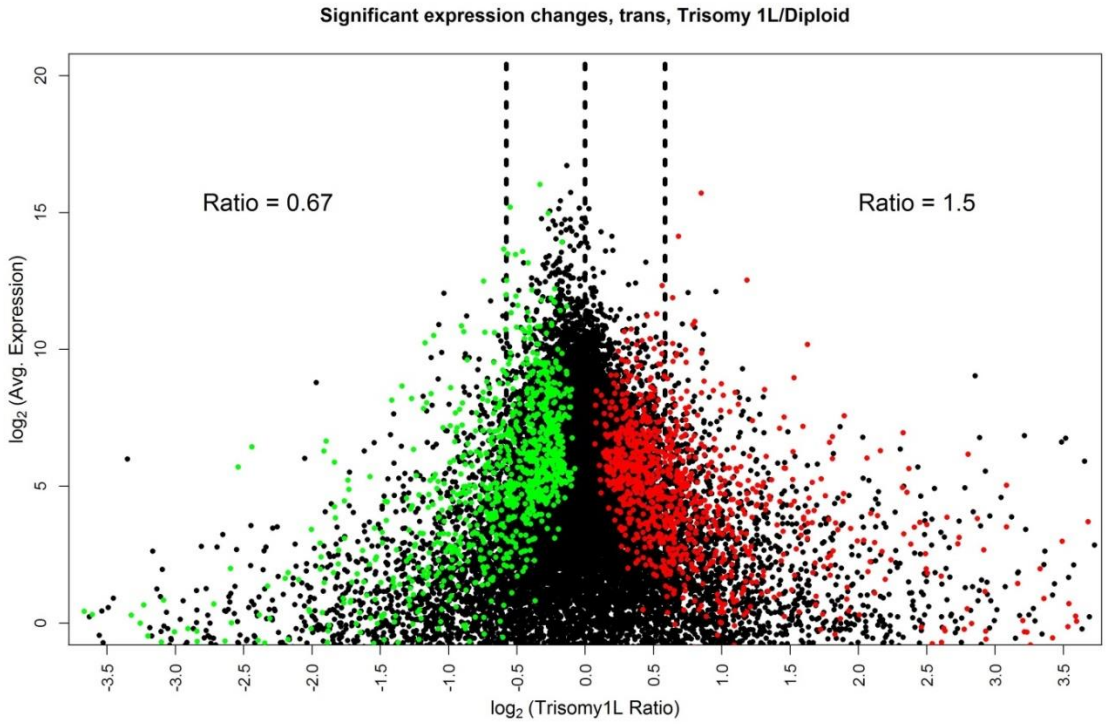


**Figure 6f – Volcano plot for trans genes, trisomic/diploid, second dosage series**

For explanation see Figure 6a.

| Cis comparison | Total genes | Ratio sig. below 1 | Ratio sig. above 1 |
|---|---|---|---|
| hap/dip | 1734 | 64 | 76 |
| tri/dip | 1807 | 22 | 712 |
| tetra/dip | 1729 | 16 | 1050 |
| di/hap | 1669 | 20 | 811 |
| mono/dip (set 2) | 1678 | 909 | 4 |
| tri/dip (set 2) | 1717 | 5 | 673 |
| | | | |
| Trans comparison | Total genes | Ratio sig. below 1 | Ratio sig. above 1 |
| hap/dip | 24657 | 1043 | 1285 |
| tri/dip | 25655 | 1427 | 1069 |
| tetra/dip | 24338 | 2284 | 1506 |
| di/hap | 23598 | 2279 | 1890 |
| mono/dip (set 2) | 24589 | 939 | 1039 |
| tri/dip (set 2) | 24385 | 957 | 1012 |

**Table 1 – Summary of t-test results**

For haploid (compared to diploid control), trisomy 1L (compared to diploid), tetrasomy 1L (compared to diploid), disomy 1L (compared to haploid), monosomy 1L (compared to diploid), and trisomy 1L (compared to diploid), summary table of gene expression effects. Cis (1L) and trans (non-1L) effects are separated. Total number of genes differs due to removal of data points on fringes of ratio distribution table.

| Comparison of cis ratios | p-value |
|---|---|
| haploid/diploid vs. trisomic/diploid | 2.20E-16 |
| haploid/diploid vs. tetrasomic/diploid | 2.20E-16 |
| haploid/diploid vs. disomic/haploid | 2.20E-16 |
| disomic/haploid vs. tetrasomic/diploid | 2.20E-16 |
| | |
| Comparison of trans ratios | p-value |
| haploid/diploid vs. trisomic/diploid | 5.33E-15 |
| haploid/diploid vs. tetrasomic/diploid | 2.20E-16 |
| haploid/diploid vs. disomic/haploid | 2.20E-16 |
| disomic/haploid vs. tetrasomic/diploid | 2.20E-16 |

**Table 2 – KS test results for comparison of ratios**

Kolmogorov-Smirnov (KS) tests were performed in order to compare expression ratio sets. All comparisons resulted in extremely low p-values, implying each comparison consists of two significantly different ratio sets. 2.20E-16 is the lowest value produced by the R programming language for this test.

# Chapter 3 – Functional analysis of gene dosage effects in a series of aneuploids for maize chromosome arm 1L

## Introduction and methods

The dynamics of gene regulation, and therefore dosage sensitivity, are potentially different for various functional groups of genes. As an example, transcription factors (TFs) may regulate both other TFs and other functional groups, while a metabolism-involved gene may not directly regulate expression of any others. TFs may act in concert with other genes as part of a complex, making the relative dosage of component parts a factor in its regulatory function. A dosage series of chromosome arm 1L in maize was analyzed for the differential impacts of genome imbalance on genes assigned to different functional groups. Ratio distributions were produced for a number of gene groups in both monosomic/diploid and trisomic/diploid comparisons. Gene sets consisted of the same genes in both dosage levels, and excluded those with a structural locus on the dosage-varied chromosome arm – therefore, all effects presented are defined as trans. Gene lists for each functional set were provided by the Plant GeneSet Enrichment Analysis Toolkit (PlantGSEA) (Yi *et al.* 2013). The dataset consists of a catalogue of gene ontology (GO) terms, with a list of maize genes assigned to each group either experimentally or through computational prediction. Only genes present in a given list and expressed in diploid maize plants were plotted in the ratio distribution figures.

**Results and discussion**

The first gene group to be considered is the set of transcription factors (GO: 0004879). The results from monosomic and trisomic comparisons are reflective of the impact of inverse relationships between cis gene dosage and trans gene expression. In the monosomic/diploid comparison, there is a large peak centered over 1.0, with a substantial portion falling to either side, implying a wide range of effects on members of the gene set. It is also of note that the number of genes increased in expression greater than 6-fold (the spike on the right-hand side of the plot) is greater that the number with an equivalent decrease in expression (the spike on the left-hand side of the plot). The portions to the right of 1.0, including the right-hand spike, have an inverse relationship to the dosage on 1L. In the trisomic/diploid comparison, there is a central peak near 1.0, as in the monosomic. However, there is a noticeable increase in the number of genes to the left of 1.0, and a corresponding decrease in the number of genes to the right. Regarding the spikes that occur on either end of the plot, the low spike is greater than the high spike, in contrast to the monosomic. Both dosage of 1L and the direction of expression changes in trans are reversed in the trisomic compared to the monosomic, preserving the prevalence of the inverse effect. A Kolmogorov-Smirnov (KS) test comparing the set of monosomic ratios with the set of trisomic ratios does not indicate the two sets are significantly different from each other. KS tests for all gene groups are summarized in Table 3.

A second functional group to be considered is photosynthesis-related genes (GO:0015979). This group is of direct relevance to the leaf tissue from which RNA was extracted in the maize 1L dosage series. The results are noticeably distinct from those of the TF gene list, particularly in the monosomic/diploid comparison. The ratio distribution

plot shows a general decrease of expression in the monosomic, directly correlating with the decreased dosage of 1L. The trisomic/diploid comparison also shows a general decrease of gene expression, in keeping with an inverse effect. Comparison of the monosomic and trisomic sets of ratios with a KS test indicates a significant difference between them. The loss of consistency with the generally-observed inverse trans effect in monosomics, but not trisomics, may be a reflection of molecular kinetics. In the TF group, listed genes may occur anywhere in a regulatory network, which is to say a gene may have many or few regulators that determine its expression. As the photosynthesis genes are functional rather than regulatory genes, they may be weighted towards the end of regulatory pathways, and therefore depend on many regulators which determine their level of expression. One possible explanation for decreasing expression in both monosomics and trisomics is that with an increasing number of regulatory components, genomic imbalance in either direction becomes more disruptive to the system. Alternatively, it may be the case that photosynthesis-related genes are especially reflective of the general health of the plant.

Three additional gene sets were considered in terms of distinct reactions to aneuploidy. Genes involved in the G protein signaling pathway (GO:0007186) behaved similarly to the transcription factor set. Both groups have extensive regulatory impacts on other genes, and are highly dependent on interactions with other molecules to complete their functions. The sets of monosomic and trisomic ratios are not found to be significantly different from each other, as measured by a KS test. A set of ribosomal genes (GO:0003735) was also considered. The ratio distribution plots do not indicate a large scale response to either aneuploid condition, and the two sets are not found to be

significantly different according to the KS test. Genes comprising the proteasome complex (GO: 0005839) were found to be significantly different by the KS test, a feature this gene group shares with the set of photosynthesis genes. A relatively small sample size may have contributed to this output, but it is also possible that this group shares some basic properties of expression regulation with the photosynthesis genes, and perhaps others that do not have a functional focus on other genes' regulation.

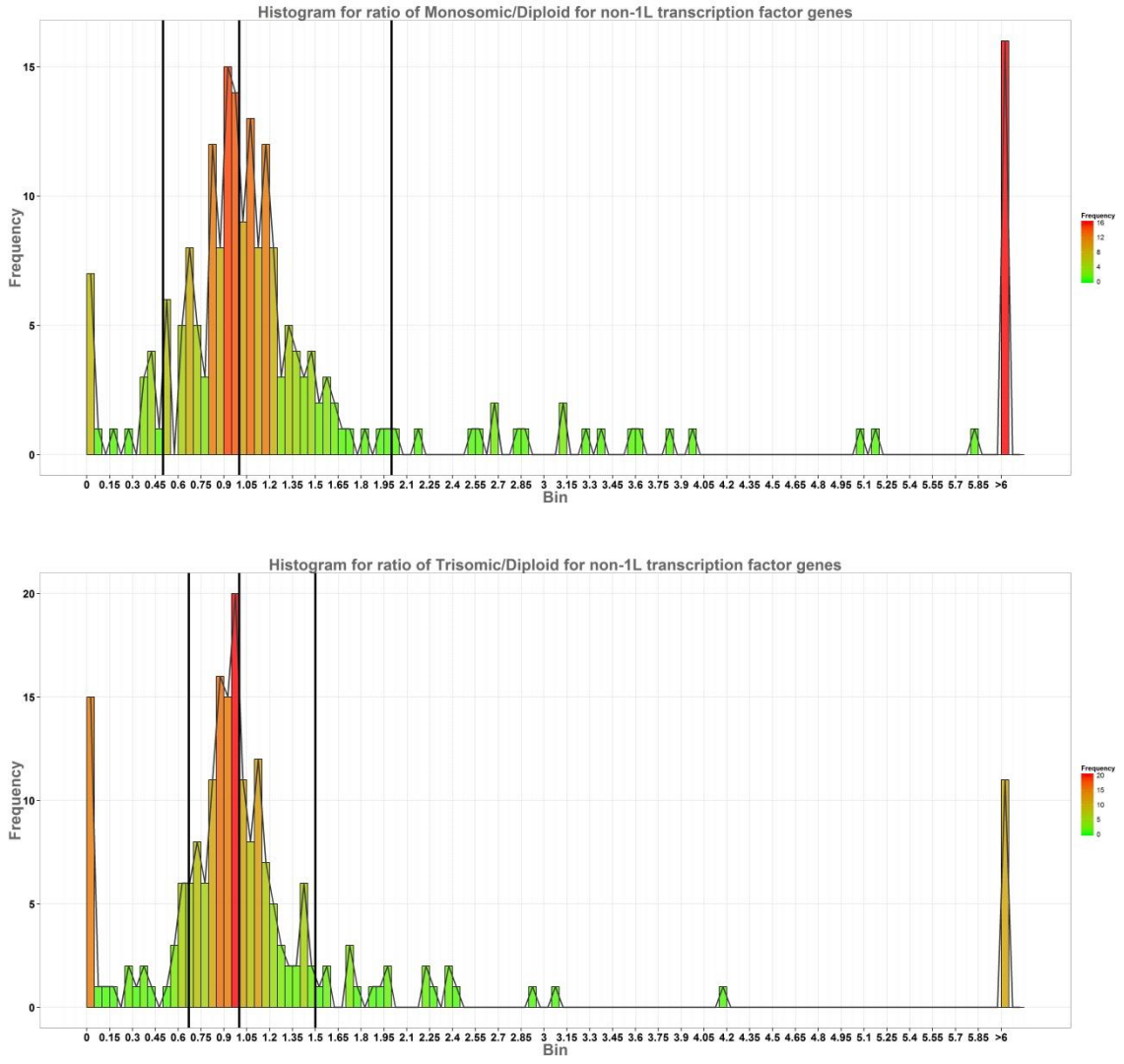


**Figure 7a – Ratio distribution plots for transcription factor genes**

The top plot concerns ratio distributions for trans genes in a monosomic plant. The bottom plot represents the same set of genes in a trisomic plant. Vertical guides in the top plot are placed at 0.50, 1.00, and 2.00, while in the bottom plot they are placed at 0.67, 1.00, and 1.50.

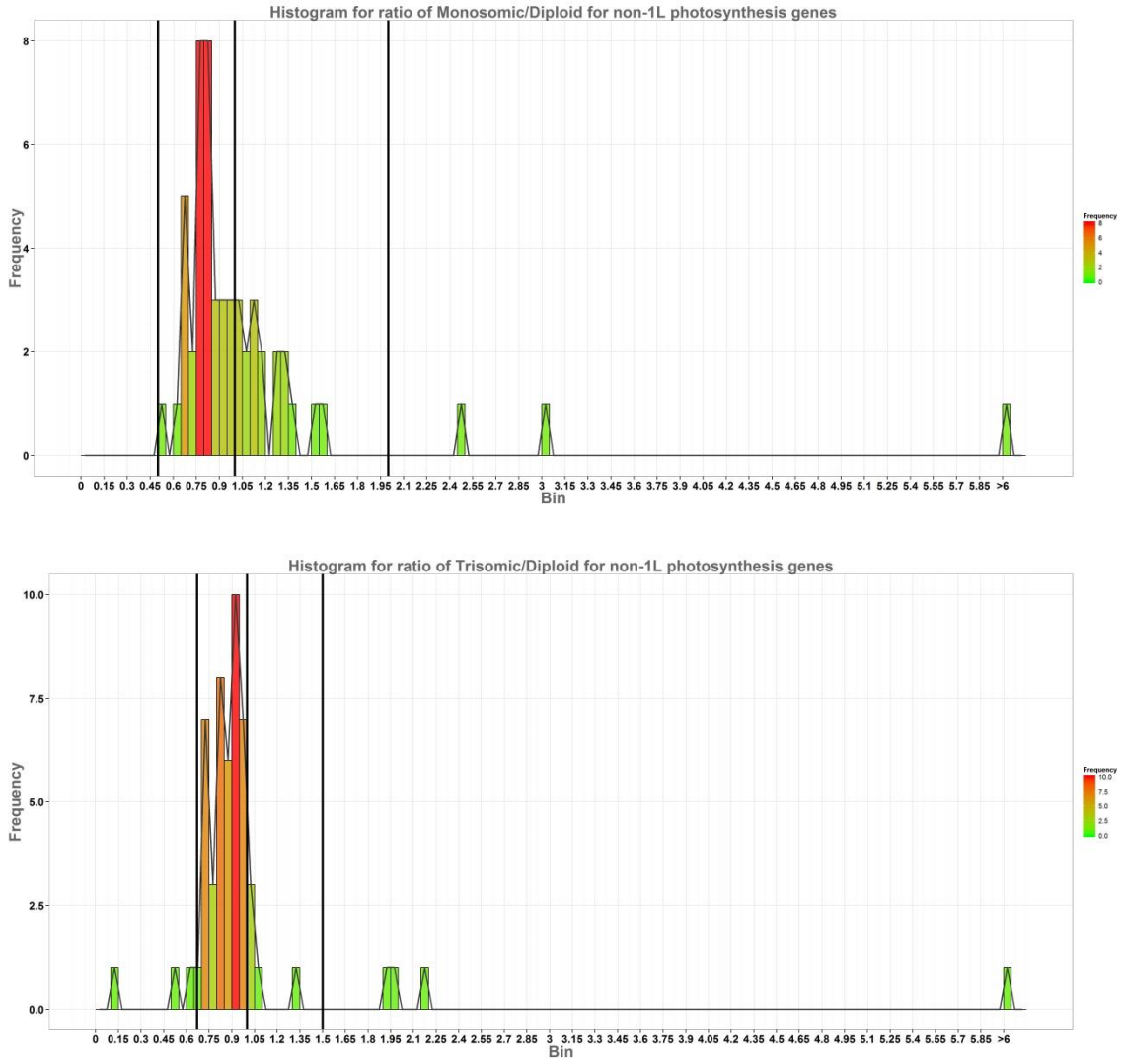


**Figure 7b – Ratio distribution plots for photosynthesis genes**

For explanation see Figure 7a.

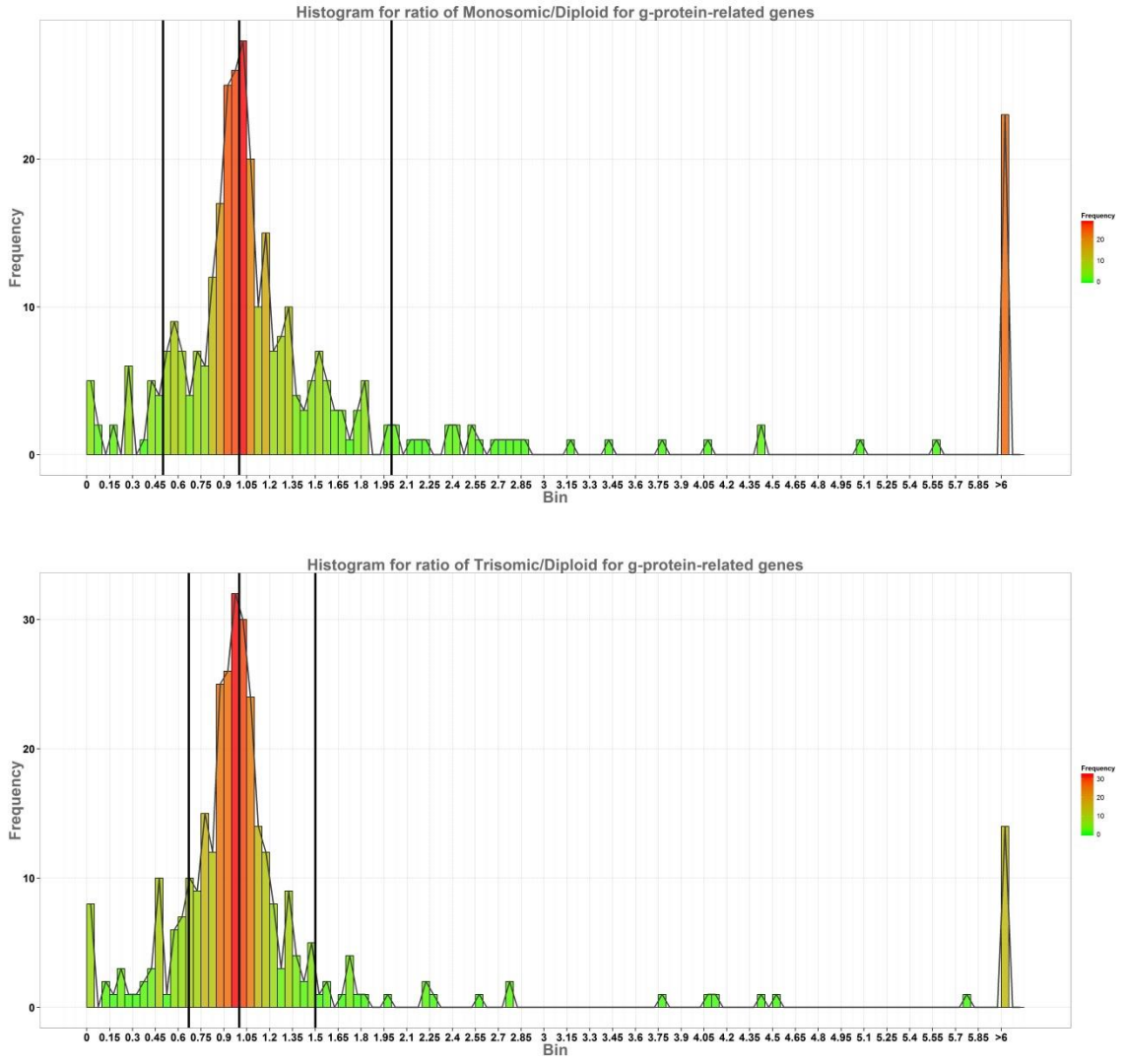


**Figure 7c – Ratio distribution plots for G protein-related genes**

For explanation see Figure 7a.

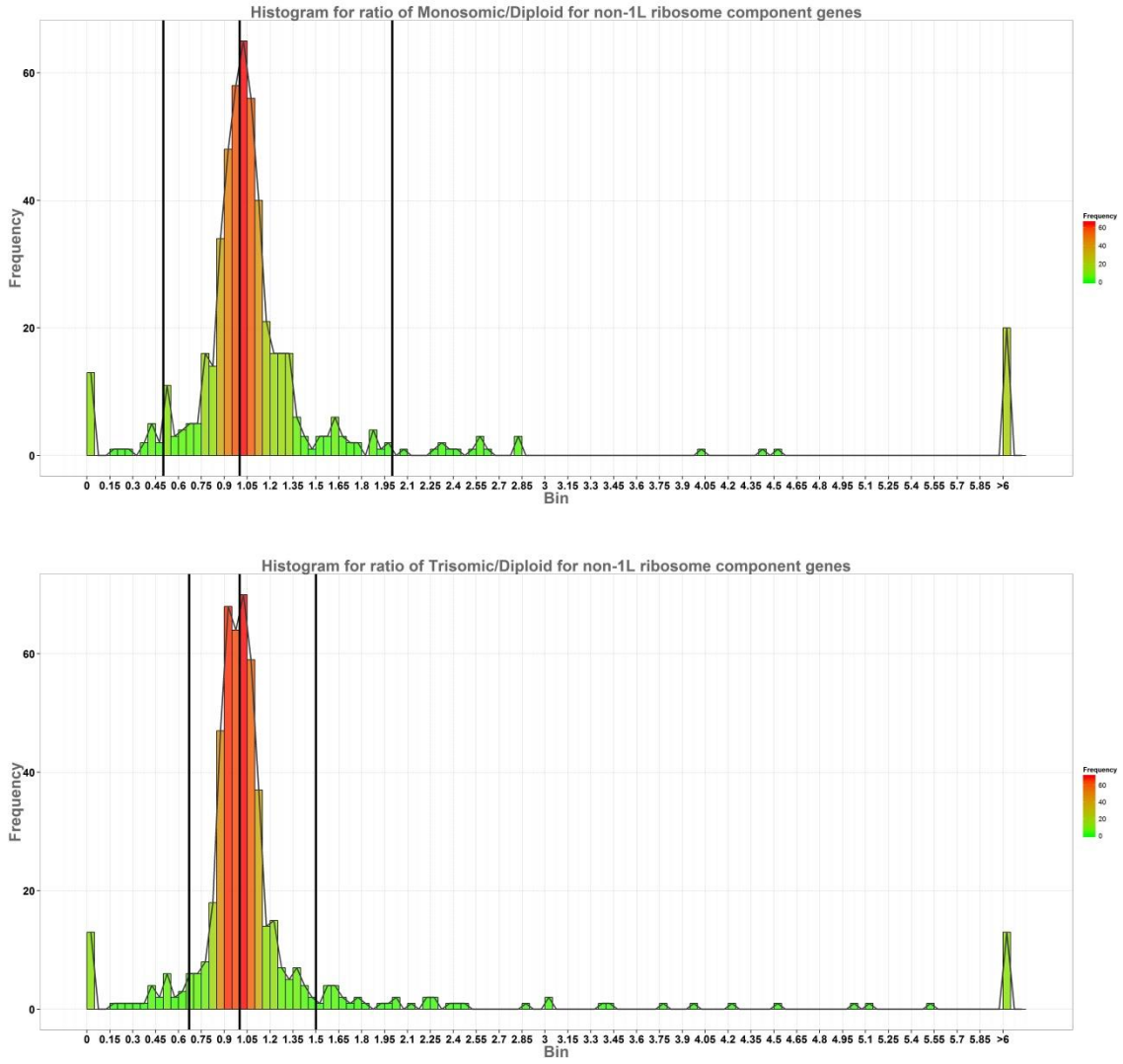


**Figure 7d – Ratio distribution plots for ribosome component genes**

For explanation see Figure 7a.

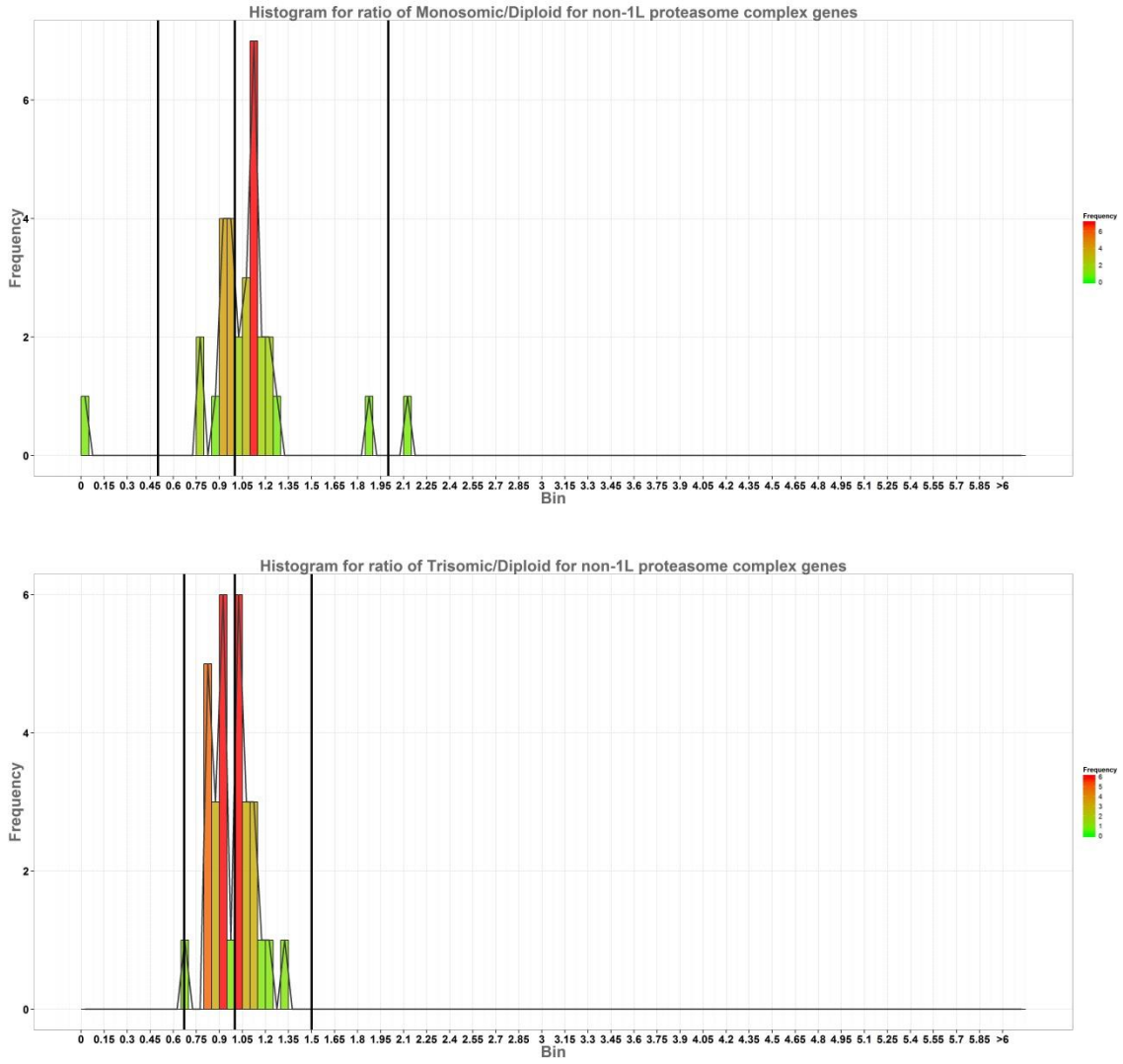


**Figure 7e – Ratio distribution plots for proteasome complex genes**

For explanation see Figure 7a.

| Gene list | p-value, monosomic vs. trisomic ratios |
|---|---|
| Transcription factors | 0.01594 |
| Photosynthesis | 0.1389 |
| G protein signaling pathway | 0.001809 |
| Ribosome components | 0.008948 |
| Proteasome complex | 0.1474 |

**Table 3 – KS test results for gene functional groups**

Kolmogorov-Smirnov (KS) tests were performed for each of the five listed gene function groups, comparing the expression ratios of monosomic/diploid and trisomic/diploid. A p-value above 0.05 indicates the two ratio sets are significantly different from each other. A p-value below 0.05 indicates the two sets are similar enough that one could be a subset of the other; that is, the expression changes in a similar way at both dosage levels.

# Chapter 4 – The Effects of Genomic Imbalance on Organisms of Different Kingdoms

## Introduction

Aneuploidy has been studied in numerous organisms, with an eye towards understanding the mechanism most directly responsible for its extensive effects on phenotype. A number of possible explanations have the power to account for its impacts on human health in conditions such as Down syndrome, its absence as a detectable phenomenon on an evolutionary scale, and its relative severity compared with polyploidy. Among these hypotheses, the most direct is that additional proteins produced from some of the genes in a hyperploid condition are toxic at all levels above normal. This would allow for direct cis effects to account for phenotypic effects, and indeed there is evidence of extensive direct cis effects at a transcriptional level. However, two factors complicate this model: it does not account for the greater severity of aneuploidy relative to polyploidy; and it does not account for the presence of trans effects on gene expression. In fact, the presence of trans effects is not universally agreed upon in the literature. Work presented in this dissertation suggests trans effects, both direct and inverse, are extensive, and affected genes would need to be considered as a potential source of phenotypic effects. However, if a study proceeds from the assumption that trans effects of aneuploidy are rare, then trans gene expression can be used to normalize expression data from control euploids and experimental aneuploids; trans gene expression is presumed to be the unchanged.

A second hypothesis suggests the method presented in this dissertation, and is referred to as the gene balance hypothesis. This model does not assume trans gene expression is unaffected, and the RNA expression data presented is normalized only to transcriptome size, which is itself the result of the technical constraints of RNA sequencing. Possible changes to transcriptome size as a whole are not detectable with this method. Presuming proportional cis effects and no trans effects occurred, increased expression by cis genes may still produce an apparent decrease of trans gene expression. For example, trisomy of 1L, accounting for 5% of the maize genome and present at 150% of its normal level, would produce a transcriptome of 102.5% of its normal size, and trans genes would appear to be depressed to 97.5% of their normal level. Decreases substantially below that level, however, can be identified as genuine inverse trans effects, as is seen in the maize data presented here.

**Methods**

Data presented in a number of published articles on the topic of gene expression in aneuploids was reanalyzed to allow for the possibility of trans effects. The two studies presented here were conducted with a mouse model (WILLIAMS *et al.* 2008) and a yeast model (TORRES *et al.* 2007), respectively. The mouse model analyzed gene expression data in four different trisomies as compared to a diploid control (WILLIAMS *et al.* 2008). These were trisomies of chromosomes 1, 13, 16, and 19. Notably, mouse chromosome 16 is orthologous to human chromosome 21, and mouse trisomy 16 is in fact used to model the mechanisms of Down syndrome (human trisomy 21). The yeast model utilized a set of disomies for the majority of chromosomes as compared to a haploid control (TORRES

*et al.* 2007). Some comparisons of multiple disomies (that is, disomic for two chromosomes) against a haploid were also conducted. Microarray data for both studies was presented in pre-normalized form as supplemental material to the relevant articles.

Mouse gene expression data was analyzed to determine expression ratios between trisomy and diploid for each gene. Ratios were then plotted with histograms to display the number of genes expressed in a given range of ratios. These histograms are referred to as ratio distribution plots. Ratio distribution plots were not produced in either mouse or yeast in the original publications of the data, so the trends in gene expression visualized here would not have been detected. These figures are distinct from the ratio distribution plots presented for maize, in that there is a prominent spike at or near 1.00 in each of the mouse plots, while in the maize plots there were spikes at the two extreme ends, at <0.05 and >6.00.

**Results and discussion**

Ratio distributions were produced separately for cis and trans genes, shown in Figure 8. Cis genes are defined here as those with a locus on the triplicate chromosome, meaning the cis set consists of different genes in each trisomy. Trans genes are defined as all those in the rest of the genome, which remain in a diploid condition. Cis genes in all four trisomies show similar effects, the movement of a large number of genes towards a ratio of 1.50, proportional with the increase in gene dosage. The plots show that the effect acts differently on different genes, however. Only a portion of a given set is actually expressed at a ratio close enough to 1.50 to form a visible peak there; a large number of genes are found between 1.00 and 1.50, indicating a partial buffering of the direct dosage

effect. Likewise, some genes are overexpressed compared to their dosage; some are dosage compensated, showing little change from their diploid level of expression; and a smaller number of genes show a decrease of expression in the trisomy. The prevalence of these effects is masked if the summary statistics of expression ratios are simplified to mean or median changes of expression.

Trans effects are less extensive than cis effects, consistent with data presented for maize aneuploids. However, there are noticeable deviations from 1.00, and these occur in a non-symmetrical fashion. In each of the four trisomies, inverse trans effects (a decrease in expression) are more prevalent than direct trans effects (an increase in expression).

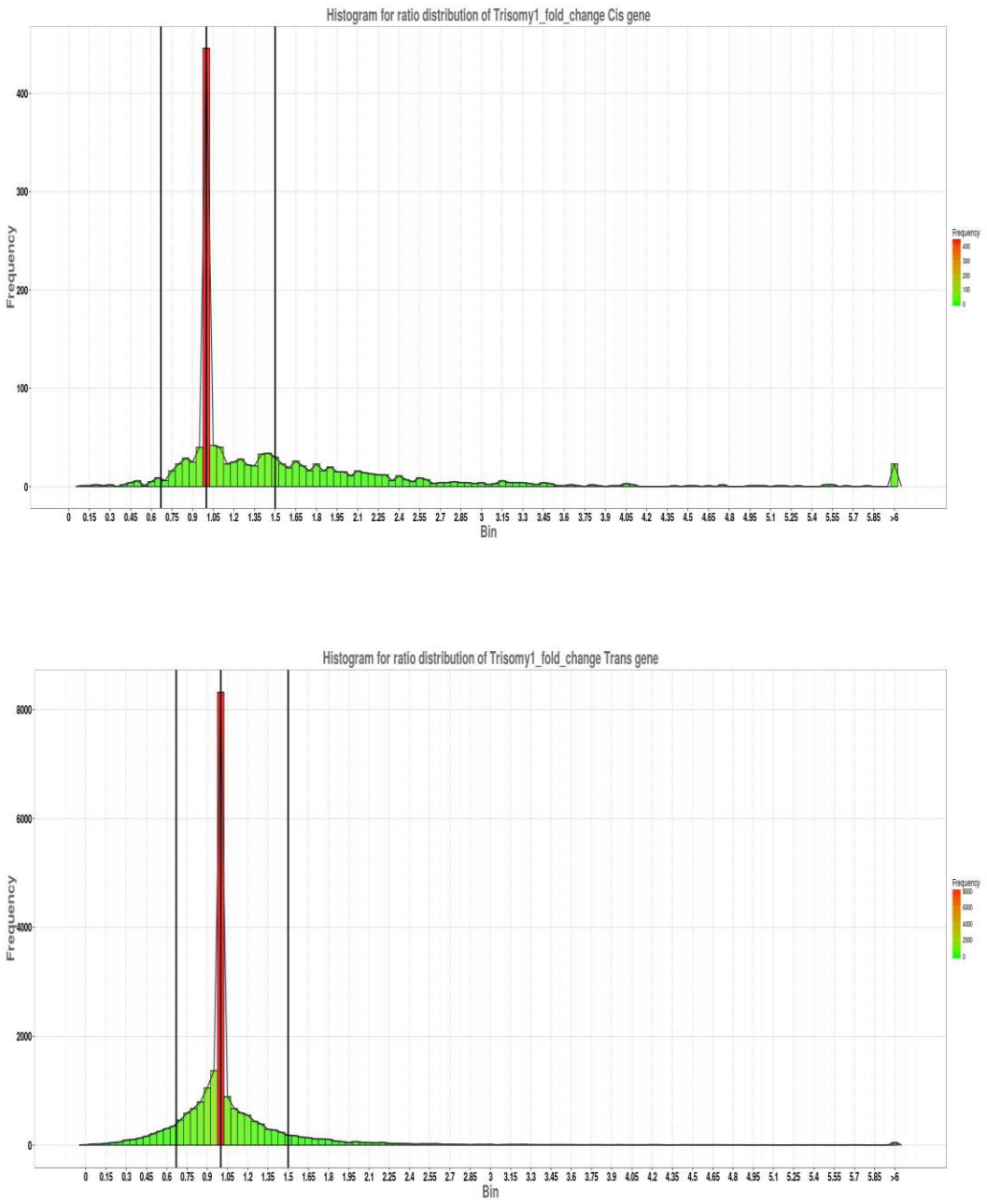


**Figure 8a – Ratio distribution plots for trisomy 1/diploid**

The top plot displays the ratio distribution for cis genes, and the bottom plot for trans genes. The relative distribution of ratios to the left or to the right is distinct between the two plots. Ratios for cis genes trend to the right, indicating an increase of expression relative to a diploid, while for trans genes they trend to the left, indicating a decrease of expression. Vertical guides are placed at 0.67, 1.00, and 1.50.

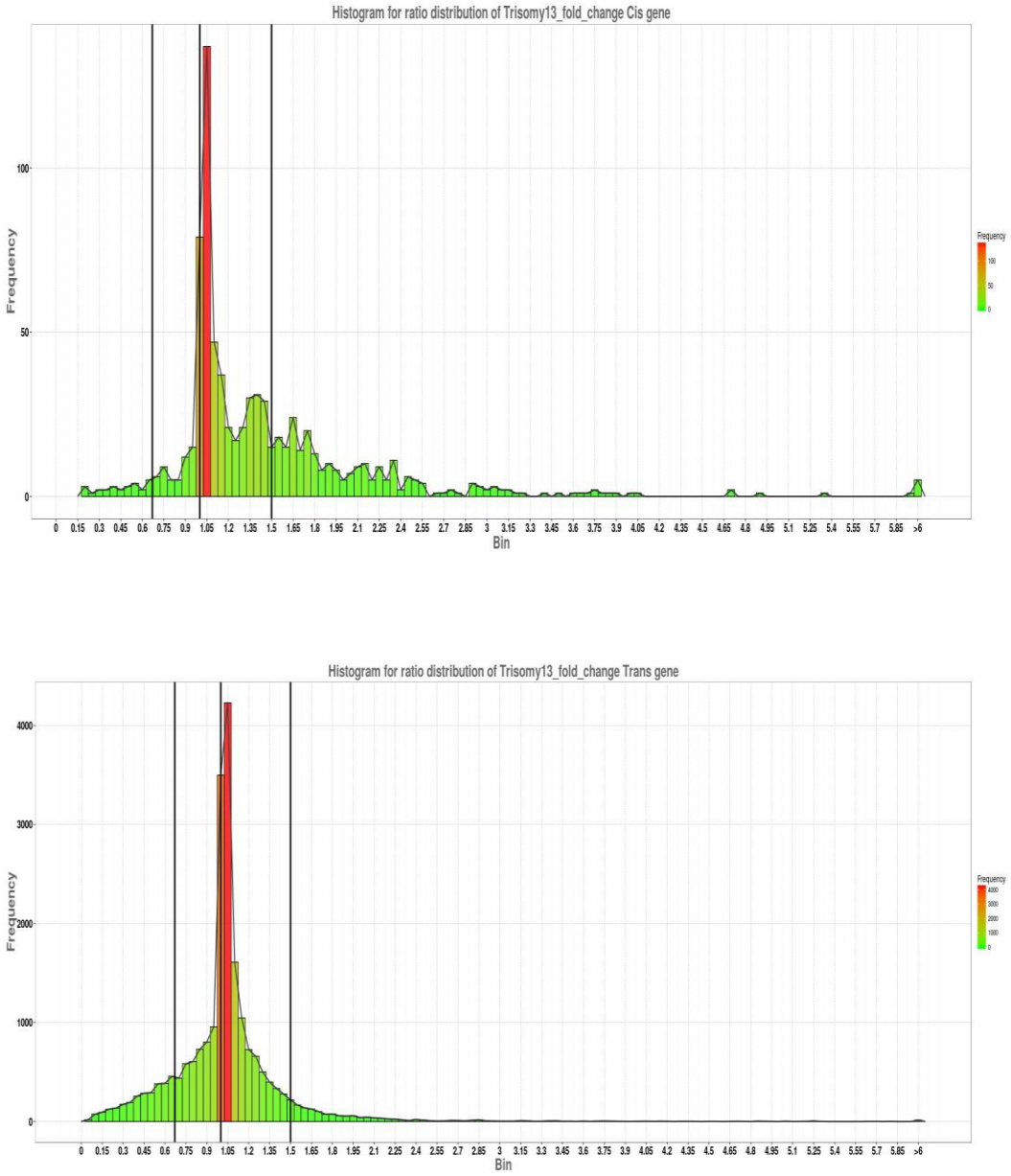


**Figure 8b – Ratio distribution plots for trisomy 13/diploid**

For explanation see Figure 8a.

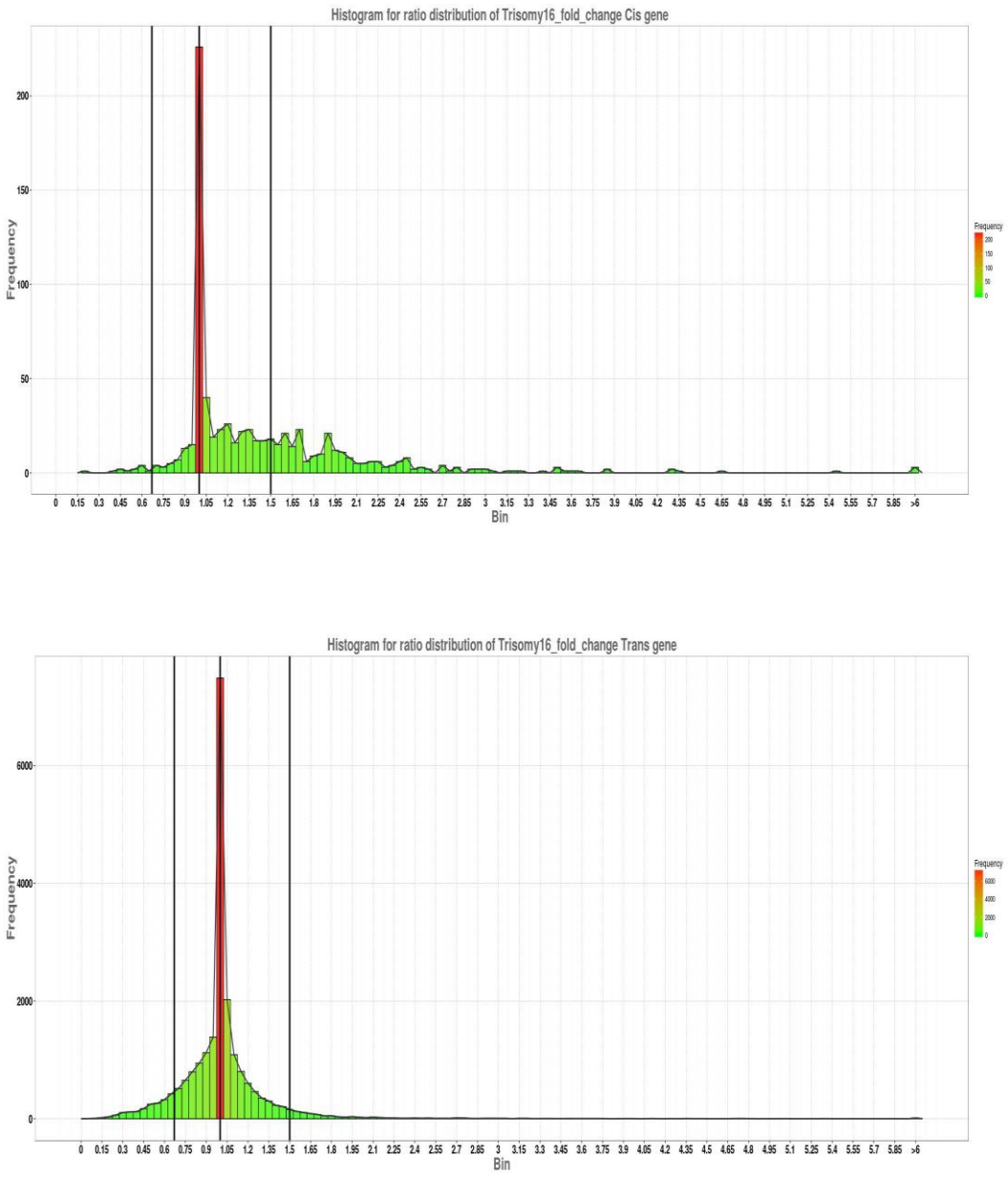


**Figure 8c – Ratio distribution plots for trisomy 16/diploid**

For explanation see Figure 8a.

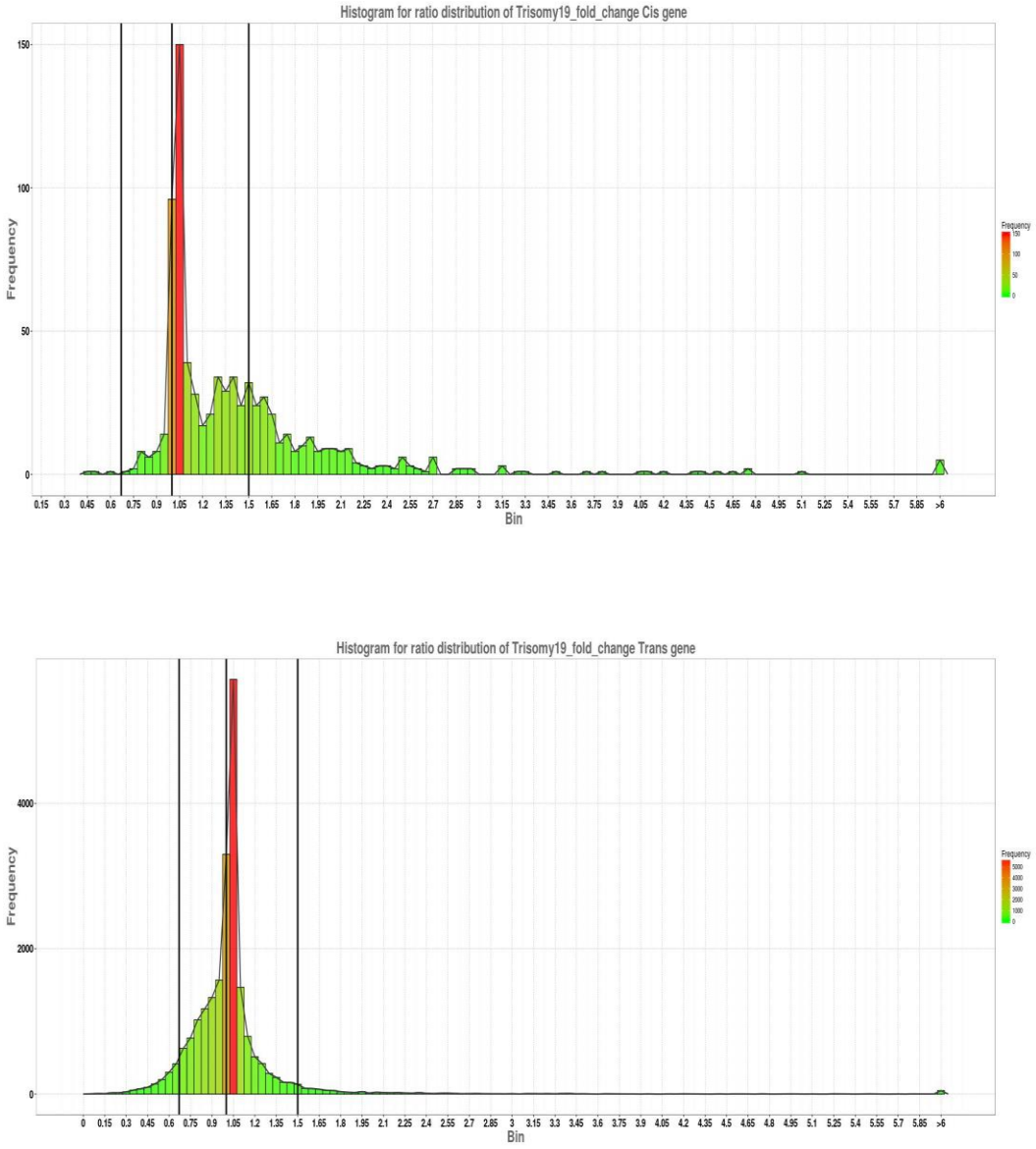


**Figure 8d – Ratio distribution plots for trisomy 19/diploid**

For explanation see Figure 8a.

Data from the set of yeast disomies was also used to produce ratio distribution plots, shown in Figure 9. The results provide an opportunity to show the effect of normalization method on perceived trends in expression. For each of the disomies in the yeast set, plots were generated for both cis and trans sets of genes. The signal strength from each gene in a ratio of disomic/haploid averaged over biological replicates was plotted in a distribution. This approach uses all genes for a test of trends and can quantify the magnitude of these shifts from normality because of the extensive statistical power of thousands of genes.

Expression data was collected from a large number of disomies in a haploid background, and ratio distribution plots were created for all of those involving a single chromosome. This analysis was undertaken as part of a project focused on issues of genomic balance, coordinated by Dr. James Birchler and Dr. Jianlin Cheng. Figures for Chapter 4 were produced by Md Soliman Islam, under the advisement of Dr. Cheng.

The general trends present in maize aneuploids may be observed in published data from organisms in different kingdoms. This suggests the effects of aneuploidy are the result dosage-sensitive mechanisms that are not specific to any particular group of organisms, but instead are universally shared.

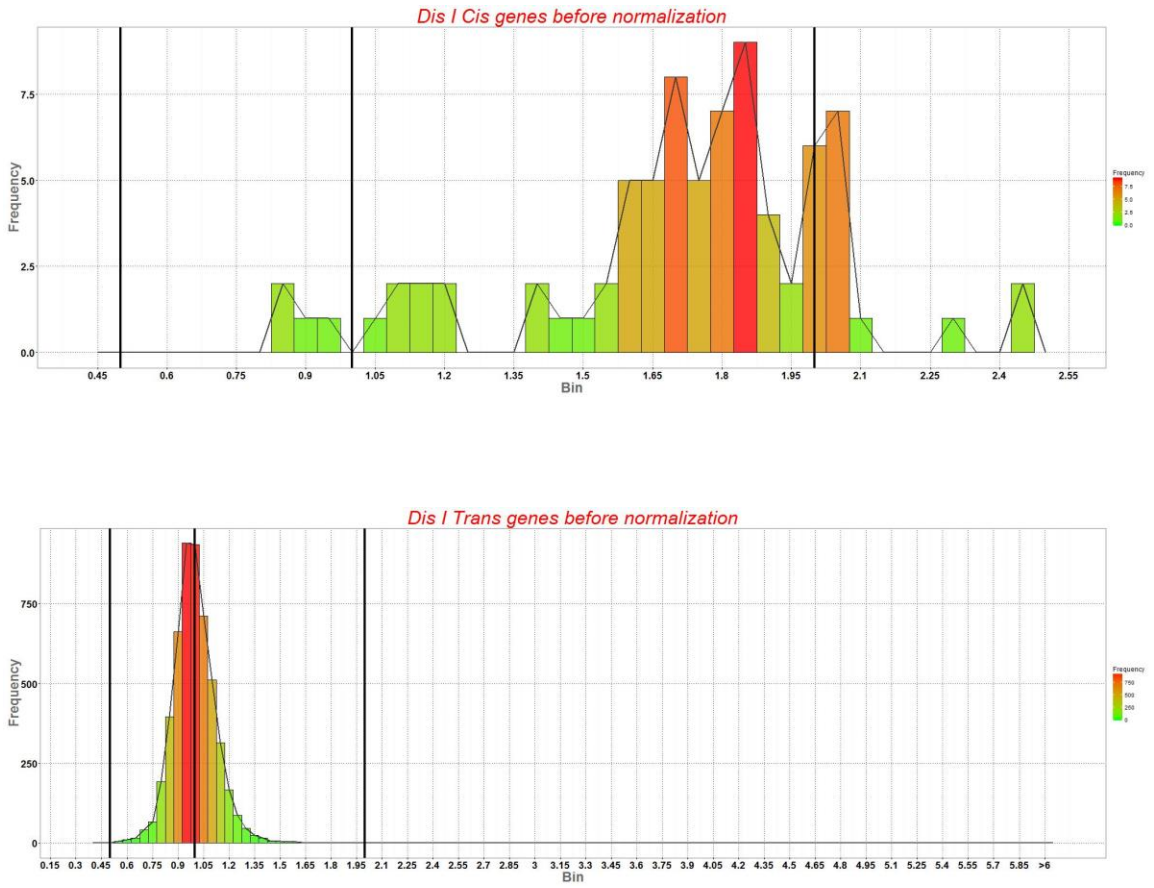


**Figure 9a – Ratio distribution plots for disomy I/haploid**

The top plot shows the distribution of expression ratios for cis genes, and the bottom plot for trans genes. In the series of disomic haploids, the presence and magnitude of inverse trans effects was variable in different disomies. Disomy I, for example, does not display a notable shift to the left of 1.00. Disomy IV (Figure 9c) shows the inverse trans effect prominently. All cis plots indicate the presence of direct dosage effects, though they are variable in magnitude across different disomies. Vertical guide bars are placed at 0.50 (inverse of cis dosage), 1.00 (no change from euploid), and 2.00 (direct value of cis dosage).

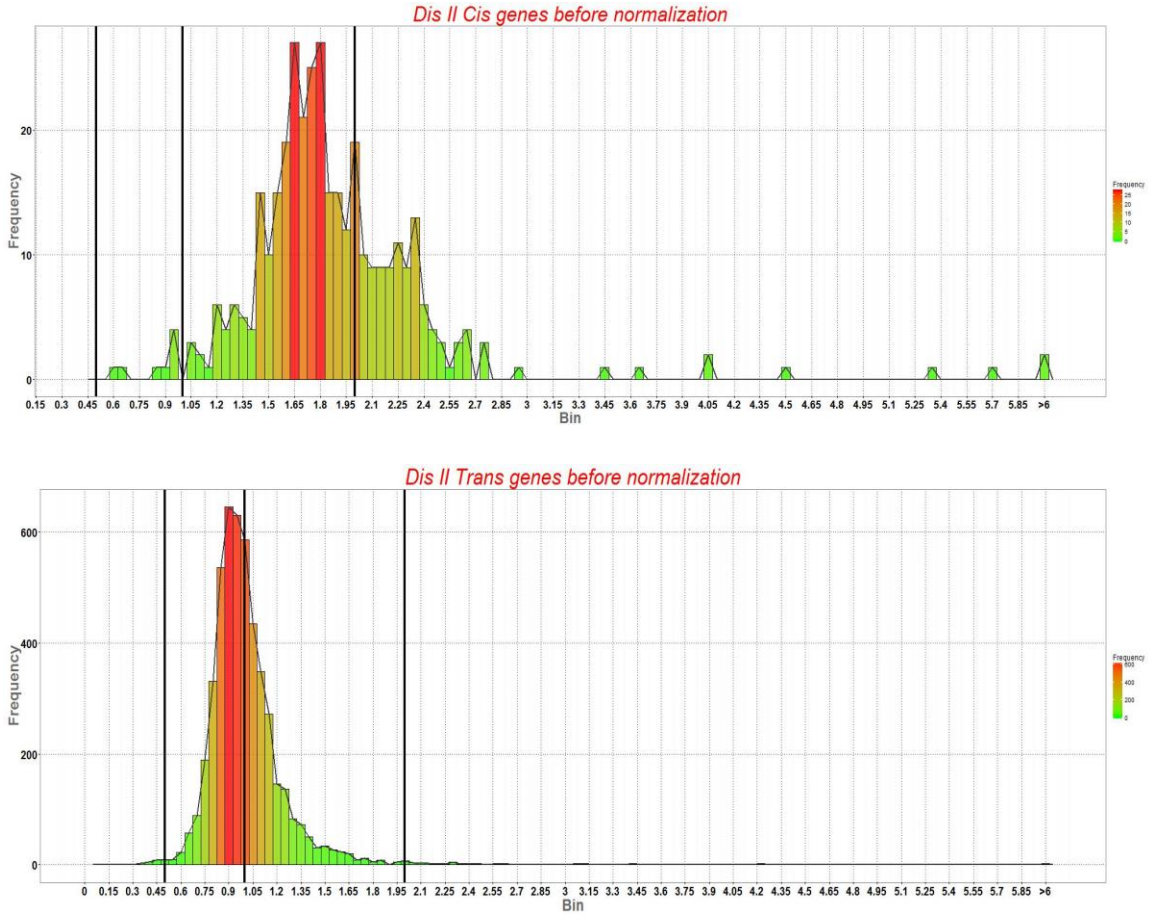


**Figure 9b – Ratio distribution plots for disomy II/haploid**

For explanation see Figure 9a.

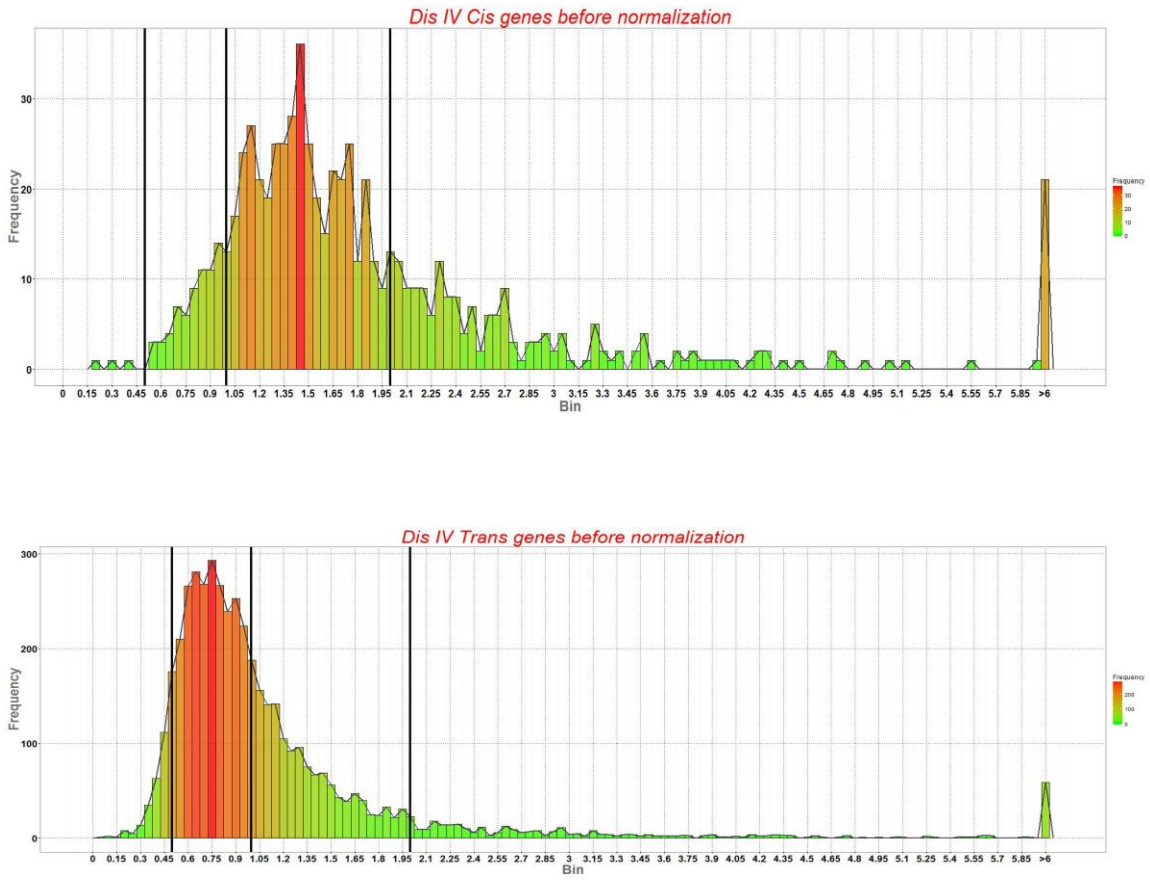


**Figure 9c – Ratio distribution plots for disomy IV/haploid**

For explanation see Figure 9a.

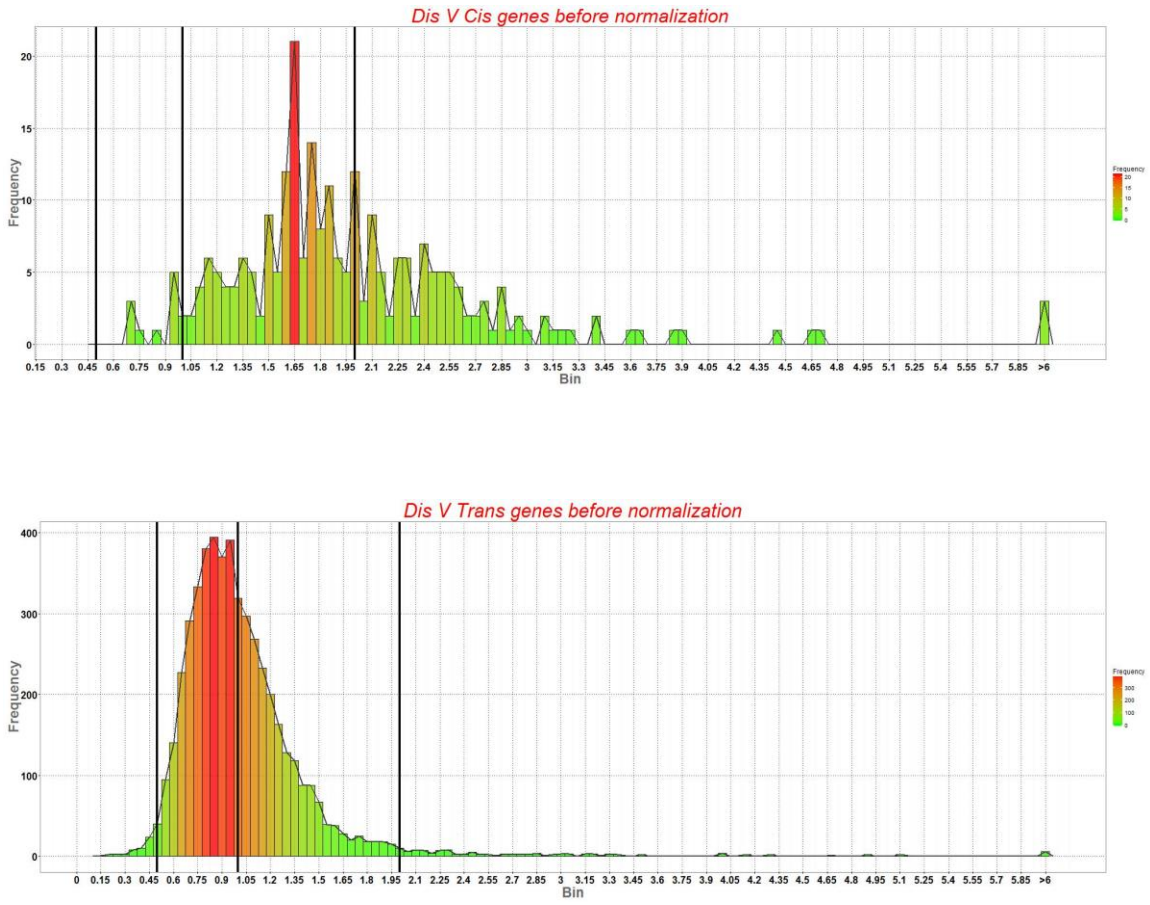


**Figure 9d – Ratio distribution plots for disomy V/haploid**

For explanation see Figure 9a.

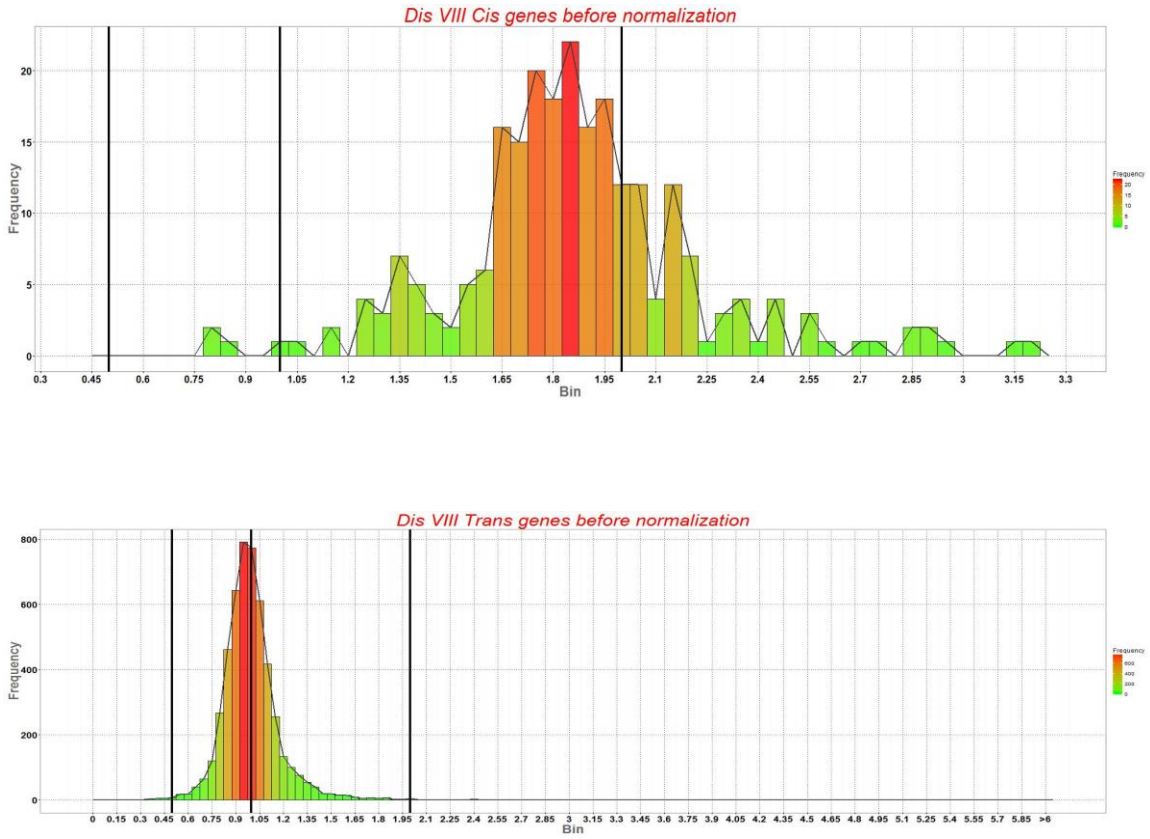


**Figure 9e – Ratio distribution plots for disomy VIII/haploid**

For explanation see Figure 9a.

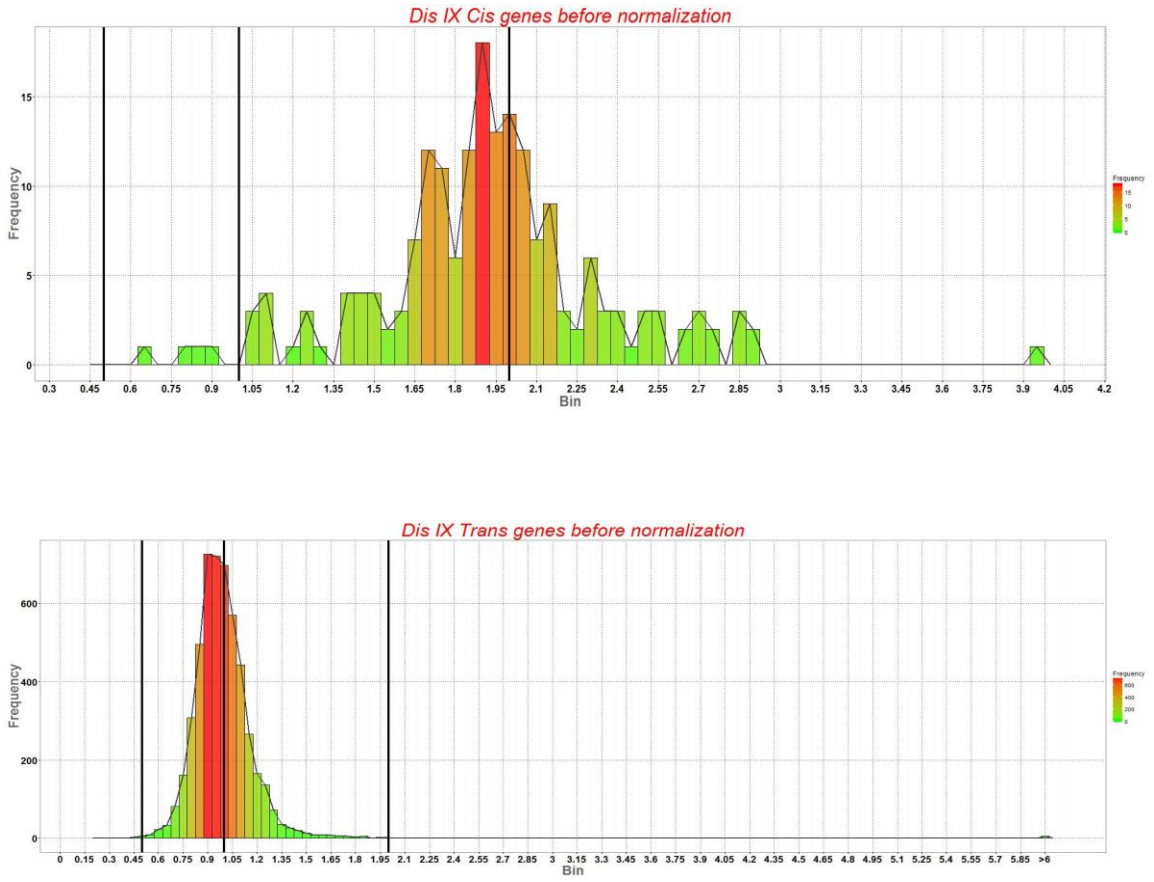


**Figure 9f – Ratio distribution plots for disomy IX/haploid**

For explanation see Figure 9a.

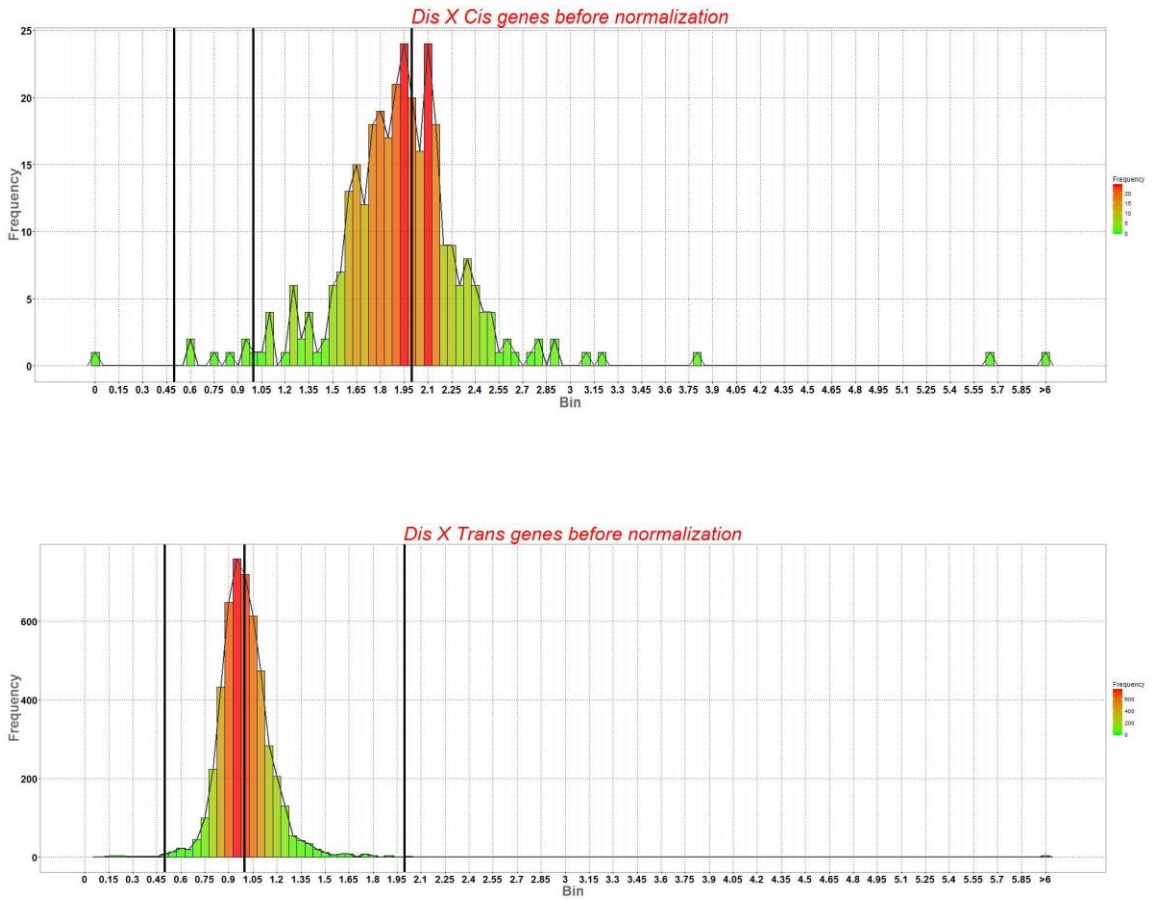


**Figure 9g – Ratio distribution plots for disomy X/haploid**

For explanation see Figure 9a.

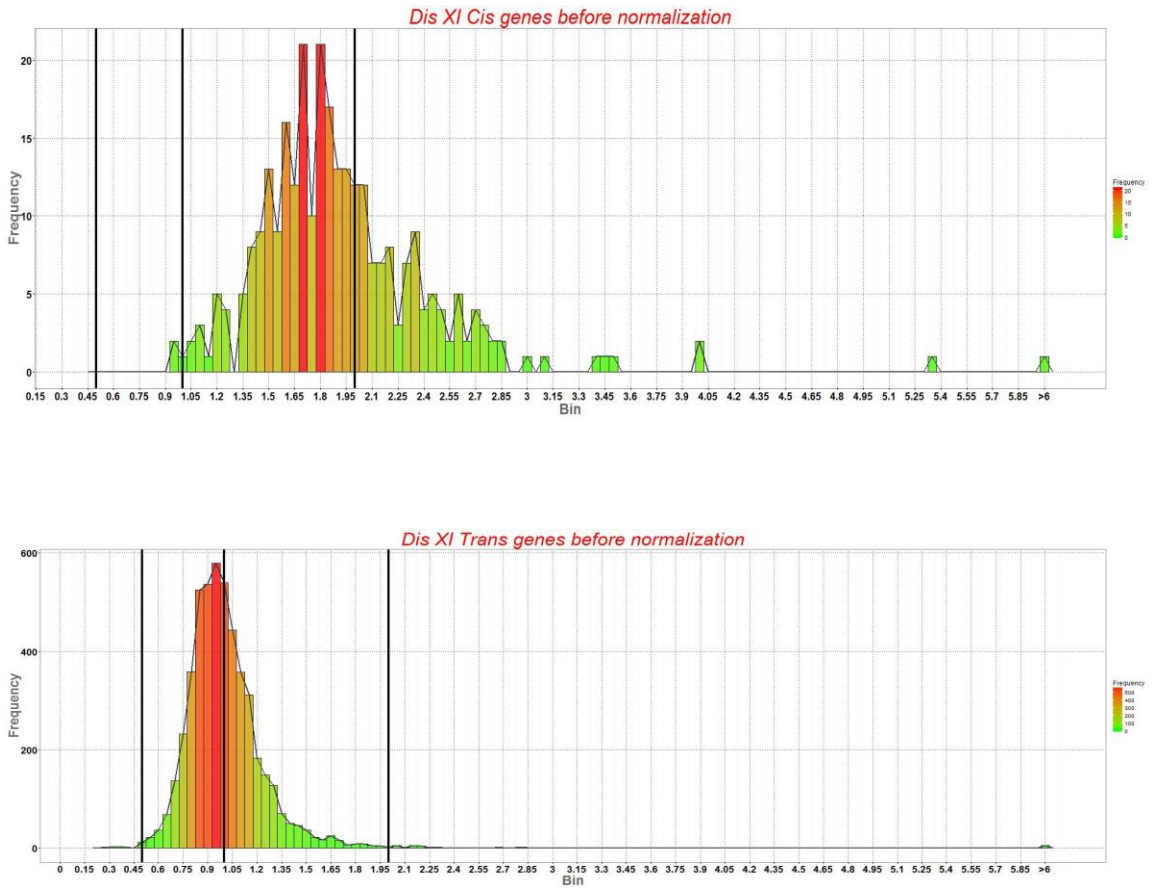


**Figure 9h – Ratio distribution plots for disomy XI/haploid**

For explanation see Figure 9a.

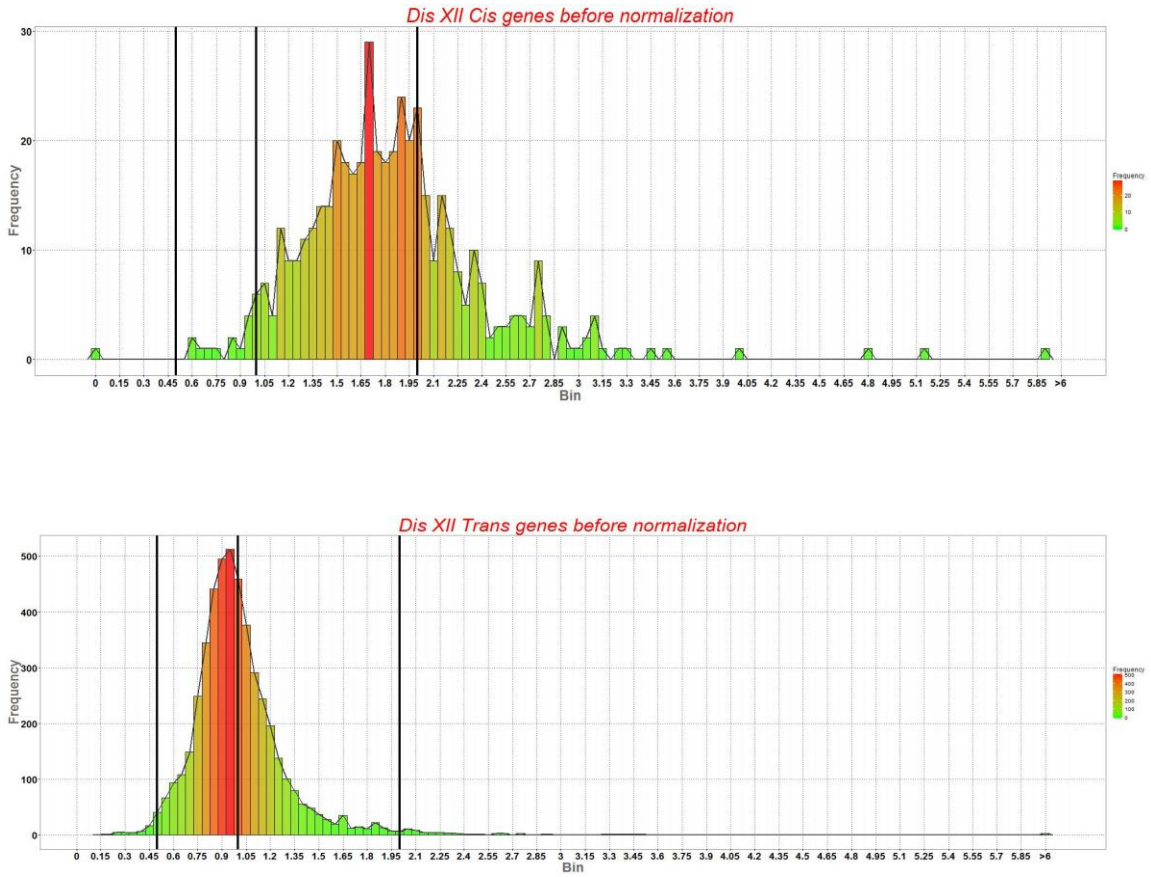


**Figure 9i – Ratio distribution plots for disomy XII/haploid**

For explanation see Figure 9a.

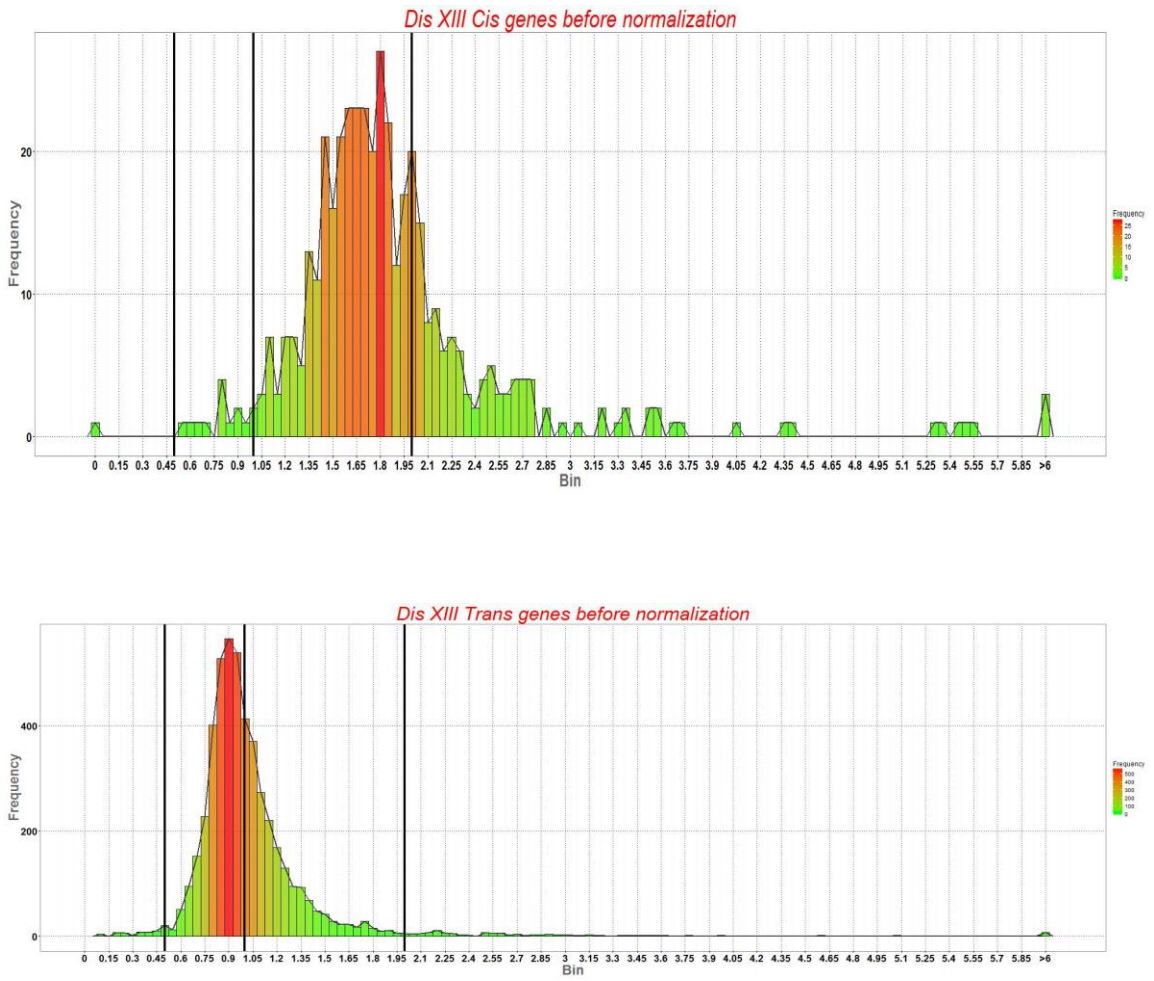


**Figure 9j – Ratio distribution plots for disomy XIII/haploid**

For explanation see Figure 9a.

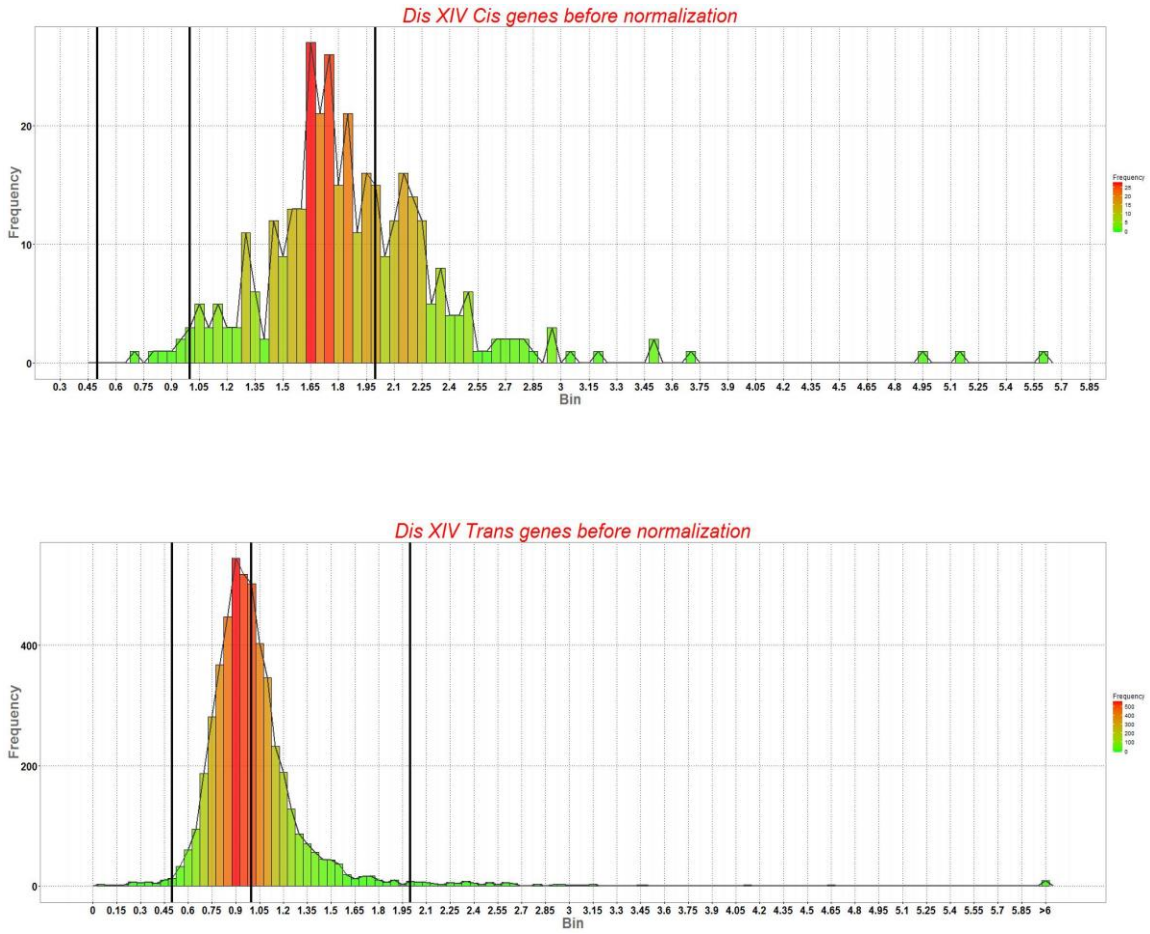


**Figure 9k – Ratio distribution plots for disomy XIV/haploid**

For explanation see Figure 9a.

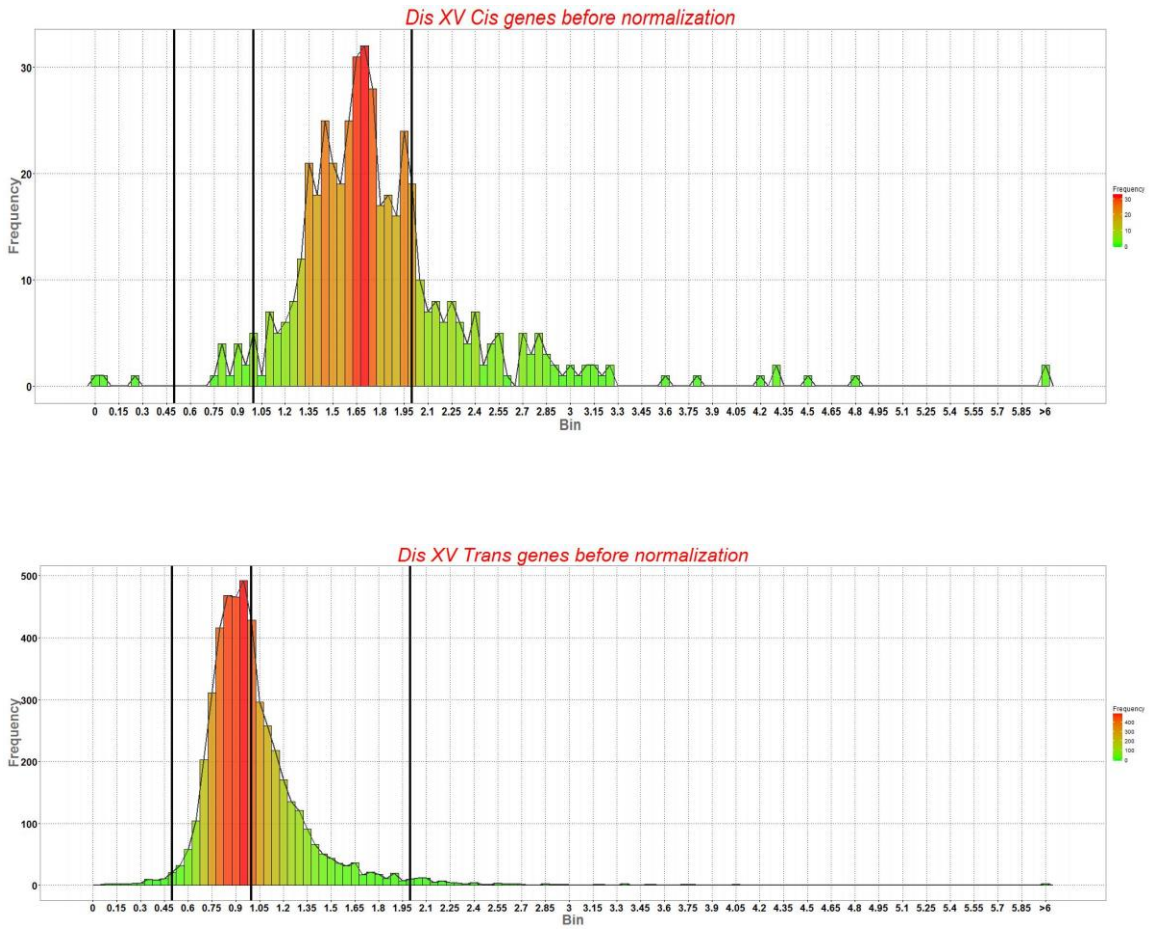


**Figure 9l – Ratio distribution plots for disomy XV/haploid**

For explanation see Figure 9a.

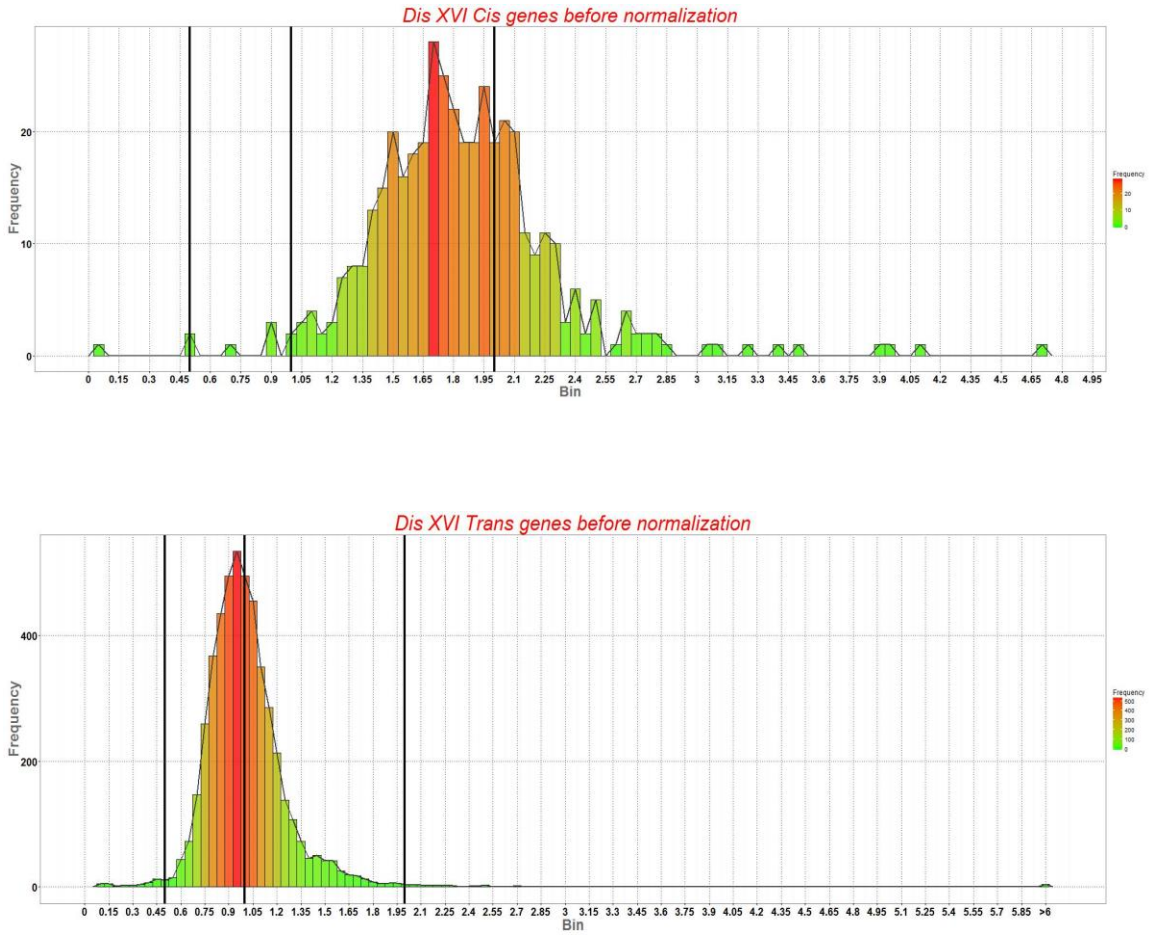


**Figure 9m – Ratio distribution plots for disomy XVI/haploid**

For explanation see Figure 9a.

# Conclusion

Aneuploidy results in widespread changes in gene expression, across the whole genome. This appears to happen in the same way in organisms that have exceedingly distant relationships to each other, implying a set of common, conserved mechanisms for regulation of gene expression, which are in some fashion dosage sensitive. A number of questions are provoked by this understanding. One possible avenue for future study is to determine which genes are most responsible for producing the inverse effect – that is, which genes, when varied in cis, produce a trans reaction in a large number of specific target genes. If highly efficacious regulatory genes like this can be identified, they could be used to answer a more general question: what is their mechanism of action? Insights into the mechanisms of dosage effects, such as the Gene Balance Hypothesis, may be applied both broadly (to unresolved questions about the interaction of molecules in cellular systems, for example) and deeply (to identify potentially important genes in the effects of aneuploidy in mammals).

# Bibliography


ALFENITO, M. R., and J. A. BIRCHLER, 1993 Molecular characterization of a maize B chromosome centric sequence. Genetics **135:** 589-597.

AUGER, D. L., and J. A. BIRCHLER, 2002 Maize tertiary trisomic stocks derived from B-A translocations. J Hered **93:** 42-47.

BENNETZEN, J. L., 2000 Transposable element contributions to plant gene and genome evolution. Plant Mol Biol **42:** 251-269.

BENTLEY, D. R., S. BALASUBRAMANIAN, H. P. SWERDLOW, G. P. SMITH, J. MILTON *et al.*, 2008 Accurate whole human genome sequencing using reversible terminator chemistry. Nature **456:** 53-59.

BHADRA, U., M. PAL-BHADRA and J. A. BIRCHLER, 1999 Role of the male specific lethal (msl) genes in modifying the effects of sex chromosomal dosage in Drosophila. Genetics **152:** 249-268.

BHADRA, U., M. PAL-BHADRA and J. A. BIRCHLER, 2000 Histone acetylation and gene expression analysis of sex lethal mutants in Drosophila. Genetics **155:** 753-763.

BIRCHLER, J., L. SUN, H. FERNANDEZ, R. DONOHUE, W. XIE *et al.*, 2011 Re-evaluation of the function of the male specific lethal complex in Drosophila. J Genet Genomics **38:** 327-332.

BIRCHLER, J. A., 1979 A study of enzyme activities in a dosage series of the long arm of chromosome one in maize. Genetics **92:** 1211-1229.

BIRCHLER, J. A., 1981 The genetic basis of dosage compensation of alcohol dehydrogenase-1 in maize. Genetics **97:** 625-637.

BIRCHLER, J. A., 2013 Aneuploidy in plants and flies: the origin of studies of genomic imbalance. Semin Cell Dev Biol **24:** 315-319.

BIRCHLER, J. A., and M. R. ALFENITO, 1993 Marker systems for B-A translocations in maize. Journal of Heredity **84:** 135-138.

BIRCHLER, J. A., M. P. BHADRA and U. BHADRA, 2000 Making noise about silence: repression of repeated genes in animals. Curr Opin Genet Dev **10:** 211-216.

BIRCHLER, J. A., D. J. CHALFOUN and D. M. LEVIN, 1990a Recombination in the B chromosome of maize to produce A-B-A chromosomes. Genetics **126:** 723-733.

BIRCHLER, J. A., J. C. HIEBERT and K. PAIGEN, 1990b Analysis of autosomal dosage compensation involving the alcohol dehydrogenase locus in *Drosophila melanogaster*. Genetics **124:** 679-686.

BIRCHLER, J. A., and K. J. NEWTON, 1981 Modulation of protein levels in chromosomal dosage series of maize: the biochemical basis of aneuploid syndromes. Genetics **99:** 247-266.

BIRCHLER, J. A., and R. A. VEITIA, 2007 The gene balance hypothesis: from classical genetics to modern genomics. Plant Cell **19:** 395-402.

BIRCHLER, J. A., and R. A. VEITIA, 2010 The gene balance hypothesis: implications for gene regulation, quantitative traits and evolution. New Phytol **186:** 54-62.

BIRCHLER, J. A., and R. A. VEITIA, 2012 Gene balance hypothesis: connecting issues of dosage sensitivity across biological disciplines. Proc. Natl. Acad. Sci. U. S. A. **109:** 14746-14753.

BLAKESLEE, A. F., 1934 New Jimson weeds from old chromosomes. Journal of Heredity **25:** 81-108.

BLAKESLEE, A. F., J. BELLING and M. E. FARNHAM, 1920 Chromosomal duplication and Mendelian phenomena in Datura mutants. Science **52:** 388-390.
BRIDGES, C. B., 1916 Non-Disjunction as Proof of the Chromosome Theory of Heredity. Genetics **1:** 1-52.
CALLEWAERT, B., F. MALFAIT, B. LOEYS and A. DE PAEPE, 2008 Ehlers-Danlos syndromes and Marfan syndrome. Best Pract Res Clin Rheumatol **22:** 165-189.
CAREY, J. C., 2012 Perspectives on the care and management of infants with trisomy 18 and trisomy 13: striving for balance. Curr Opin Pediatr **24:** 672-678.
CEREDA, A., and J. C. CAREY, 2012 The trisomy 18 syndrome. Orphanet J Rare Dis **7:** 81.
CHAN, J. Y., 2011 A clinical overview of centrosome amplification in human cancers. Int J Biol Sci **7:** 1122-1144.
CONANT, G. C., and K. H. WOLFE, 2007 Increased glycolytic flux as an outcome of whole-genome duplication in yeast. Mol Syst Biol **3:** 129.
DISTECHE, C. M., 2012 Dosage compensation of the sex chromosomes. Annu Rev Genet **46:** 537-560.
EDGER, P. P., and J. C. PIRES, 2009 Gene and genome duplications: the impact of dosage-sensitivity on the fate of nuclear genes. Chromosome Res **17:** 699-717.
FESCHOTTE, C., N. JIANG and S. R. WESSLER, 2002 Plant transposable elements: where genetics meets genomics. Nat Rev Genet **3:** 329-341.
GORDON, D. J., B. RESIO and D. PELLMAN, 2012 Causes and consequences of aneuploidy in cancer. Nat Rev Genet **13:** 189-203.
GUO, M., and J. A. BIRCHLER, 1994 Trans-acting dosage effects on the expression of model gene systems in maize aneuploids. Science **266:** 1999-2002.
HALL, H., P. HUNT and T. HASSOLD, 2006 Meiosis and sex chromosome aneuploidy: how meiotic errors cause aneuploidy; how aneuploidy causes meiotic errors. Curr Opin Genet Dev **16:** 323-329.
HASSOLD, T. J., and P. A. JACOBS, 1984 Trisomy in man. Annu Rev Genet **18:** 69-97.
HJERRILD, B. E., K. H. MORTENSEN and C. H. GRAVHOLT, 2008 Turner syndrome and clinical treatment. Br Med Bull **86:** 77-93.
JONES, R. N., W. VIEGAS and A. HOUBEN, 2008 A century of B chromosomes in plants: so what? Ann Bot **101:** 767-775.
KING, R. W., 2008 When 2+2=5: the origins and fates of aneuploid and tetraploid cells. Biochim Biophys Acta **1786:** 4-14.
LANA-ELOLA, E., S. D. WATSON-SCALES, E. M. FISHER and V. L. TYBULEWICZ, 2011 Down syndrome: searching for the genetic culprits. Dis Model Mech **4:** 586-595.
LUCCHESI, J. C., W. G. KELLY and B. PANNING, 2005 Chromatin remodeling in dosage compensation. Annu Rev Genet **39:** 615-651.
MEYERS, B. C., S. V. TINGEY and M. MORGANTE, 2001 Abundance, distribution, and transcriptional activity of repetitive elements in the maize genome. Genome Res **11:** 1660-1676.
NICOLAIDIS, P., and M. B. PETERSEN, 1998 Origin and mechanisms of non-disjunction in human autosomal trisomies. Hum Reprod **13:** 313-319.
PAVELKA, N., G. RANCATI and R. LI, 2010 Dr Jekyll and Mr Hyde: role of aneuploidy in cellular adaptation and cancer. Curr Opin Cell Biol **22:** 809-815.

PFAU, S. J., and A. AMON, 2012 Chromosomal instability and aneuploidy in cancer: from yeast to man. EMBO Rep **13:** 515-527.
RANDOLPH, L. F., 1941 Genetic Characteristics of the B Chromosomes in Maize. Genetics **26:** 608-631.
RICKE, R. M., J. H. VAN REE and J. M. VAN DEURSEN, 2008 Whole chromosome instability and cancer: a complex relationship. Trends Genet **24:** 457-466.
RIO, D. C., M. ARES, JR., G. J. HANNON and T. W. NILSEN, 2010 Purification of RNA using TRIzol (TRI reagent). Cold Spring Harb Protoc **2010:** pdb prot5439.
ROEBER, F. K., G. A. GORDILLO and H. H. GEIGER, 2005 In vivo haploid induction in maize-performance of new inducers and significance of doubled haploid lines in hybrid breeding. Maydica **50:** 275-283.
ROMAN, H., 1947 Mitotic Nondisjunction in the Case of Interchanges Involving the B-Type Chromosome in Maize. Genetics **32:** 391-409.
SEMON, M., and K. H. WOLFE, 2007 Consequences of genome duplication. Curr Opin Genet Dev **17:** 505-512.
SUN, L., A. F. JOHNSON, R. C. DONOHUE, J. LI, J. CHENG *et al.*, 2013 Dosage compensation and inverse effects in triple X metafemales of Drosophila. Proc Natl Acad Sci U S A **110:** 7383-7388.
SUN, X., and J. A. BIRCHLER, 2009 Interaction study of the male specific lethal (MSL) complex and trans-acting dosage effects in metafemales of Drosophila melanogaster. Cytogenet Genome Res **124:** 298-311.
TORRES, E. M., T. SOKOLSKY, C. M. TUCKER, L. Y. CHAN, M. BOSELLI *et al.*, 2007 Effects of aneuploidy on cellular physiology and cell division in haploid yeast. Science **317:** 916-924.
VEITIA, R. A., and J. A. BIRCHLER, 2010 Dominance and gene dosage balance in health and disease: why levels matter! J Pathol **220:** 174-185.
VEITIA, R. A., S. BOTTANI and J. A. BIRCHLER, 2013 Gene dosage effects: nonlinearities, genetic interactions, and dosage compensation. Trends Genet. **29:** 385-393.
VISOOTSAK, J., and J. M. GRAHAM, JR., 2006 Klinefelter syndrome and other sex chromosomal aneuploidies. Orphanet J Rare Dis **1:** 42.
WILLIAMS, B. R., V. R. PRABHU, K. E. HUNTER, C. M. GLAZIER, C. A. WHITTAKER *et al.*, 2008 Aneuploidy affects proliferation and spontaneous immortalization in mammalian cells. Science **322:** 703-709.
YAMAKAWA, K., 2012 Towards the understanding of Down syndrome using mouse models. Congenit Anom (Kyoto) **52:** 67-71.
YI, X., Z. DU and Z. SU, 2013 PlantGSEA: a gene set enrichment analysis toolkit for plant community. Nucleic Acids Res **41:** W98-103.
ZHU, X., M. GERSTEIN and M. SNYDER, 2007 Getting connected: analysis and principles of biological networks. Genes Dev **21:** 1010-1024.

# Supplemental Figures

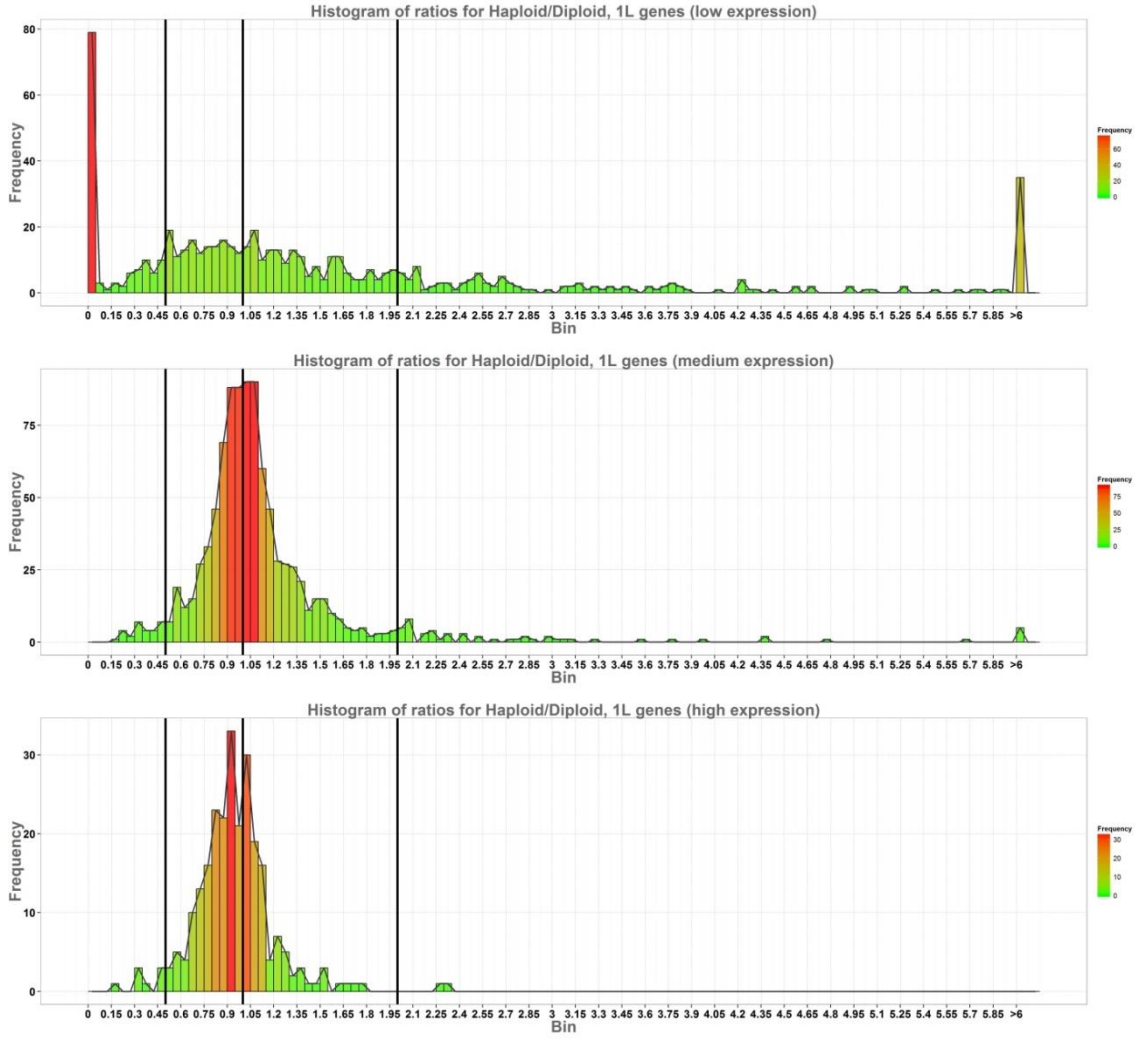


**Figure S3a – Ratio distribution plots for cis genes, haploid/diploid**

The three plots in each supplemental figure were produced using the same data as displayed in Figures 3 and 4. To produce the three plots, ratios were divided into three sets, based on the expression quantity per transcriptome (low: <1 FPKM; medium: 1-30 FPKM; high: >30 FPKM). Extreme outlier peaks, with expression ratios below 0.05 or above 6.00, are found to be confined to the low-expression plots.

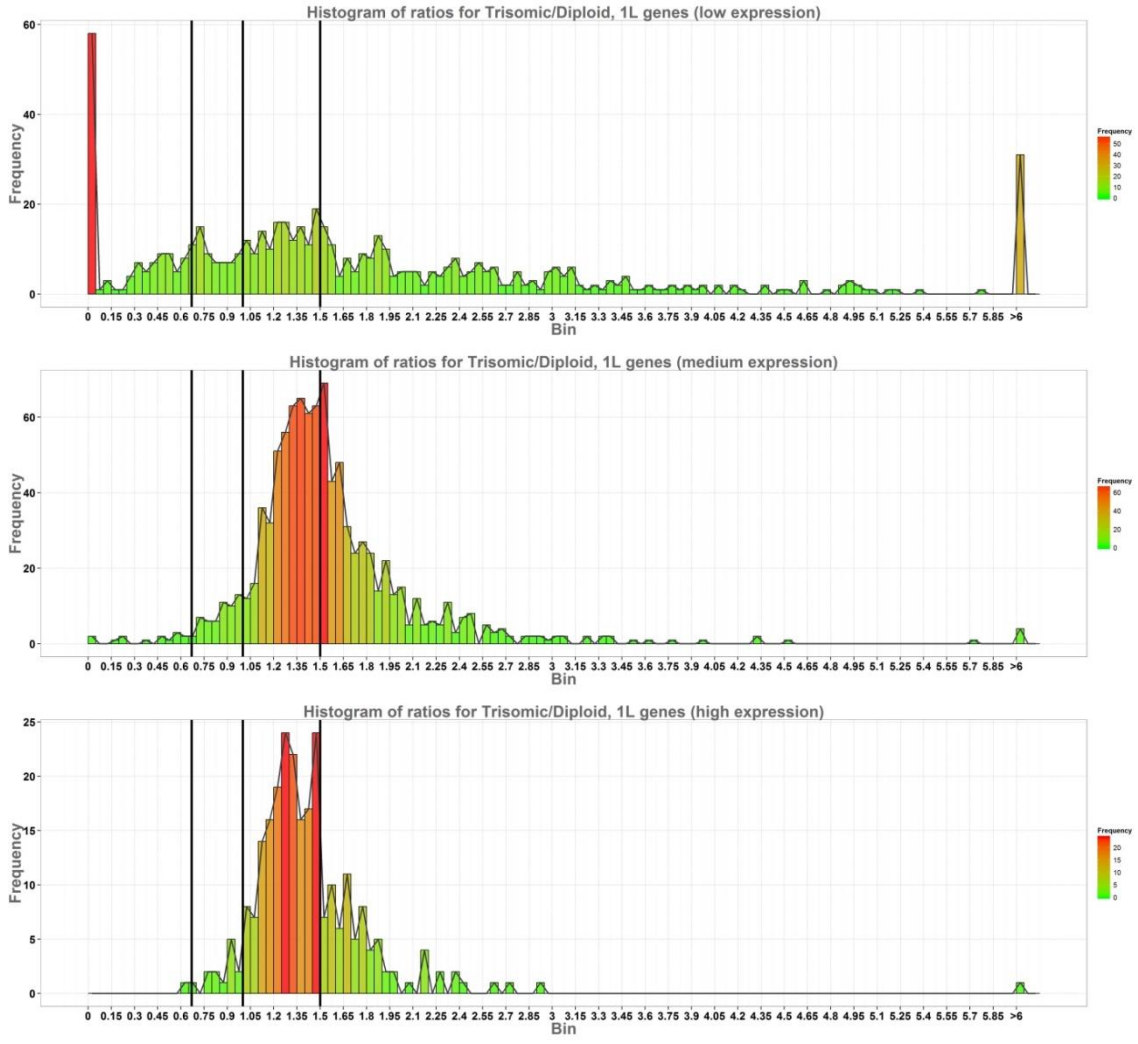


**Figure S3b – Ratio distribution plots for cis genes, trisomic/diploid**

For explanation see Figure S3a.

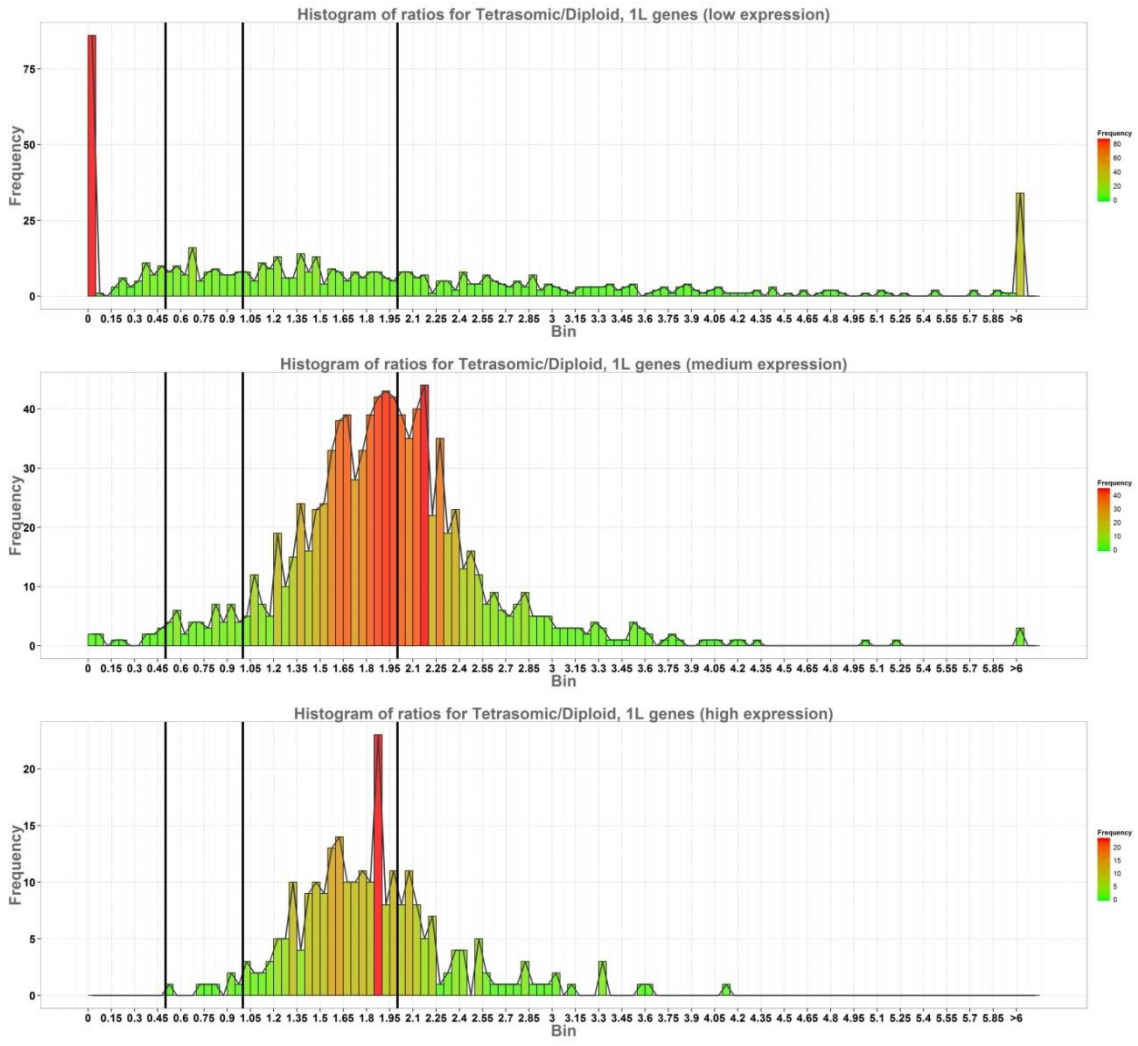


**Figure S3c – Ratio distribution plots for cis genes, tetrasomic/diploid**

For explanation see Figure S3a.

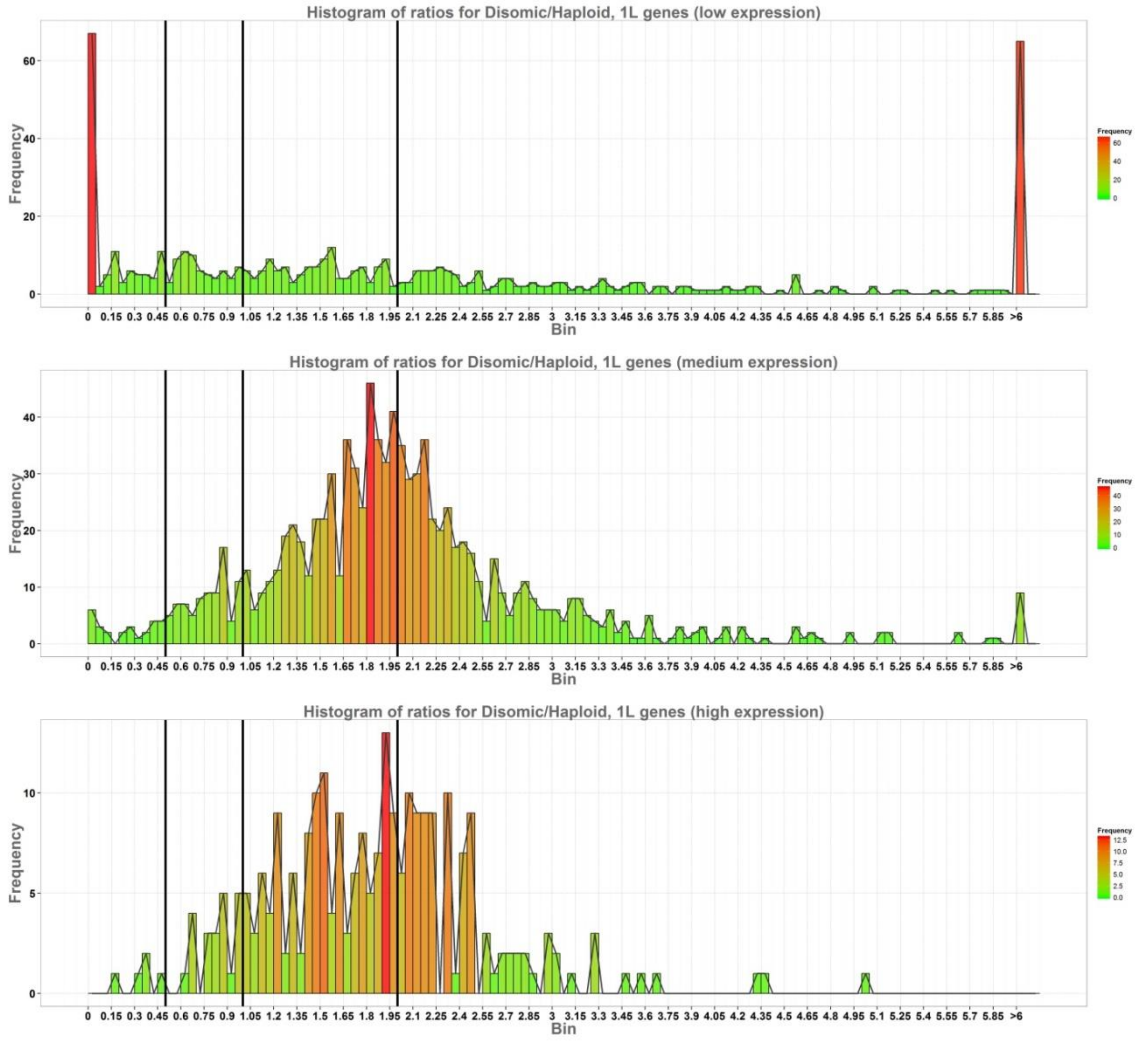


**Figure S3d – Ratio distribution plots for cis genes, disomic/haploid**

For explanation see Figure S3a.

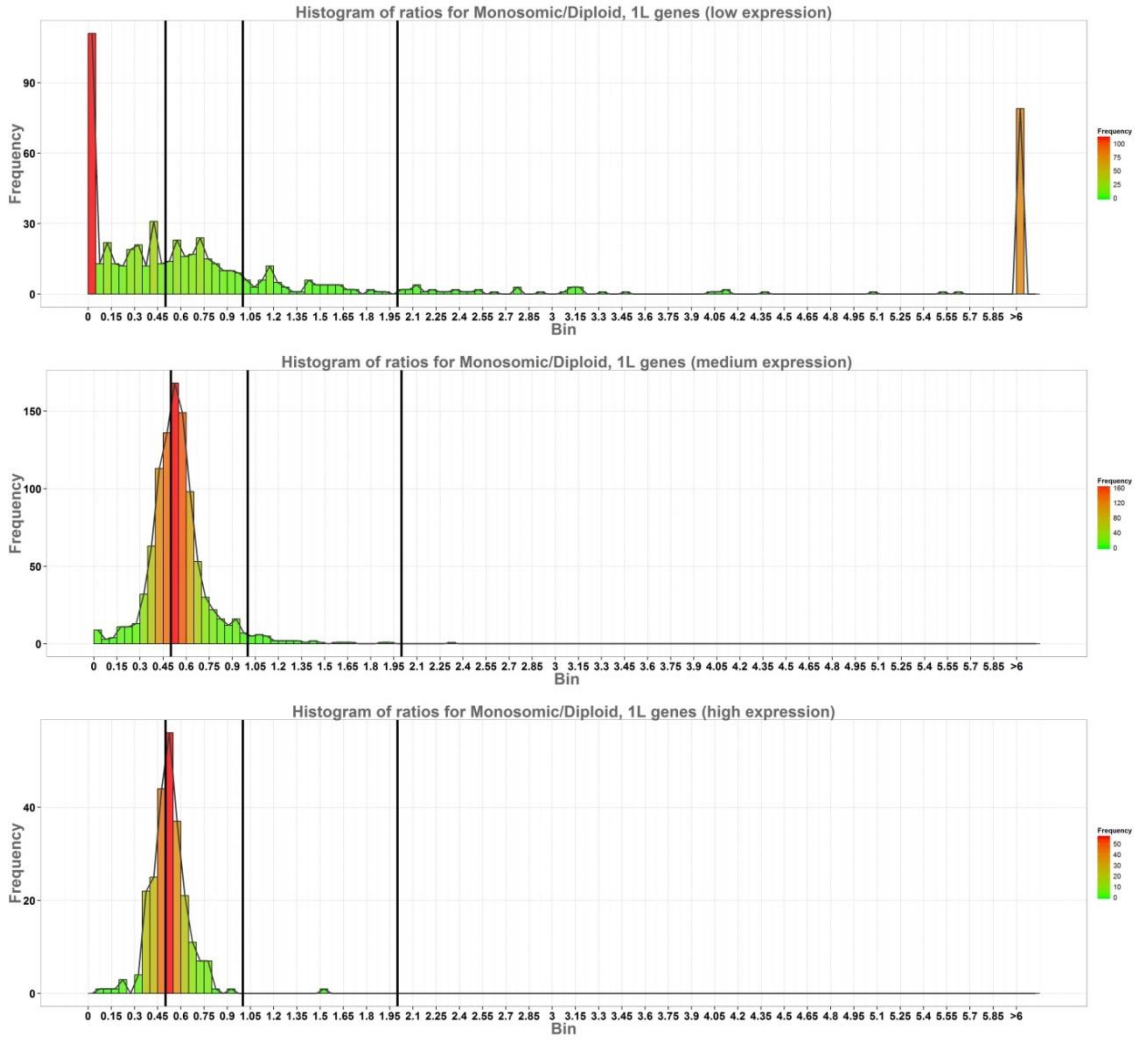


**Figure S3e – Ratio distribution plots for cis genes, monosomic/diploid, second dosage series**

For explanation see Figure S3a.

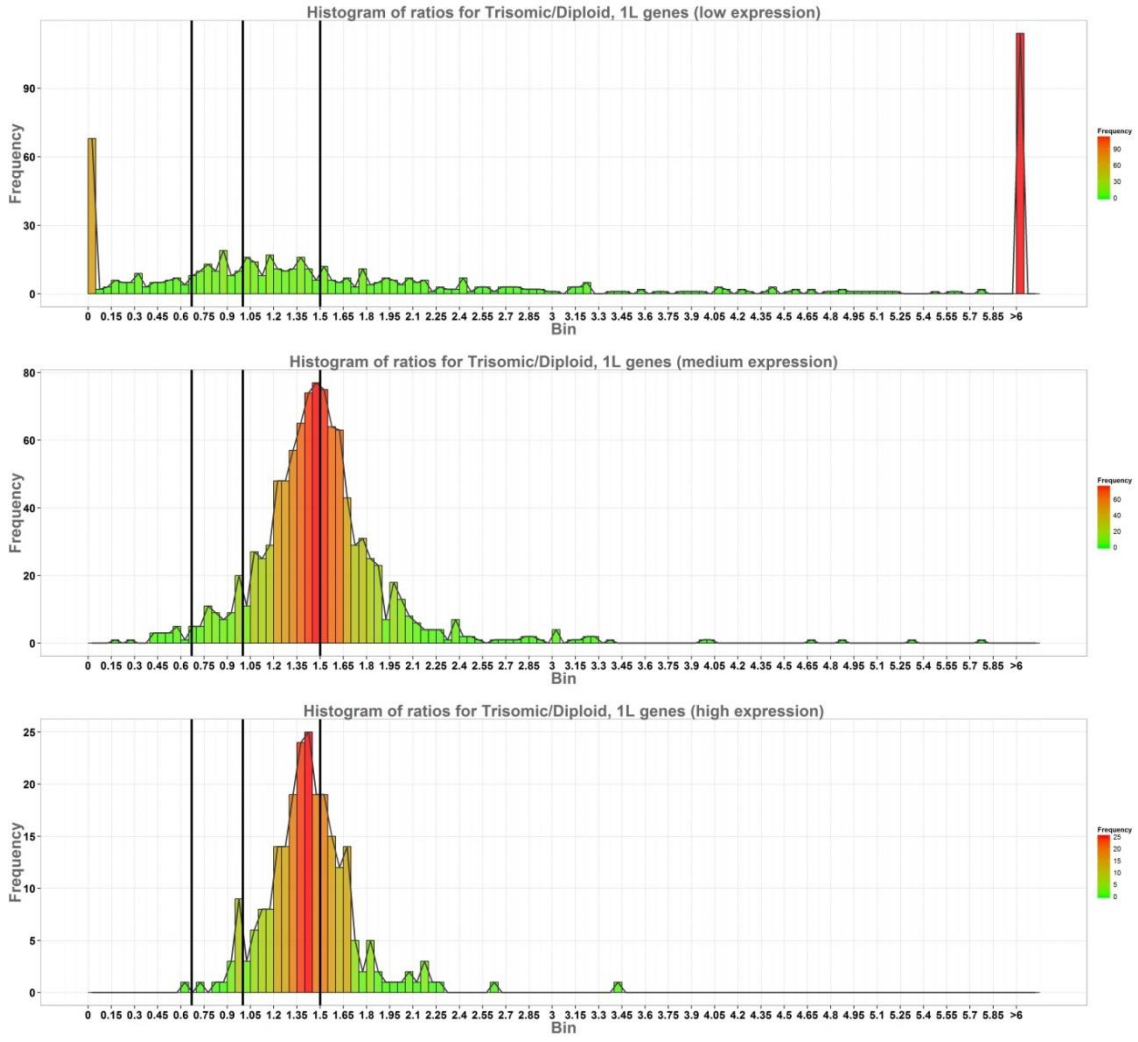


**Figure S3f – Ratio distribution plots for cis genes, trisomic/diploid, second dosage series**

For explanation see Figure S3a.

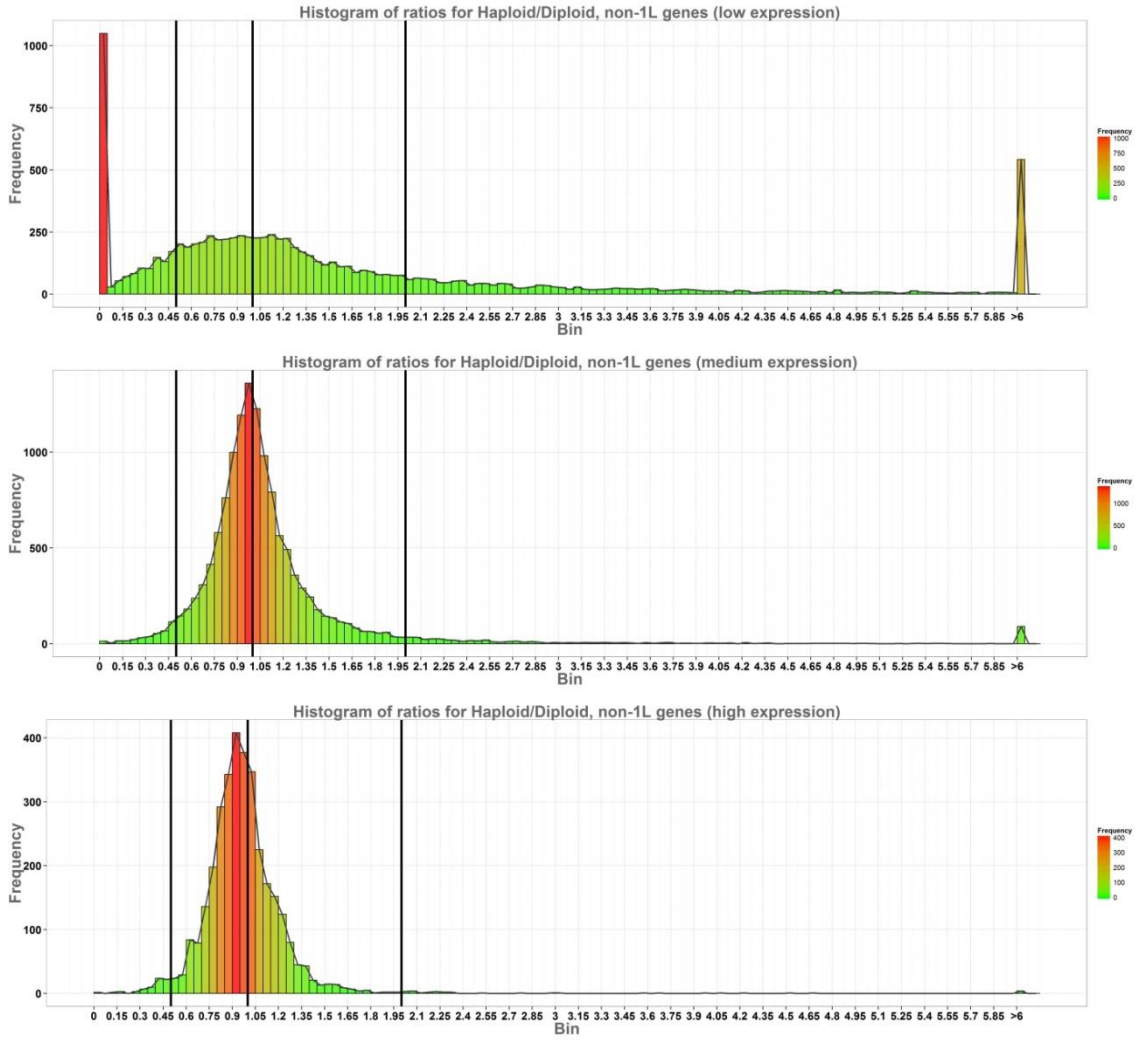


**Figure S4a – Ratio distribution plots for trans genes, haploid/diploid**

For explanation see Figure S3a.

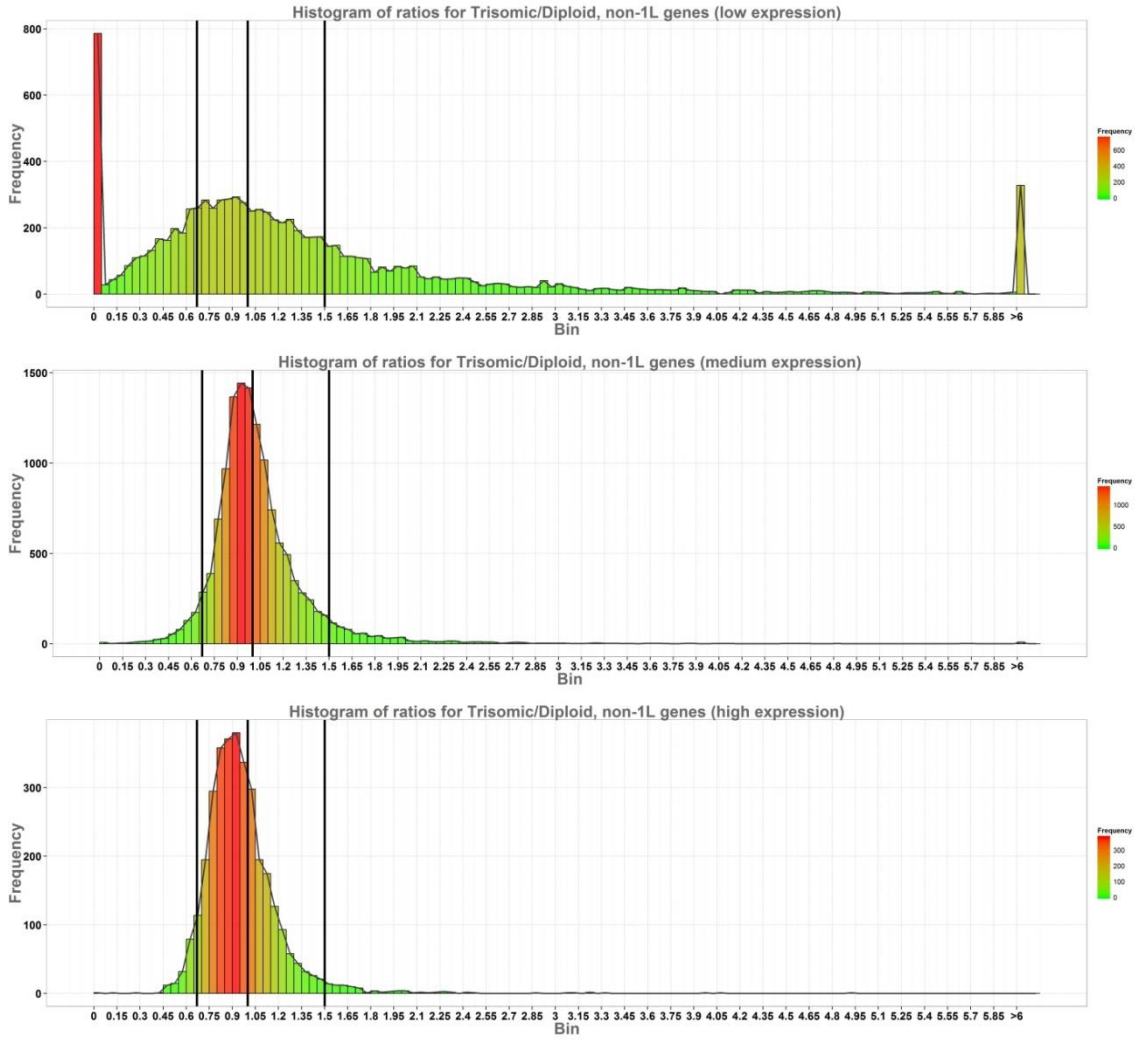


**Figure S4b – Ratio distribution plots for trans genes, trisomic/diploid**

For explanation see Figure S3a.

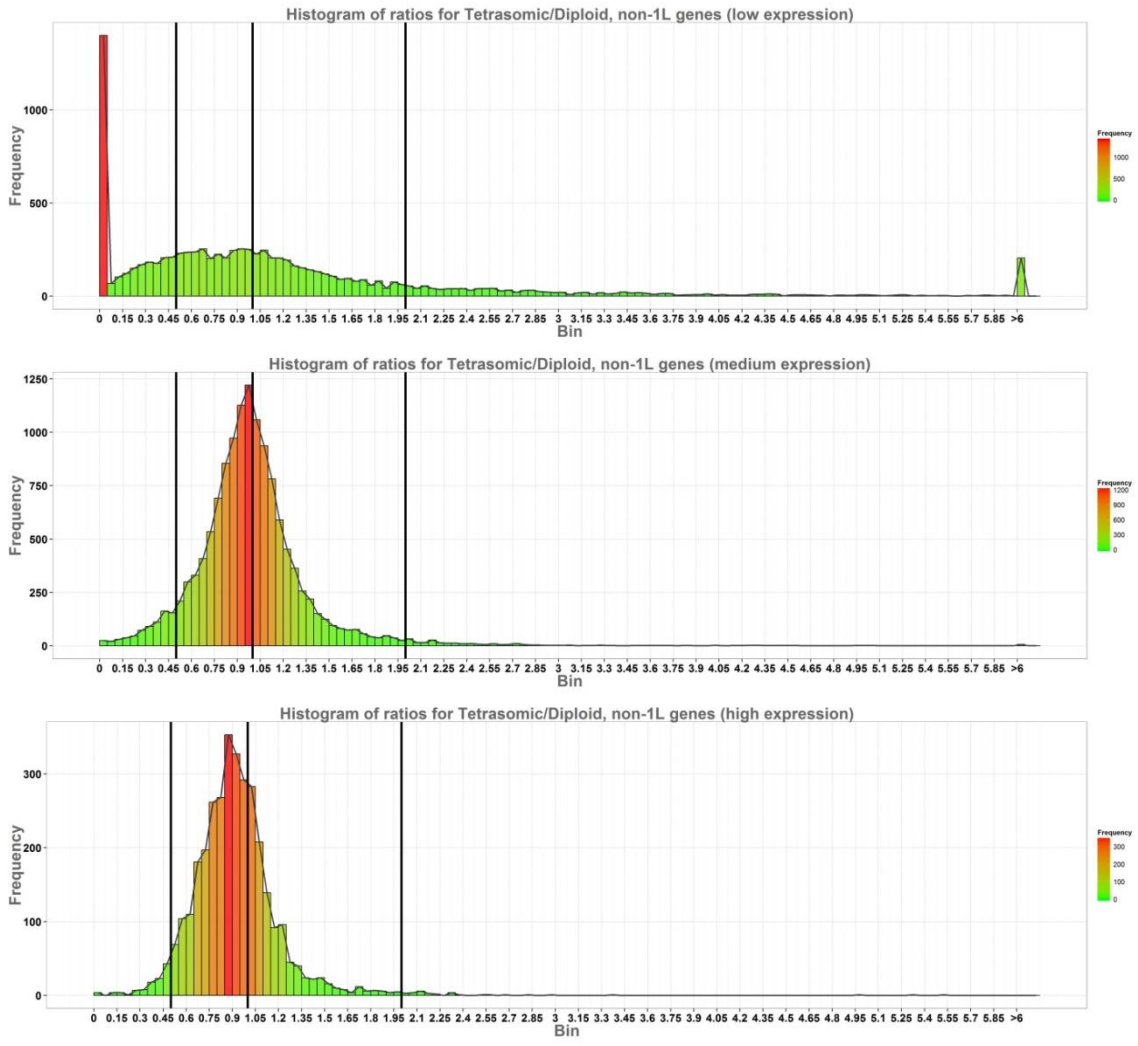


**Figure S4c – Ratio distribution plots for trans genes, tetrasomic/diploid**

For explanation see Figure S3a.

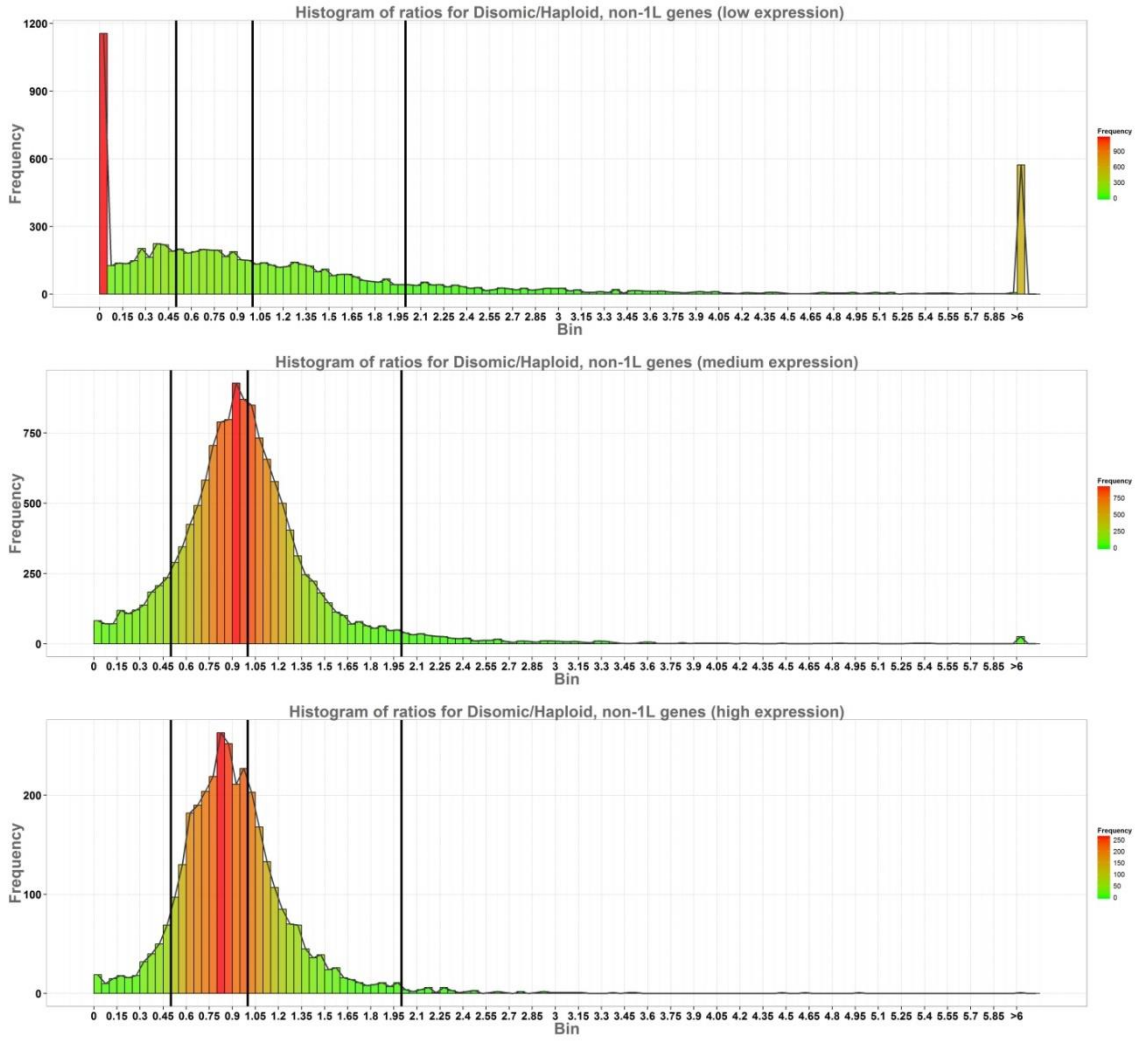


**Figure S4d – Ratio distribution plots for trans genes, disomic/haploid**

For explanation see Figure S3a.

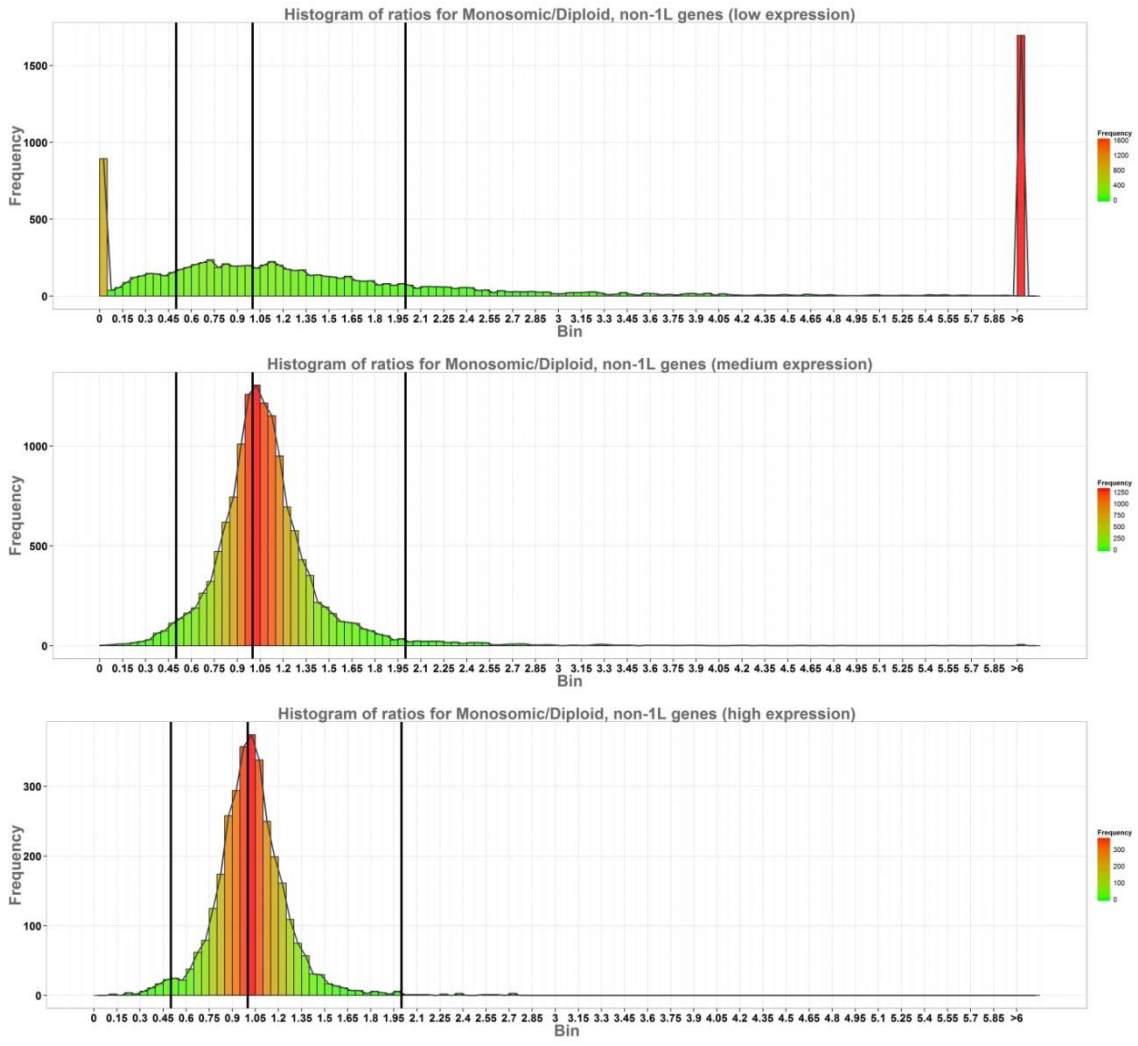


**Figure S4e – Ratio distribution plots for trans genes, monosomic/diploid, second dosage series**

For explanation see Figure S3a.

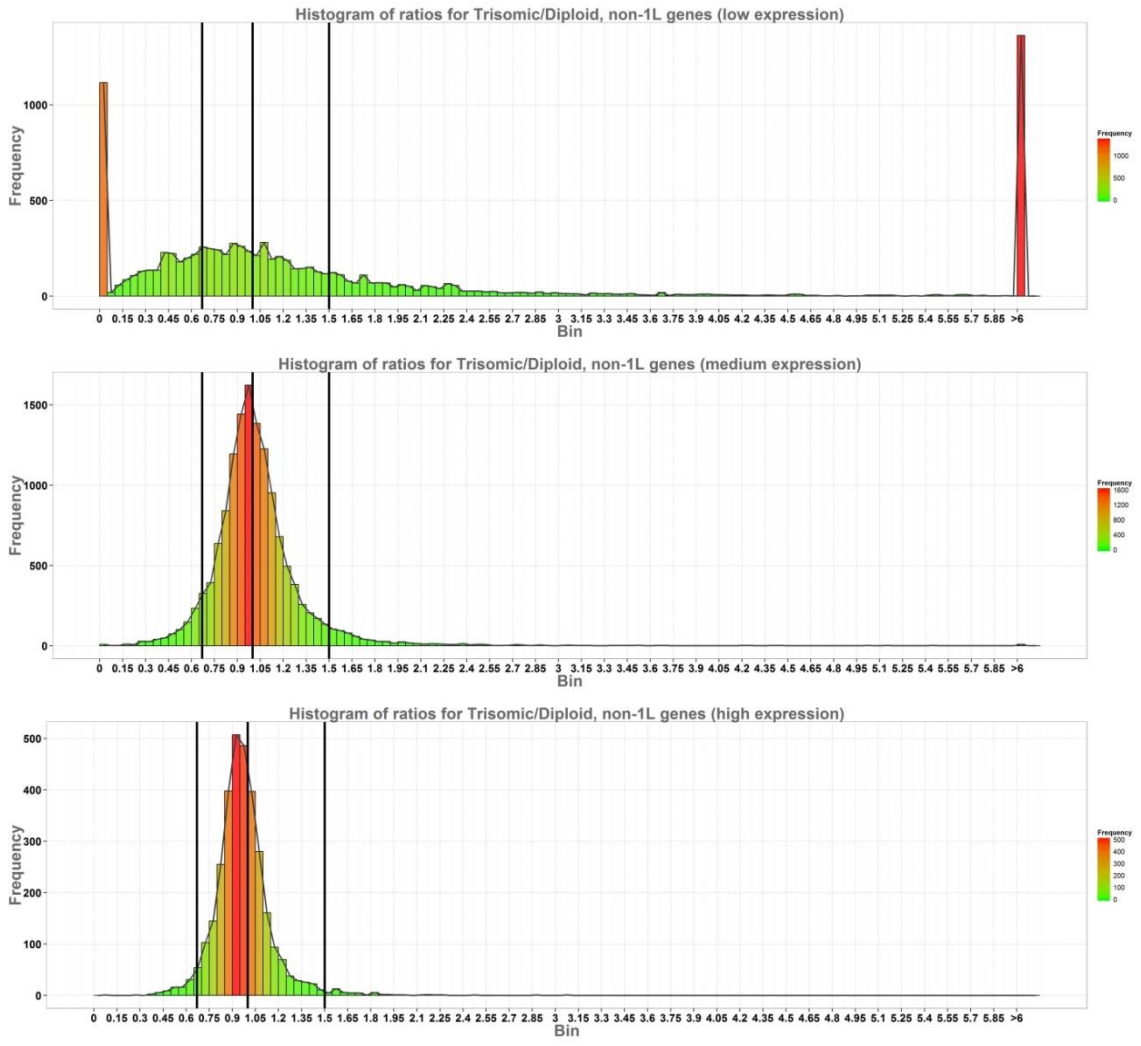


**Figure S4f – Ratio distribution plots for trans genes, trisomic/diploid, second dosage series**

For explanation see Figure S3a.

# VITA

Adam Franklin Johnson grew up in Colorado Springs, Colorado. He completed an Associate's degree from Pikes Peak Community College in 2004, and a Bachelor's degree with a major in linguistics from the University of Colorado in Boulder in 2006. Over the years, he has held an unusual variety of jobs, including roller-skate rink DJ and bookmobile driver. While working in the auto insurance industry after college, he decided to pursue a second Bachelor's degree, and graduated with a major in biology from the University of Colorado in Colorado Springs in 2011. He entered the University of Missouri's graduate program that year, and joined the lab of Dr. James Birchler in 2012. Adam and his better half, Linh To Ngo, share a love of travel, a reading habit, and a cat.