Lehrstuhl für Genetik der
Technischen Universität München

Genetic and transcriptome analysis of autopolyploid *Arabidopsis thaliana* (L.) Heynh.

Zheng Yu

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangungdes akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender:	Univ.- Prof. Dr. A. Gierl
Prüfer der Dissertation:	apl. Prof. Dr. R. A. Torres Ruiz
	Univ.- Prof. Dr. E. Grill

Die Dissertation wurde am 13.10.2014 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 24.11.2014 angenommen.

Table of contents

Zusammenfassung

Polyploidie, der Besitz von mehr als den doppelten Chromosomensatz, ist ein wichtiger Faktor in der Evolution von Eukaryonten gewesen. In der Regel werden zwei Formen der Polyploidie besonders herausgestellt: Allopolyploidie und Autopolyploidie. Erstere entsteht im ersten Schritt aus der Hybridisierung zweier Spezies auf die eine Chromosomen-Verdoppelung folgt. Letztere entsteht aus der direkten Verdoppelung der Chromosomensätze innerhalb ein und derselben Spezies. Autopolyploidie ist in Pflanzen häufiger anzutreffen als noch vor kurzem angenommen, hat aber im Vergleich zur Allopolyploidie wenig Aufmerksamkeit erhalten. Synthetisch allopolyploidisierte Pflanzen weisen beträchtliche Transkriptom-Veränderungen auf, die zum Teil wahrscheinlich auf die Vereinigung ehemals divergenter regulatorischer Schaltwege zurückzuführen sind. Im Gegensatz dazu haben Autopolyploide relativ uniforme Genome was nur wenige Veränderungen der Genexpression vermuten ließ.

In dieser Arbeit wurde mit Hilfe von Colchizin eine Serie von autotetraploiden *Arabidopsis thaliana* Pflanzen unterschiedlicher Ökotypen generiert. Zusätzlich wurden Triploide aus der Kreuzung dieser Tetraploiden und gewöhnlicher Diploiden erstellt. Die Tetraploiden wurden mit verschiedenen morphologischen und zytologischen Mitteln und zum Teil mittels LC-MS charakterisiert.

Im Vergleich zu Diploiden zeigten autotetraploide *Arabidopsis thaliana* relativ häufig Transkriptom-Änderungen. Diese bzw. deren Frequenz war von der Herkunft der Genome abhängig. Die Expressionsänderungen betrafen Gengruppen, die schon in Allotetraploiden auffällig gewesen waren, wie auch neue Gruppen. Die Expressionsänderungen waren stabil, nicht stochastisch, entwicklungs-spezifisch und teilweise mit Methylierungsänderungen der DNA verbunden. Der Vergleich der Genexpression zwischen diploiden und triploiden *A. thaliana* des Col-0 Ökotyps wies nur wenige Unterschiede auf. Nur drei der gefundenen annotierten Gene, darunter *MRD1*, waren sowohl in tetraploiden wie auch triploiden Col-0 Pflanzen überexprimiert. Diese Befunde weisen auf einen interessanten Unterschied zwischen der Wirkung gerad- und ungeradzahliger Chromosomenausstattung hin.

Ein besonderer Fokus wurde auf *MRD1* gelegt. Verschiedene, durch Kreuzung generierte, tetraploide und triploide Hybridpflanzen wiesen einen deutlichen Einfluss der Genomherkunft und der DNA-Methylierung auf die Expression von *MRD1* auf.

Daten des *Arabidopsis* Massive Parallel Signature Sequencing Projekts weisen diesbezüglich auf eine mögliche Rolle von kurzen RNAs hin.

In einem weiteren Teil der Arbeit wurde ein Teil einer “Recombinant Inbred Line“-Population (RIL-Population) mit Colchizin behandelt um eine Konvertierung in Tetraploide zu erreichen. Sechsundzwanzig dieser Linien wurden als tetraploide mittels Durchflußzytometrie bestätigt. Das Samengewicht war in den tetraploiden Linien signifikant höher als in den diploiden Ursprungslinien. Eine Kartierung wurde von Polyploidie-Effekten versucht. Für ein Genomintervall könnte eine Kopplung für die Entwicklung abnormer Blüten, die gelegentlich in tetraploiden RIL-Pflanzen auftaucht, gefunden worden sein.

Summary

Polyploidy, the acquisition of more than two sets of chromosomes, has been an important factor in the evolution of eukaryotes. Two forms of polyploidy are often considered: allopolyploidy and autopolyploidy. The former originates from interspecies hybridization followed by genome duplication. The latter originates from intraspecies genome duplication events. Autopolyploidy is more common in plants than traditionally assumed, but has received little attention compared with allopolyploidy. Synthetic allopolyploid plants exhibit considerable transcriptome alterations, part of these are likely caused by the reunion of previously diverged regulatory hierarchies. In contrast, autopolyploids have relatively uniform genomes, suggesting lower alteration of gene expression.

In this study, a series of unique *Arabidopsis thaliana* autotetraploids from different ecotypes was generated by using colchicine. In addition, triploid plants were generated through crosses of autotetraploids with diploids. The tetraploid lines were analyzed by different morphological and cytological means and in part by LC-MS.

A. thaliana autotetraploids showed, in comparison to diploids, relatively frequent transcriptome alterations that strongly depended on their parental genome composition. They included altered expression of both new genes and gene groups previously described from allopolyploid *Arabidopsis*. Alterations in gene expression were stable, nonstochastic, developmentally specific, and associated with changes in DNA methylation.The gene expression alteration between diploid and triploid *A. thaliana* of Col-0 ecotype concerned very few genes. Only three of the detected annotated genes, including *MRD1*, were over-expressed in tetraploids as well as in triploids. The difference between tri- and tetraploids points to an interesting odd- vs. even chromosome number effect.

A special focus was laid on *MRD1*. Further analysis in various tetraploid and triploid hybrids demonstrated the impact of the genome origin and the DNA methylation state on the expression of *MRD1*. Data from the *Arabidopsis* Massive Parallel Signature Sequencing Project point to the possible involvement of small RNAs in the expression of this locus.

In a separate section of this work, part of a Recombinant Inbred Line (RIL) population was subjected to colchicine treatment in order to generate a tetraploid RIL population. Twenty-six of these lines were confirmed by flow cytometry analysis as tetraploids. The seed weight of these lines was significantly higher if compared to diploid lines. A “polyploidy effect” mapping was performed. One genome region may link to the generation of abnormal flowers occasionally found RIL tetraploid plants.

Acknowledgments

I would like to thank Prof. Dr. Alfons Gierl for providing me the opportunity to carry out my PhD study in his laboratory at the “Lehrstuhl für Genetik”.

I would like to express my extreme gratitude to my research supervisor, Prof. Dr. Ramon A. Torres-Ruiz, for his unceasing support, guidance, research and academic advice during the course of my study and writing. Thank you so much for being so patient with me like a father, for so much teaching and discussions to broaden my knowledge base, as well as for providing a good research team and working place, scientific resources to expand my repertoire of experimental techniques throughout my study.

I would like to express my sincere appreciation to Dr. Monika Frey and Dr. Erich Glawischnig for support, guidance, kindness, and academic advice.

I thank Dagmar Engl, Stephan Kotschote and Kerstin Stegmüller (both GPC-Biotech), Michael Pfaffl, Dirk Haller (Technische Universität München), Dr. Gert Daniel (Bayer Landesanstalt für Landwirtschaft, Freising, Germany), Jörg Durner, Dr. Uta von Rad, Dr. Tony Schaeffner (Helmholtz Zentrum), and Götz Frommer (Agilent) for their helpful support.

My work on this dissertation has been aided in direct and indirect ways by numerous colleagues at the “Lehrstuhl für Genetik”, including Otti Peis, Dr.Birgit Treml, Dr. Miriam Vogg, Dr. Rafał Kazimierz Jończyk, Dr. Lilla Römisch-Margl, Dr. Regina Dick, Dr. Thomas Rauhut, Dr. Katrin Schullehner, Dr. Gerti Spielbauer, Dr. Ruohe Yin, Andreas Fieselmann, Regina Huttl, Peggy Muller, Carolin Ziegler, Petra Wick, Peter Dobos, Dr. Ulrich Genschel and Dr. Holger Schmidt.

I would like to thank my husband, my father and my parents in law, for their support. Thank my two little kids for the laughs they bring to my life.

Abbreviation Index

A	Adenine
ABRC	*Arabidopsis* Biological Resource Center
ABCB1	Adenosine Triphosphate Binding Cassette subfamily B member 1
Act	Actin
adhA	Alcohol dehydrogenase A
AGP1	Arabinogalactan Protein 1
AP2	APETALA2
ARF	Auxin Response Factor
ATEXPA	Arabidopsis Thaliana Expansin A
A. thaliana	*Arabidopsis thaliana*
ATP	Adenosine Triphosphate
ATPase	Adenylpyrophosphatase
Bidest	Bidistilled
Bor	Borky
bp	Base Pair(s)
BSA	Bovine Serumalbumin
Bur	Burren
CAPS	Cleaved Amplified Polimorphic Sequences
cDNA	Complementary DNA
CIA	Chloroforme:iIsoamylalcohol (24:1)
cM	Centimorgan
Col (C)	Columbia
Ct	Catania
conc	Concentrated
CTAB	Cetyltrimethylammoniumbromide
℃	Degrees Celsius
Cy3 or Cy5	Cyanine 3 or Cyanine 5 dye
DAPI	4',6-Diamidino-2-Phenylindol-Dihydrochloride
dATP	Deoxyribo-Adenosine Triphosphate
dCTP	Deoxyribo-Cytidine Triphosphate
DEPC	Diethylpyrocarbonate
dest	Distilled
dGTP	Deoxyribo-Guanosine Triphosphate
DIG	Digoxigenin
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleotide-Triphosphate
DRM2	Domain Rearranged Methyltransferase 2
dTTP	Deoxyribo-Thymidine Triphosphate
EDTA	Ethylendiaminetetraaceticacid
ENP	*ENHANCER OF PINOID*
enp	*Enhancer of Pinoid*
EtOH	Ethanol
FC	Fold Change
FLC	*Flowering Locus C*
FMF	Familial Mediterranean Fever
FRI	*FRIGIDA*
G	Guanine
g	Gram
GO	Gene Ontology
GUS	β-Glucoronidase
h	hour(s)

HAK 5	High Affinity K^+ Transporter 5
HCl	Hydrogenchloride
HEC2	HECATE2
H_2O	Water
H_2O_{dd}	Deionised water
H_2O_{dest}	Distilled Water
H_3PO_4	Phosphoric Acid
HPT	Hygromycin Phoshpotrasferase
IAA	*Indoleacetic Acid induced*
kb	Kilo base pairs
KCl	Potassium chloride
KOH	Potassium hydroxyde
L	Liter
LAC 1	Laccase 1
LB	Luria-Bertani
LC-MS	Liquid Chromatography-Mass Spectrometry
L*er* (L)	Landsberg *erecta*
LOD	Logarithm of Odds
M	Molar
MatDB	MIPS *Arabidopsis thaliana* data base
Mb	Megabase(s)
Met	Methyltransferase
mg	Milligram
min	Minute(s)
MIPS	Munich Information Center for Protein Sequences
mL	Milliliter
mM	Millimolar
MOPS	3-(N-morpholino)Propanesulfonic acid
MPK 4	Map Kinase 4
MPSS	Massively Parallel Signature Sequencing
MRD	*Mto 1 Responding Down*
MS	Murashige and Skoog
Mto	Methionine
MN	Macherey Nagel (company)
NADPH	Nicotinamide adenine dinucleotide phosphate
Nd	Niederzenz
NEB	New England Biolabs (company)
µg	Microgram
µL	Microliter
N	Normal
NaAc	Sodiumacetate
NaCl	Sodiumchloride
NaOH	Sodiumhydroxide
NASC	Nottingham *Arabidopsis* Stock Center
Nd (N)	Niederzenz
NEP1	Necrosis and ethylene inducing Peptide 1
ng	Nanogram
nm	Nanometer
no.	Number
PAS	PASTICINO
PCR	Polymerase Chain Reaction
pg	Picogram
PID	*PINOID*
pid	*pinoid*

PIN	*PINFORMED*
Pmol	Picomol
PVP	Polyvinyl Piridine
q RT PCR	Quantitative Reverstranscription Polymerase Chain Reaction
QTL	Quantitative Trait Locus
RdDM	RNA directed DNA Methylation
RDR2	RNA Dependent RNA Polymerase 2
RFLP	Restriction Fragment Length Polymorphism
RI	Recombinant Inbred
RIL	Recombinant Inbred Lines
RIN	RNA Integrity Number
RISC	RNA Induced Silencing Complex
RNA	Ribonucleic Acid
ROS	*Repressor of silencing*
rpm	Revolutions Per Minute
RT	Room Temperature
SAM	Shoot Apical Meristem
SAUR	*Short Auxin Up-regulated RNAs*
SDI	*Sensitive to Dosage Imbalace*
SDS	Sodiumdodecylsulfate
Sec	Second(s)
SNP	Single Nucleotide Polymorphisms
SSLP	Short Sequence Length Polymorphisms
SSR	Simple Sequence Repeats
STM	*Shootmeristemless*
T	Thymine
TAIR	The *Arabidopsis* Information Resource
Taq-polymerase	*Thermus Aquaticus*-DNA-polymerase
TBE	Tris-Borate-EDTA-buffer
TE	Tris-EDTA-buffer
TEA	Triethanolamine
Tris	Tris-(hydroxymethyl)aminomethane
tRNA	Transfer ribonucleic acid
Ts	Tossa de Mar
U	Unit
UV	Ultra Violet
v/v	Volume per volume
w/v	Weight per volume

1. Introduction

1.1 Polyploidisation in evolution and speciation

Polyploidy is important for plant evolution

Polyploidy is the heritable condition of possessing more than two complete sets of chromosomes. Most polyploids have an even number of sets of chromosomes, with four being the most common (tetraploidy). Polyploidy is a common phenomenon in the evolution of plants (Wendel, 2000) and some animal clades (Ohno, 1970). It was estimated that 50 to 80% of angiosperms are polyploids, including crop plants such as alfalfa (Medicago sativa), potato (Solanum tuberosum), wheat (Triticum aestivum), oat (Avena sativa), cotton (Gossypium hirsutum), and coffee (Coffea arabica) (Wendel, 2000). Moreover, complete sequencing of the nuclear genome has revealed evidence of ancient polyploidy throughout angiosperms and in other eukaryotes (Soltis and Soltis, 2009). All plant nuclear genomes sequenced to date showed evidence of ancient genome duplication: *Arabidopsis*, *Oryza*, *Populus*, *Vitis*, and *Carica* (reviewed by Soltis and Soltis, 2009). The EST (expressed sequence tags) approach identified ancient polyploidy in a number of crops, including *Zea* (maize), *Glycine* (soybean), and *Gossypium* (cotton) (reviewed by Soltis and Soltis, 2009). The genome increase of these paleopolyploids was then gradually dismantled during evolution. This “genome downsizing” (Bennett and Leitch, 2005) was accompanied by whole genome or chromosome rearrangements as well as non-stochastic retention of functionally related gene duplications (Blanc and Wolfe, 2004; Adams et al., 2004; Adams and Wendel, 2005; Wang et al., 2004/2006a).

Extensive genomic rearrangements, including exchanges between genomes and gene loss, often arise with the onset of polyploidization (Levy and Feldman, 2002). Genome doubling also significantly affects gene expression, resulting in epigenetically induced gene silencing. Novel phenotypes were known to emerge from this genomic amalgam, including some with high visibility to natural selection, such as organ size and flowering time. Thus, polyploidy could be a prominent and significant force in plant evolution, at temporal scales ranging from ancient to contemporary, and with profound effects at scales ranging from molecular to ecological (Adams et al. 2004).

Genomic rearrangement following polyploidysation promotes the polyploid speciation

Differential gene loss (i.e. loss of some duplicates but not others) following polyploidy is responsible for much of the deviation in co-linearity among relatively closely related plants, such as cereals. When differential gene loss is considered, co-linearity of genes on orthologous chromosomes among cereal (and other) genomes is higher than previously recognized (Adams and Wendel, 2005). Particularly illustrative of the process of gene removal following genome duplication were the studies from maize that demonstrated that about half of all duplicated genes had been lost in the approximately 11 million years since the polyploidy event that gave rise to the progenitor of maize (Messing et al., 2004; Lai et al., 2004; Ilic et al., 2003). On a longer evolutionary timescale, the cumulative effects of these twin processes of genome doubling and gene loss have created modern angiosperm genomes that exhibit clustered, hierarchical networks of synteny with only partial gene membership of any single linkage group, as well as the differential survivorship of duplicated genes. This phenomenon extended beyond angiosperms, as elegantly shown by a comparative analysis of genome sequences in common baker's yeast, *Saccharomyces cereviseae*, and its close relative *Kluyveromyces waltii*. Sequence data unambiguously demonstrates a 2:1 relationship of linkage groups for these two lineages, with the duplicated chromosomes in yeast having mostly decayed into collective single-copy status (from the review by Adams and Wendel, 2005).

The retention of duplicate genes is non-random and can happen either when the duplicated genes are completely functionally redundant but the dosage effect presents a selective advantage (Osborn et al., 2003) or when their function diverges. Functional divergence can occur by neofunctionalization (a gene copy acquires a new function) or by subfunctionalization (the copies retain different subsets of the functionality of the ancestral gene (Force et al., 1999).

Arabidopsis thaliana has undergone at least two and probably three paleopolyploidy events during the evolutionary history. Approximately 27% of the gene pairs that were formed by polyploidy have been retained and more than half of these gene pairs show evidence of functional divergence (Blanc and Wolfe, 2004).

1.2 Mechanisms of polyploidy formation

Polyploids can arise by somatic doubling, by the fusion of unreduced gametes, and by means of a triploid bridge (Fig. 1, Rieseberg and Willis, 2007; Comai, 2005).

In our study, the polyploids were generated by somatic doubling with the chemical, colchicine, which disrupts the spindle apparatus resulting in the failure of the sister chromatide to separate. However, unreduced gametes are common in plants and likely represent the most frequent route to polyploidy (Ramsey et al., 2007). Diploid gametes, which arise infrequently, typically fuse with haploid ones and produce triploid zygotes, which are unstable and can either be sterile or contribute to further polyploid gametes, depending on the species. The fusion of diploid gametes leads to tetraploid zygotes, which are potentially stable (Ramsey et al, 2007).

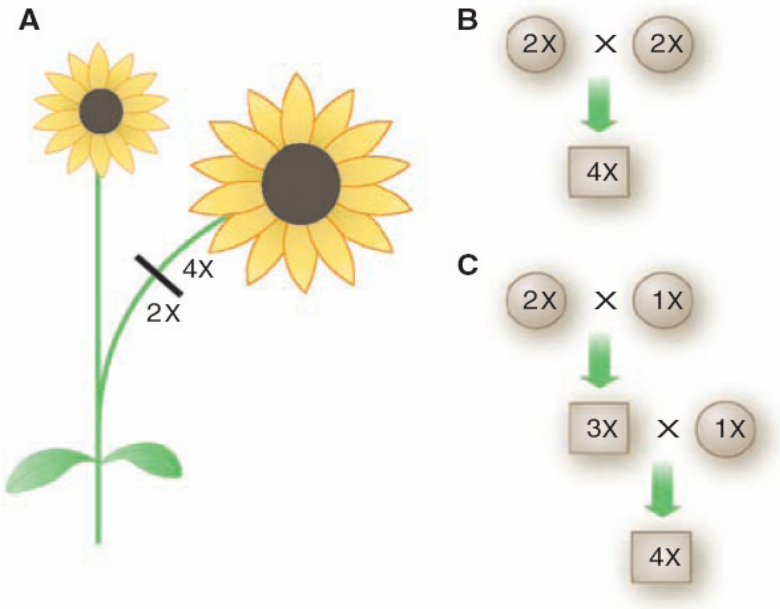


Fig. 1: Mechanisms by which polyploids can arise
(A) Somatic doubling, in which chromosome number is doubled in vegetative tissue that gives rise to reproductive organs. (B) Fusion of unreduced gametes that are produced when cell walls fail to form in the final stage of meiosis. (C) A triploid bridge, in which unreduced and reduced gametes form triploids. If the triploids also produce unreduced gametes, the triploid gametes may fuse with reduced gametes from diploid individuals to generate stable tetraploids (taken from Rieseberg and Willis, 2007).

1.3 Types of polyploids

Depending on the origin of the genome, polyploids can be classified into allopolyploids and autopolyploids. Allopolyploids result from genome doubling of interspecies hybrids, whereas autopolyploids result from intraspecies genome doubling. Thus, an autotetraploid will contain four copies of each chromosome (all four being homologs), whereas an allotetraploid will contain two of each pair of the counterpart chromosomes derived from two different species (homeologous chromosomes). Both allopolyploid and autopolyploid are extremes, which are connected by a continuum of overlapping forms (Grant, 1981; Soltis et al., 2003). Among known polyploid plants, allopolyploids show a taxonomic predominance.

However, increasing evidence indicates that the actual appearance of autotetraploid plants in nature might be significantly underestimated. Because autopolyploidy may be difficult to detect, its frequency may still be underestimated and underappreciated (reviewed by Soltis and Soltis, 2009).

1.4 Polyploidy can lead to immediate and extensive changes in gene expression

Polyploidy is a fundamental but relatively under-explored biological process. It is wide spread but little is known about how duplicate genes and genomes function in the early stages of hybridization, and how the duplicate genes maintain and diverge functions during plant evolution and crop domestication. Resynthesized polyploids with known progenitors are excellent materials for dissecting gene expression and genomic changes in early stages and comparisons with older polyploids.

To succeed, newly occurring polyploids must overcome notable challenges: genomic instability based on aberrant chromosome segregation during meiosis (Ramsey and Schemske, 2002; Comai, 2005; Mallet, 2007), and rapid adaption to selective environmental pressures that includes competition, for instance, with their diploid progenitors (Soltis et al., 2003, Osborn et al., 2003). Cascades of novel expression patterns, altered regulatory interactions and new phenotypic variation occurred after polyploidisation (reviewed by Adams and Wendel, 2005).

Differential gene expression study in Allopolyploid

Studies on differential gene expression and transcriptomics have mainly focused on (neo-)allotetraploids such as wheat, cotton, maize and prominently resynthesized *Arabidopsis suecica*, which originated from *A. thaliana* and *A. arenosa* (Comai et al., 2000; Kashkush et al., 2002; Madlung et al., 2002; Wang et al., 2004; Adams et al., 2004; Wang et al., 2006a, 2006b). The transcriptome analysis of two independently generated *A.suecica* lines uncovered 5-6% genes to display expression divergence from the midparent value. The majority of these genes also displayed differential expression between the two parents. Remarkably, most of these (>90%) were non-additively repressed genes that normally exhibit higher expression levels in *A. thaliana* than in *A. arenosa*, a phenomenon, which is consistent with the silencing of *A. thaliana* rRNA genes subjected to nucleolar dominance and with overall suppression of the *A. thaliana* phenotype in the synthetic allotetraploids (Chen et al., 1998). Interestingly, the non-additive regulation often affects gene groups involved in particular physiological and developmental pathways (Wang et al., 2006b). For example, nonadditive regulation of FRI and FLC loci mediates flowering-time

variation in *Arabidopsis* Allopolyploids. Synthetic *Arabidopsis* allotetraploids contain two sets of *FLC* and *FRI* genes originating from *A. thaliana* and *A. arenosa*, respectively, and flower late. Inhibition of early flowering is caused by upregulation of *A. thaliana FLC* (At*FLC*) that is trans-activated by *A. arenosa FRI* (Aa*FRI*). Two duplicate *FLC*s (Aa*FLC1* and Aa*FLC2*) originating from *A. arenosa* are expressed in some allotetraploids but silenced in other lines. The expression variation in the allotetraploids is associated with deletions in the promoter regions and first introns of *A. arenosa FLC*s. The strong At*FLC* and Aa*FLC* loci are maintained in natural *Arabidopsis* allotetraploids, leading to extremely late flowering (Wang et al., 2006a). Furthermore, *FLC* expression correlates with histone methylation and acetylation. This indicates that interactive roles of regulatory sequence changes, chromatin modification and trans-acting affect the fate of duplicate genes and adaptation of allopolyploids during evolution (Wang et al., 2006a).

Silencing and relative expression levels of genes duplicated by polyploidy can be variable in different parts of the plant, indicating differential regulation of the two homoeologs during plant development (Adams et al., 2003; Wang et al., 2004; Madlung et al., 2005). In cotton, there is considerable variation in the relative expression levels and silencing patterns of duplicated gene pairs among organ types, especially in different floral whorls, beginning at the first generation after polyploidy. For example, expression of a calmodulin binding protein and an ubiquitin-protein ligase ranged from silencing of one homoeolog in some organs to equal expression of both homoeologs in other organs (Adams et al., 2004). Complete partitioning of expression between homoeologs can occur in different organs. For example, only one homoeolog of the *alcohol dehydrogenase A* (*adhA*) gene is expressed in cotton petals and only the other homoeolog in styles (Adams et al.,2003), indicative of subfunctionalization. The developmentally regulated patterns of reciprocal *adhA* homoeolog silencing are strikingly similar in natural cotton polyploids and in two synthetic allopolyploids. These observations suggest the possibility of instantaneous, epigenetic regulatory alteration that might be evolutionarily stable for more than one million years (review by Adams and Wendel, 2005).

Differential gene expression study in autopolyploid

In allotetraploids, many of the observed gene expression phenomena might probably result from conflicts of the two concurring genomes, which have diverged during evolution. In autopolyploids where the genomes are identical, this source of variability is not given and gene expression alterations are expected to be low. This is confirmed in potato autopolyploids by analyzing part of their transcriptome. The analysis of 9,000 genes in potato autopolyploids revealed few very weak differences

in comparison with diploids (Stupar et al., 2007). A control experiment included in the analysis of the *Arabidopsis suecica* transcriptome compared di- and tetraploid seedlings of the *Arabidopsis thaliana* ecotype L*e*r and uncovered only 88 genes differentially expressed (Wang et al., 2006b). This is reminiscent of the dosage-dependent regulation of a dozen genes as observed in yeast autoploids (Galitski et al., 1999).

Paramutation-like interaction of the epialleles in autopolyploids

Mittelsten-Scheid and her coworkers uncovered the paramutation-like interaction of the stable epialleles in autotetraploid *A. thaliana*. In their work, they generated the autotetraploid derivatives of line *A.thaliana* C. The diploid line *A. thaliana* C, homozygous with respect to a *hygromycin phosphotransferase* (*HPT*) transgene, showing uniform hygromycin resistance over many generations of self-pollination and mendelian segregation when outcrossed with diploid wild-type plants (Mittelsten-Scheid et al., 1996). Among the autotetraploid lines, one had only resistant progeny (C4R: tetraploid, resistant) and two had completely sensitive progeny (C4S1 and C4S2: tetraploid sensitive). They reciprocally crossed C4R with C4S1 or C4S2. F1 seedlings of all crosses (RRSS) were uniformly hygromycin-resistant. Tetraploid F2 progeny should have a minimum of 2.8% (1 of 36) seedlings containing only the previously active epialleles (RRRR) and be fully resistant to hygromycin. However, none of the 742 germinated seedlings survived in prolonged hygromycin selection among three independent F2 populations. This suggested that encounter of R with S epialleles in the RRSS hybrid resulted in heritable inactivation of R epialleles that persisted even after segregation away from the inactivating S allele. This resembles paramutation, the interaction of a dominant paramutagenic allele (silent and causing silencing) with a susceptible paramutable allele (able to acquire silencing), leading to heritable conversion of the paramutable into a new paramutagenic allele. Thus, they demonstrated that epialleles in tetraploid plants (but not in diploids) interact in *trans* and lead to heritable gene silencing persisting after segregation from the inactivating allele. Such interactions probably contribute to rapid adaptation and evolution of polyploid plant species (Mittelsten-Scheid et al., 2003).

1.5 Aneuploids

Polyploidisation often involves aneuploidy (the condition resulting from extra or missing chromosomes relative to the normal chromosome number of a species) as a byproduct. The occurrence of univalents and multivalents during polyploid meiosis complicates the orderly separation of homologs/homeologues. Univalents and trivalents by necessity divide unequally during anaphase I because there is no mechanism to evenly divide the chromosomes of an odd-number configuration (though by chance, unbalanced divisions may compensate each other, for example by a 2–1 separation of a trivalent, and a 0–1 division of a univalent). The divisions of tetraploids are more complicated to assess. Some ten types of quadrivalent configurations can be formed, depending on which homologous/homeologous chromosomes happen to cross over (Singh 1993). Quadrivalent configurations can broadly be divided into ring configurations (each homolog/homeolog forming two chiasma) and chain configurations (each homolog/homeolog forms one or two chiasma), analogous to the ring and rod configurations of diploids. Among ring and chain configurations, one may distinguish alternate orientations (proximate homologs/homeologs oriented in opposite directions) and adjacent orientations (proximate homologs/homeologs oriented, to varying degrees, in the same direction). Alternate quadrivalent orientations, sometimes called zigzag orientations, are believed to nearly always generate equal (2–2) chromosome disjunctions (Garber 1955, McCollum 1958), whereas disjunctions from adjacent orientations will include both balanced and unbalanced separations. The unbalanced separation results in aneuploidy (reviewed by Ramsey and Schemske, 2002).

The relative excess or deficiency of specific chromosome types in aneuploid, results in gene dosage imbalance. The proper functioning of cells and organisms relies on molecular complexes, which require a delicate balance between components for proper operation (Papp et al., 2003). Even a slight departure from this balance can have dramatic phenotypic or developmental consequences (Papp et al., 2003; Veitia, 2005) as exemplified by many genes identified in humans as tumor suppressors (Fodde and Smits, 2002) and as essential or regulatory genes in yeast (Giaever et al., 1999) and *Drosophila* (Birchler et al., 2001). In aneuploids, where dosage variations affect whole chromosomes rather than single genes, the consequences can be severe when the copy numbers of many dosage-sensitive genes are altered at once. Therefore, an alteration of gene dosage as occurs in aneuploids typically has unfavorable consequences. However, aneuploidy is not always deleterious and can be persistent. For example, aneuploid cells are normally found in certain tissues such

as the brain and the placenta, where they appear to play a functional role (Kingsbury et al., 2005, Weier et al., 2006). Aneuploid individuals are common in plants and in yeast and provide a pool of phenotypic variation not present in the euploid population. Comai and his coworkers showed that the viability of the aneuploid (from the natural tetraploid Wa-1) associates with the distortion of the *Sensitive to Dosage Imbalance* (*SDI*) allele. The distortion was greatest in the aneuploids facing the strongest viability selection (Henry et al., 2007).

1.6 Aims of this work

The studies of polyploidy have mainly focused on allopolyploidy. The importance of autopolyploidy in plant evolution may be underestimated (Soltis and Soltis, 2009). In allotetraploids, many of the observed gene expression phenomena might probably result from conflicts of the two concurring genomes. Analysis of autopolyploid may exclude this source of variation and produce the "pure polyploidy effect". However, the previous work on autopolyploidy is limited. In potato, only part of the genome was analyzed. In *Arabidopsis*, only one developmental stage (seedling) from a single tetraploid line (N3900/CS3900) from the ecotype Landsberg has been analyzed (Wang et al., 2006b). The aims of this work, therefore, were to generate autotetraploids from different ecotypes and to find whether significant gene expression alterations could be found among these newly synthesized autopolyploids. The work aimed to address several questions: whether the alterations of the gene expressions are developmental stage dependent; or whether the alterations are ecotype dependent; whether epigenetic effects accompany gene expression alterations; whether the expression pattern of the polyploids after neopolyploidsation is relatively stable or will be altered rapidly in the following generations and whether the gene expression pattern will be altered vigorously when the ploidy level is elevated to three, i.e. an odd number of chromosomes.

The transcriptomic analysis is a way "from gene to phenotype" to get the "polyploidy effects". This work also aimed to find "polyploidy effects" with the way "from phenotype to gene". To this end, the phenotype alteration, especially the seed weight alteration from diploid to tetraploid was subjected to a mapping strategy using the diploid and tetraploid recombinant inbred lines in order to narrow down possible responsible genomic region(s) for the phenotype alteration in tetraploids.

2. Material and Methods

All laboratory reagents used in this work are obtained commercially (analytical grade) from the Companies BioRad (USA), Fluka (Switzerland), Merck (Darmstadt, Germany), Roth (Karlsruhe, Germany), Sigma-Aldrich (USA), Serva (Heidelberg, Germany), Duchefa (Netherlands), Promega (USA), Qiagen(USA), New England Biolabs (USA), Agilent (USA), Applied Biosystems (USA), Roche(USA), Macherey Nagel (Düren, Germany), Partec(Münster, Germany), GE Healthcare(UK)

Oligonucleotides were obtained from Biomers (Ulm, Germany).

The molecular biology protocols, unless otherwise stated, were as described by Sambrook and Russell (2001).

2.1 Material

2.1.1 Plant material

Ecotype-lines	Source
Ler (NW20)	NASC / ABRC
Col-0 (N1092)	NASC/ABRC
Nd-0 (N1390)	NASC/ABRC
Nd-1 (CS22619)	NASC/ABRC
Bur-0 (CS22656)	NASC/ABRC
Bor-1 (CS22590)	NASC/ABRC
Ts-1	NASC/ABRC
Ct-1 (CS22639)	NASC/ABRC
Pro-0	NASC/ABRC
Ler-1 (CS22618)	NASC/ABRC
C24 (CS22620)	NASC/ABRC
Zürich	Ortrun Mittelsten-Scheid (GMI Wien)
RI lines	Lister and Dean Col x Ler-RI-population (NASC, Lister and Dean, 1993)

Transgenic lines	Source
CYCAt1:CDB:GUS	D. Celenza (via M.-T. Hauser, University BOKU Wien)
DR5rev.::GFP	G. Jürgens/J. Friml (Univ. Tübingen)
SALK_077688	NASC
SAIL_819_F03	NASC
SALK_034132	NASC
SALK_118101	NASC

SALK_080613	NASC
SALK_014624	NASC

2.1.2 Bacteria

Strain	Genotype
E. coli XL1 Blue	supE44 hsdR17 recA1 endA1 gyrA46 thi relA1 laclac [F' proAB+ laclq lacZΔM15 Tn10 (Tetr)]; (Bullock et al., 1987)

2.1.3 Kits and enzymes

Enzyme	**Usage**	**Source**
McrBC	Methylation profiling	NEB
Tag polymerase	PCR	Promega
Kit	**Usage**	**Source**
Reverse transcription reagents	RT-PCR	TaqMan
Lightcycler FastStart DNA master^{plus} SYBR Green I	qRT-PCR	Roche
Lightcycler480 SYBR Green I Master	qRT-PCR	Roche
NucleoSpin	DNA isolation and purification	MN
RNeasy Plant Mini Kit	total RNA isolation	Qiagen
Illustra^{TM} GFX^{TM} PCR DNA and Gel Band purification kit	DNA band purification	GEHealthcare
SV Total RNA isolation System	total RNA isolation	Promega
Low RNA Input Linear Amplification Kit two color	Microarray	Agilent
RNA Spike In kit, two color	Microarray	Agilent
Gene Expression Hybridization Kit	Microarray	Agilent

2.1.4 Solutions

100x Denhardt's

2% (w/v) Ficoll 400

2% (w/v) Polyvinylpyrrolidon

2% (w/v) BSA

Dilute them in H_2O_{dest} , sterilize with filtration and aliquot and keep the aliquots in -20℃ freezer.

2x CTAB-Buffer

CTAB	2%
PVP	1%
NaCl	1.4M
Tris·HCl (pH 8.0)	100mM
EDTA (pH 8.0)	20mM

Take about 3/4 of the desired volume of H_2O_{dest} and add the proper quantities of the solids under continuous stirring. Then add the rest of the components and of the water and stir until the solids are completely dissolved.

CIA

Chloroform:Isoamylalcohol	24:1

DEPC-H_2O

Dissolve 0.1% DEPC in H_2O_{bidest}, and place it over night under the fume hood, then autoclave.

dNTP-Solution (for PCR-Reactions)

dATP	2 mM
dCTP	2 mM
dGTP	2 mM
dTTP	2 mM

Dissolve in H_2O_{dest}, aliquot and store at –20 ℃.

0.5 M EDTA pH 8.0

EDTA	0.5M

Add 3/4 of the necessary H_2O_{dest} and adjust to pH 8.0 with 10N NaOH under continuous stirring. Then fill up to the final volume. (EDTA will only dissolve completely when pH 8.0 is reached.)

Ethidiumbromide-Stock

Ethidiumbromide	10 mg/mL

Formamide

Deionise the formamide (about 200 mL) over night with one spoon of serdolite. Store light protected at 4°C.

6x Loading Dye (for Agarose Gels)

I: bromphenol blue	0.25%
Ficoll type 400	15%
II: xylene cyanol	0.25%
Ficoll type 400	15%

For both buffers dissolve the solids in small portions in 3/4 H_2O_{dest} of the total volume, then fill up to the final volume.

Solutions I – III for Plasmid - Isolation (Midi-Prep)

Solution I:

Glucose	50 mM
Tris·HCl pH 8.0	25 mM
EDTA pH 8.0	10 mM

Fill up with H_2O_{dest} to the desired volume and autoclave.

Solution II:

NaOH	0.2 M
SDS	1%

Solution III:

NaAc	3 M

Adjust to pH 4.8 with glacial acetic acid, then fill up to the desired volume with H_2O_{dest} and autoclave.

3 M NaAc pH 5.2

NaAc (water free)	3 M

Fill up to 3/4 of the desired volume with H_2O_{dest}, then adjust to pH 5.2 with glacial acetic acid. Finally fill up to the total volume.

Phenol

It is necessary to get the pH > 7.8, because pH < 7.8 will result in loss of DNA in the organic phase during phenol/chloroform extraction.

- Add 0.1% hydroxychinolin (works as an antioxidant, inhibitor of RNases and

chelator of metal ions)

- Add 1 volume of 1M Tris pH 8.0
- Stir at RT for 15 min
- Turn off the magnet stirrer to let the phases separate, take off the upper phase with a glass pipette
- Add 1 volume of 0.1M Tris pH 8.0 and repeat the process until the pH in the lower phenolic phase is more than 7.8. The pH in the supernatant should be between 6.5 and 7.5 (use pH sticks from Merck).
- Cover with 0.1 M Tris pH 8.0 and 0.2% ß-Mercaptoethanol.

Phenol can be stored light protected at 4°C for several months.

10% SDS

SDS	10%

Fill up under the fume hood with H_2O_{dest} to the desired volume and stir until the SDS is completely dissolved.

5x TBE

Tris	54 g
Boric acid	27.5 g
0.5M EDTA (pH 8.0)	20 mL

Add the solids and the EDTA to 3/4 of the desired volume (1 L) under continuous stirring. When all components are dissolved add up to the final volume with H_2O_{dest}.

50xTAE

40 mM Tris,
5 mM sodium acetate,
1 mM EDTA
Adjust to pH 7.8 with glacial acetic acid

1x TE pH 8.0

Tris- HCl	10 mM
EDTA·Na_2	1 mM

Dissolve the solids or mix the corresponding solutions (pH 8.0) under continuous stirring in 3/4 H_2O_{dest} of the desired volume. When all components are dissolved adjust to pH 8.0 with concentrated HCl (37%), then fill up to the final volume with $H_2O_{dest.}$ Autoclave the solution.

T 0.1E Buffer

Tris-HCl	10 mM
EDTA·Na_2	0.1 mM

Dissolve the solids or mix the corresponding solutions (pH 8.0) under continuous stirring in 3/4 H_2O_{dest} of the desired volume. When all components are dissolved adjust to pH 8.0 with concentrated HCl (37%), then fill up to the final volume with $H_2O_{dest.}$ Autoclave the solution.

1 M Tris-HCl pH 8.0

Tris-HCl	1M

Add the Tris to 3/4 H_2O_{dest} of the desired volume and adjust to pH 8.0 with concentrated HCl (37%) under continuous stirring. Fill up to the final volume and autoclave.

Solution for RNA-Extraction from plant material

R1 Buffer

8 M Guanidin·HCl
20 mM EDTA (TITRIPLEX III)
20 mM MES
pH 7,0

Autoclave the solution. Keep the soltuion in room temperature and protect it from light. Before using add 50 mM ß-Mecaptoethanol.

Solution for Northern Blots

10x MOPS-Buffer

MOPS	200 mM
NaAc	50 mM
EDTA	10 mM

Adjust to pH 7.0 and fill up with H_2O_{dest} to the desired volume.
The buffer is not autoclaved, it is stored light protected at 4℃.
1x MOPS buffer is used as running buffer for formaldehyde-gels.

RNA-loading buffer (NBSB)

100 µl Formamid deionised
20 µl 10x MOPS
38 µl 37% Formaldehyde

10 µl 50% Bromphenolblue

The RNA loading buffer must be prepared freshly before using

0,1% Toluidinblue color solution

0,1% Toluidinblue

10% EtOH

Heringssperm-DNA Stocksolution

10 mg/ml Heringssperm

Dissolve it in H_2O_{dest}. Incubate the solution in 100°C water bath for 5 min then immediately put it on ice. Aliquot the solution and keep them in -20°C.

Hybridization solution and prehybridization solution

50% Formamid deionised

6x SSC

5x Denhardt's

0,5% SDS

Fill up with DEPC-H_2O. For the prehybridization solution and add 0,1 mg/ml denatured. Add Heringssperm before using.

5x Oligo-Mix

1 Vol solution A

2,5 Vol solution B

1,5 Vol solution C

Aliquot and keep in -20°C.

Solution 0 1,27 g $MgCl_2 \cdot 6H_2O$ Dissolve In 50 ml 1,25 M Tris-HCl pH 8,0.

Solution A 1 ml solution 0

18 µl ß-Mercaptoethanol

15 µl 100 mM dATP

15 µl 100 mM dTTP

15 µl 100 mM dCTP

15 µl 100 mM dGTP

Solution B 2 M HEPES-NaOH pH 6,6

Solution C 90 OD 260-Units random Hexanucleotide (Pharmacia) dissolve in 1 ml 1x TE.

2x Oligo-STOP

7,5 M NH_4Ac

1 mg/ml Heringssperm

0,5 M EDTA

Dissolve in H_2O_{dest}, aliquot and keep in 4°C.

20 x SSC pH 7,0

3 M NaCl

0,3 M $Na_3Citrate_2H_2O$

Dissolve in H_2O_{dest} and adjust the pH value to 7.0 with1 N HCl .

Autoclave the solution.

dYT-Medium

16 g Pepton

10 g Yeast-Extract

5 g NaCl

15 g Agar-Agar (for the solid medium)

Fill up to 1 L with H_2O_{dest} and autoclave.

Water agar plates

3,5-4 g Agar-Agar dissolve in 400 ml H_2O_{dest} and then autoclave.

½ MS-Medium

1x MS-Salt

2,2 g Murashige & Skoog Medium

Dissolve in 500 ml H_2O_{dest}, adjust the pH value to 5,8 with KOH then autoclave the solution.

2x Succrose

20 g Succrose

9 g Agar-Agar

Fill up to 500 ml with H_2O_{dest} and autoclave

Mix 1 Volume of the 1x MS salt with 1 Volume of the 2x Succrose solution after autoclaving them separately, wait until the mixed solution cools down to around 50°C, add appropriate antibiotics and pour the solution into plates.

Carnoys´solution

EtOH : acetic acid; 3:1; v:v

2.1.5 Oligomers

For RT-PCR and qRT-PCR

At4g32280F; cagcgtttgtttgccttgaat

At4g32280R; tcgccgtgatcctcctacc

At2g40610F; tctagaaactggggacaaaactg

At2g40610R; taggcacaatgaaaatacaacca

At5g48900F; ttccttcttccttcttcttctta

At5g48900R; tcgtcgtttggatcactcact

At5g12050F; ttgggagaatacgagaaacagaag

At5g12050R; cggggacgataacattgacc

At3g21330F; cttcttccggtttcacttttcat

At3g21330R; tccgtcgaggtgttgttgctatt

At5gclusterF; caccaaaagggtttcttgc

At5gcluster R; ccgagaagtcacattgatgaa

At5g66590F; ggctttaacccatcacatcatctt

At5g66590R; ctggttcgcgccgtattttc

At1g18140F; tccgccggagaggaatactt

At1g18140R; tagcgaaaccaaaaccctaacaaa

At5g62520F; ggattacgtgagaacccaagtc

At5g62520R; gcggctcgctaaacccatac

At5g57760F; aagatggtgaagtgtaagaaggtt

At5g57760R; taattagactccacatacatcagc

At5g64310F; ggcagccgcattgactccag

At5g64310R; cataaccgccacagatccgaaact

At2g25460F; ccgcgatggtggaggtgaaatg

At2g25460R; gtggctggggaggaaaaagagtcg

At4g13420F; tcccgtgaaccgagtgac

At4g13420R; cgacccttcccgacctaatcttgt

At5g60250F; actcgcggcttctctatcttctca

At5g60250R; gtgggttttgccttttgctttag

At5g01380F; accgacgtaacccttccaacatc

At5g01380R; gccatcttagcagccacgacttc

Actin2 F; ttgttccagccctcgtttgt

Actin2 R; cctggacctgcctcatcatact

At1g53480 F; taatgagaaatgggggagaaaaga

At1g53480 R; ccacaacgagatgaattacaccac

At1g53490 F; ctacccgcgacaccaagaacc

At1g53490R; agctgagggcgggaacgaga

For sequencing the At1g53480 and At1g53490 genomic region

SDF1; cttttcctcaaatactacattt

SDR1; gtcgtttttctctcctgcaaca

SDF2; ctggccatatctcatctttagg

SDR2; tcagaagatgggcaagaggt

SDF3; cctcttgctccatcatctgac

SDR3; aagtggttaagctgaaatacgca

SDF4; ggaaaagaagaaaagctatag

SDR4; agttagttcagtggagaatgg

SDF5; ccattctccactgaactaact

SDR5; ttgctatacctaagctcttt

SDF6; aaagagcttaggtatagcaa

SDR6; ttgctatacctaagctcttt

SDF7; aaacaacatcttaagtaaaat

SDR7; gtcgaggctttattggagag

SDF8; ctctccaataaagcctcgac

SDR8; caaccacaacgagatgaatt

SDF9; caccgtatacaggatgtcct

SDR9; attaatctcaactacaataac

SDF10; tgacttgagtgagcaataa

SDR10; cttttaaaccaccaaaattca

For methylation profiling control

TA2F; aaacgatgcgttgggataggtc

TA2R; atactctccacttcccgtttttctttttta

ta25c11F; actcgctacgagaaagaggat

ta25c11R; accaaaccggaacaataaacc

2.2 Methods

2.2.1 Plant Breeding

2.2.1.1 Plant Breeding on Soil

For the growing of *Arabidopsis thaliana* on earth low nutrient piqueing soil type P or T

(Bayerische Gärtnerei Genossenschaft) is used. It is sieved and mixed with silica sand at a volume ratio of 2:1. The mixture is autoclaved, filled into plastic pots and soaked with tap water. The seeds can then be sown, which should be done in a way to achieve a manageable density of the population.

After a two day vernalisation at 4°C the pots are placed in a Heraeus walk-in growth chamber (80–100 µmol photons/m^2s, 40% relative humidity, 18°C). For the first 3-4 days the seedlings are covered with a translucent plastic bag to prevent the earth from drying out. When the seedlings are a few days old, the bag can be removed. Alternatively seedlings can also be grown on a 1/2 MS plate and then transfered on earth using tweezes.
The plants are watered once per two day with water, which is filled into the dishes which carry the flover pots. Excessive watering should be avoided, because it leads to dwarfism.

2.2.1.2 Plant Breeding Under Sterile Conditions

Surface Sterilization of Seeds

- put seeds into a sterile Eppendorf tube
- wash seeds once with 500 µL H_2O_{dest}
- take off the water and add 500 µL 70% EtOH - wash for 2-3 min
- wash 3x with 500 µL H_2O_{dest}
- bleach (5% sodium hypochlorite 0,01% triton X 100) for 20 min under constant inversion of the tube
- wash 4x with H_2O_{dest} and at the fourth time do not remove the water

After surface sterilization the seeds are pipetted onto the medium using the remaining H_2O_{dest} (under the clean bench). Take care to separate the seeds as good as possible.

The 1/2 MS plates are sealed with Nescofilm and vernalised at 4°C before they are transferred into the plant chamber.

2.2.2 General Molecular Biology Methods

2.2.2.1 Ethanol-Precipitation of Nucleic Acids

- add 0.1 volume of 3 M NaAc pH 5.2 and 2 volume of 100% ethanol to the DNA solution, mix well and precipitate the DNA at –70°C for 30 min or at –20°C over night
- centrifuge at 14000 rpm and 4°C for at least 20 min

- discard the supernatant and wash the pellet with 70% ethanol
- centrifuge at 14000 rpm and RT for 10 min
- discard the supernatant and dry the pellet for about 5 min
- dissolve the pellet in the desired volume 1xTE pH 8.0 or H_2O_{dest}

The TE buffer contains EDTA, which forms a complex with bivalent ions. Most DNases and other enzymes like polymerases need bivalent ions as a cofactor and are therefore inhibited by EDTA. For this reason 1xTE is especially useful if the DNA is supposed to be stored for a long time. In turn 1xTE should not be used if the DNA is needed for a PCR or other processes that involve enzymes which could be inhibited by the EDTA. In this case store nucleotides etc. in T0.1E.

2.2.2.2 Isopropanol-Precipitation of Nucleic Acids

- add 0.1 volume of 3M NaAc (pH 5.2)
- add 0.6 - 1 volume(s) of isopropanol and mix well
- incubate for 15-30 min at RT
- centrifuge at 14000 rpm and RT for 10 min
- discard the supernatant and wash the pellet with 70% EtOH
- centrifuge at 14000 rpm and RT for 10 min
- pipette away the supernatant and dry the pellet for about 5 min
- dissolve the pellet in the desired volume 1xTE pH 8.0 or H_2O_{dest}

2.2.2.3 Purify the nucleic acids with Phenol/Chloroform (PC)-Extraction

- add 1 volume of phenol/chloroform (1:1) and shake well or vortex
- centrifuge at 14000 rpm and RT for 5 min
- transfer the upper, aqueous phase into a new Eppendorf tube – the white interphase must not be carried along
- add 1 volume of CIA (to get rid of phenol) and transfer the upper layer to a new tube to remove the remaining chloroform. An ethanol or isopropanol precipitation (see 2.2.2.1 and 2.2.2.2 respectively) is performed subsequently.

2.2.2.4 Restriction Digest of DNA

Restriction endonucleases are enzymes that cut DNA at specific sequences. These so called restriction sites are palindromic in many cases, and four to eight base pairs long. The DNA is cut via a hydrolysis of the phosphodiester-bonds in both strands.

For a restriction digestion the following components are necessary:

DNA-solution	x µL (10 µg)
10x BSA (in some cases)	$^1/_{10}$ volume of the total preparation
10x restriction buffer	$^1/_{10}$ volume of the total preparation
restriciton enzyme	2-5 U per µg DNA
H_2O_{dd}	to fill up to the final volume (analytically: 20µL; preparative: 100-200µL)

Salt concentration of the buffer and presence of BSA are dependent on the enzyme used. The appropriate buffer conditions are indicated by the corresponding company. In most cases the same company provides suited buffers.
The incubation time should be at least 60 min. For larger preparations, an incubation over night might be necessary. The incubation temperature is 37°C if not indicated otherwise by the manufacturer.
A double digest with two different enzymes can be performed if both enzymes have 100% activity in the same buffer. Otherwise a precipitation step needs to be carried out before the second digest.

2.2.2.5 RNaseA-Digest

Add 10-40 µg RNaseA (stock solution 10mg/mL) per mL DNA-solution and incubate at 37°C for at least 30 min. If necessary perform a phenol/chloroform extraction afterwards.

2.2.2.6 Agarose gel electrophoresis and isolation of DNA fragments

DNA was subjected to electrophoresis using 1.0 to 2.0% agarose gels containing ethidium bromide at final concentration of 0.5 µg/ml. Gels were cast and run in TAE buffer, 1Kb Plus DNA Ladder (Invitrogen) was routinely used as size marker.
Gel pieces containing desired DNA fragment were excised from the agarose gel. DNA was then isolated using GFX™ DNA and Gel Band Purification Kit (Amersham), according to the manufacturer's instructions.

2.2.2.7 Preparation of E. coli XL1 Blue competent cells

100ml dYT medium was inoculated with 5 ml overnight culture (containing tetracycline15 mg/l) of E.coli XL1 Blue. The cells were grown to an OD600 < 0.5, and then were pelletedby centrifugation for 20 min in 4°C with 4,500 rpm. The pellet was suspended and incubated in cold TBF I buffer (30 mM KAc, 50 mMMnCl, 100

mMRbCl, 10 mM CaCl2, 15% glycerol, pH 5.0 with acetic acid, sterilized byfiltration) for20 min on ice. Next, the cells were again pelleted by centrifugation in 4℃ for 5 min with 4,000 rpm. The pellet was resuspended in 3.6 ml of cold TBF II buffer (10 mM NaMOPS pH 7.0, 10 mM RbCl, 15 mM $CaCl_2$, 15% glycerol, sterilized by filtration), aliquoted, snap frozen in liquid nitrogen. Aliquots were stored in -70℃.

2.2.2.8 E. coli transformation (heat-shock)

An aliquot of competent E. coli cells (50 µl) was thawed on ice. 100-500 ng of plasmid DNA was added, mixed and incubated for 20 min on ice. An eppendorf containing transformation mix was then heated to 42℃ for 1 min and then immediately cooled on ice for 2 min. 400 ml of dYT medium was added and probe was incubated in 37℃ for an hour.

50-200 ml of transformed bacteria suspensions were plated on dYT plates with addition of the appropriate antibiotic.

2.2.2.9 E. coli growth conditions

Standard bacteria cultures were grown in volume of 3 ml (dYT or LB) at 37℃ over night in a shaker incubator (120 rpm). Cultures were inoculated with a single colony from the plates. Cultures 500 ml or 1000 ml were inoculated from the 3 ml cultures.

2.2.2.10 Preparation of plant genomic DNA

- Place 2 small metal balls into a 2 ml eppendof tube.
- Place roughly 0,1g leaf into the eppendof tube and put it into liquid nitrogen immediately.
- The leaf material then is ground with the ball mill (Retsch) at the highest speed for 30 seconds.
- After homogenizing add 600 µl 2x CTAB buffer into the eppendof tube. Vortex it shortly and incubate it in 65℃ water bath for at least 30 min.
- Add 750 µl Chloroform/Isoamylalcohol (24:1) and vortex for 15 sec.
- Centrifuge for 5 min at highest speed and move the upper phase into a new eppendof tube.
- Perform a isopropanol precipitation (see 2.2.2.2)
- Then perform a RNaseA digestion (see 2.2.2.5)
- For the DNA used for methylation profiling, a purification step should be performed with the NucleoSpin kit (MN), according the manufacture's description.

2.2.2.11 Northern Blot

Agarose/Formaldehyde Gel Electrophoresis

Prepare gel: Dissolve 0.6 g agarose in 35 ml water and cool to 60 °C in a water bath. When the flask has cooled to 60 °C, place in a fume hood and add 5 ml of 10xMOPS running buffer and 9.5 ml formaldehyde. Pour the gel and allow it to set. Remove the comb, place the gel in the gel tank, and add sufficient 1xMOPS running buffer to cover to a depth of ~ 1mm.

Prepare sample: Adjust the volume of each RNA sample to 1-2 µl with water, then add 8 µl freshly prepared RNA loading buffer (see solution for northern blots). Mix by vortexing, micro centrifuge briefly to collect liquid, and incubate 15 min at 62-65 °C. Then cool on ice immediately.

Run gel: Run the gel in 1xMOPS running buffer at 65 volt for about 2 hours.

Process for the marker : Cut the RNA marker part off the gel, stain it with 0,1% Toluidinblue color solution for 10 min in room temperature. Wash it with 0,1% EtOH for 3X 20 min. Leave it in the washing solution at the last time until the marker bands are clearly visible. Make photos of it, and measure the position of each band.

Transfer of RNA from Gel to Membrane

Prepare gel for transfer: Place the gel in an RNase-free dish and rinse with sufficient 20XSSC to cover the gel for 2x20 min.

Transfer RNA from gel to membrane:

1. Fill the glass dish with enough 20xSSC. Put a glass plate on the top of the glass dish.
2. Cut one long piece of Whatman 3MM paper, place its middle part on the glass plate and wet both edges of the paper with 20xSSC in the dish.
3. Place the gel on the filter paper and squeeze out the air bubbles by rolling a glass pipette.
4. Cut four strips of plastic wrap and place over the edges of the gel.
5. Cut a piece of nylon membrane (Pall biosupport menbranes) just large enough to cover the gel and wetted in 20XSSC. Place the wetted membrane on the surface of the gel. Avoid getting the air bubbles under the membrane.

6. Flood the surface of the membrane with 20xSSC. Cut 5 sheets of whatman 3MM paper to the same size as membrane and place on top of the membrane.
7. Put paper towels on top of the whatman 3MM paper to a height of ~6 cm, and add a weight to hold everything in place.
8. Leave overnight.

Prepare membrane for hybridization: Remove paper towels and filter papers and recover the membrane and flattened gel. Mark with a pencil the position of the wells on the membrane and ensure that the up-down and back-front orientation are recognizable. Rinse the membrane in 5xSSC, then place it on a sheet of Whatman 3MM paper and dry it in 80 °C oven for 1 hour and leave it at room temperature.

Hybridization Analysis

Prepare DNA or RNA probe (>10^8dpm/µg):

The probe labeled with α-32P dCTP radioactivity in Klenow-reaction

1. Dilute 50-100 ng DNA in 15 µl H_2Odd and denature it in 100°C water bath for 5 min.
2. Centrifuge briefly to bring the contents to the bottom of the tube, and put on ice.
3. Add 5 µl 5x Oligo mix and1 µl Klenow-Enzym (5 U/µl) into the denatured DNA. Then, (in the isotope lab) add 5 µl α32P-dCTP (3000 Ci/mmol, 10µCi/µl). Mix carefully and leave the tube in room temperature at least for 1 hour.
4. Add 1 Volume of 2x Oligo-stop-buffer and 2 Volume of 100% EtOH. Centrifuge at least for 15 min at 14000 rpm and 18°C.
5. Dissolve the pellet in 100 µl H_2Odd.
6. Denature the DNA in 100°C water bath for 5 min. Keep the probe on ice.

Hybridization:

9. Pre-hybridization: Wet the membrane in the 5xSSC and place it RNA-side-up in a hybridization chamber and add 50 ml pre-hybridization solution, then place the chamber in the 42 °C water bath and incubate with rotation for more than 1 hour.
10. Hybridization: Double-stranded probe was denatured by heating in a water bath for 5 min at 100 °C, then transfer to ice. Pipette the desired volume of

probe into the hybridization chamber and continue to incubate with rotation overnight at 42 °C.

Autoradiography:

11. The membrane is washed once for 20 min with 2x SSC/0,5% SDS, and once for 20 min with 2x SSC/0,1% SDS (at room temperature).
12. Remove final wash solution and cover the membrane with transparent plastic wrap. Do not allow membrane to dry out if it is to be reprobed.
13. Expose the membrane to phosphor screen for at least 4 hours.
14. The phosphor screen is scanned in the STORM®860 Phosphoimager (Amersham Pharmacia Biotech).

2.2.2.12 Total RNA isolation

For the RNA used in the northern blot, the isolation procedure is performed of Logemann et al. (1987 See below). Then, the mRNA is isolated from the total RNA with the PolyATract® mRNA Isolation-Kit from Promega. For the RNA used for RT-PCR, qRT-PCR, the SV Total RNA isolation System from Promeg is used. For the RNA for the microarray experiments, the RNeasy Plant Mini Kit from Qiagen is used.

Total RNA isolation protocol according to Logemann et al. (1987)

• Homogenize 1-2 g frozen plant leaves in liquid nitrogen until the leaves become fine powder.
• add 10 ml R1-buffer (including 42,5 µl ß-Mercaptoethanol) .
• Centrifuge the homogenized solution for 10 min at 14000 rpm and at 4℃ (JA20, Beckmann).
• Transfer the upper phase into a 50 ml Falcon.
• Add 10 ml Phenol/Chloroform and mix well.
• Centrifuge for 10 min at 4500 rpm and at 4 ℃ (SX4250, Beckmann CoulterTM)
• Transfer the water phase (upper layer) into a new 50 ml Falcon.
• add 10ml Chloroform, mix well and centrifuge for 10 min at 4500 rpm and transfer the upper layer to a centrifuge tube.
• add 1/10 Vol 3 M NaAc pH 5,2 and 2 Vol 100% EtOH.
• leave it in -20℃ for 2 h or overnight.
• Centrifuge the tube for 30 min at 14000 rpm (JA20, Beckmann) and at 4℃.

•Wash the pellet two times with 20 ml 3 M NaAc pH 5,2 and each time centrifuge for 10 min at 14000 rpm (JA20, Beckmann) and at 4°C.

• Wash the pellet with 20 ml 70% EtOH

• Centrifuge for 10 min at 14000 rpm (JA20, Beckmann) and at 4°C.

• Discard the supernatant and leave the pellet until it is dried. Then, dissolve the pellet in 500 µl H_2O_{dd} .

2.2.2.13 Sequencing

The DNA fragment or clone sequencing was done by the company eurofin MWG-Biotech.

2.2.2.14 Quantify the total RNA and DNA

The total RNA quantity was measured by using the NanodropND-1000 UV-Vis Spectrophotometer (Thermo).

- Start the NanoDrop software.
- Pipette 1.5µl of H_2O_{dest} onto the Nanodrop for initialization the machine.
- Choose the nucleotide measurement tab.
- Clean the Nanodrop and pipette 1.5µl H_2O_{dest} for blanking.
- Select RNA-40 as the sample type.
- Clean the Nanodrop and pipette 1.5µl RNA sample onto it and click measure tab.
- Record the concentration, the value of the absorbance ratio (260nm/280nm and 260nm/230nm).

The quality of the RNA used for microarray and RT-PCR was tested with the bioanalyzer (Agilent) and the Agilent RNA 6000 Nano kit according to the manufacture's instruction.

- Prepare the gel by filtering 550µl RNA6000 Nano gel matrix in a spin filter with centrifuging at 1500g for 10 min at room temperature. Aliquot 65µl filtered gel into 0.5ml microfuge tubes. The filtered gel can be kept in 4°C for one month.
- Add 1µl RNA6000 Nano dye concentrated into the 65µl filtered gel.
- Vortex the solution well. Spin the tube at 13000g for 10 min at room temperature, use it within one day.
- Put a new RNA 6000 Nano chip on the chip priming station.
- Pipette 9µl of gel-dye mix in the well marked "**G**".

- Make sure that the plunger is positioned at 1ml and then close the chip priming station.
- Press plunger until it is held by the clip.
- Wait for exactly 30 seconds then release clip.
- Wait for 5 sec, slowly pull back plunger to 1ml position.
- Open chip priming station and pipette 9µl of gel-dye mix in the wells marked G.
- Pipette 5µl of RNA 6000 Nano marker in all 12 sample wells and in the ladder well.
- Pipette 1µl of ladder in the well and 1µl of sample in each of the 12 sample wells.
- Vortex the chip for 1 min at 2400rpm.
- Run the chip in the Agilent 2100 bioanalyzer within 5 min.

The result provides a visual impression of the quality of the RNA. It also gives you a "RIN" (RNA integrity number) value, which quantifies the quality of the RNA. For all RNA used for microarray in this work the RIN numbers were higher than 8, most cases were higher than 9.

For the quantification of DNA, U-1100 spectro photometer (HITACHI) was used.

2.2.2.15 Methylation profiling

McrBC is an endonuclease which cleaves DNA containing methylcytosine* on one or both strands. McrBC will not act upon unmethylated DNA. Sites on the DNA recognized by McrBC consist of two half-sites of the form $(G/A)^{m}C$. These half-sites can be separated by up to 3 kb, but the optimal separation is 55-103 base pairs.

McrBC PCR methylation profiling was performed on genomic DNA that was extracted from 14-day-old rosette leaves from a pool of adozen plants grown under identical conditions as described above.This DNA was treated with McrBC in the same manner. For McrBC PCR of Col, L*er* diploid tetraploid plants as well as the Col L*er* triploid hybrids and tetraploid hybrids, DNA was isolated from a pool of rosette leaves of a dozen plants of identical age as above, and 1µg of DNA was digested with 1 U of McrBC for 12 h. 5µl of DNA from digested and mock-digested DNA were used as a template in a 20-µl PCR reaction with 30 cycles of amplification for each primer pair. Failure to amplify a product after digestion by McrBC indicates that the gene is methylated. Control primers from a methylated transposon (TA2, see material) control the digestion step.

2.2.3 Induction of *Arabidopsis* plants into polyploid

A drop of colchicine solution (<15µl) was placed on the apex of young seedlings with less than five primary leaves (one-drop method). For "sensitive" ecotypes like L*e*r the concentration of the colchicine solution was reduced to 0.1% or 0.05%. For a "resistant" ecotype like Col, the colchicine concentration can be 0,5% or 0,1%. The evaluation of cellular size, in particular trichomes, as explained in the results, enabled identifing polyploid versus non-polyploid sectors on a treated plant. This allowed the harvest plants producing polyploid seeds selectively, which in turn alleviated the identification of the desired tetraploids in the second round of analysis. In the second generation, chromosome staining and flowcytometry methods were used to finally confirm the polyploidy level of the plants.

2.2.4 Chromosome staining

For the metaphase chromosome preparation standard protocols have been followed (e. g.Maluszynska and Heslop-Harrison, 1991; Zhong et al., 1996) with slight modifications.

- Dissect the roots from seedlings grown on water agar plates and incubate for 1 h at 4ºC in 2 mM Hydroxychinolin until the tissue is transferred to Carnoys´solution. Alternatively, dissect roots and immediately submerge them in 1ml Carnoys´ solution to fix the root tips for at least 12hrs.
- Wash the root three times for ca. 5 min with H_2O_{dest} and incubate for at most 15 min at 37ºC in enzyme solution (10% Macerozyme R-10 from Duchefa; 2% Cellulase "Onozuka R-10" from Duchefa in 100mM Citric-buffer pH4.8).
- Wash the roots three times with H_2O_{dest} and place them on small Petri dishes to separate the root tips (they often separate automatically upon digestion).
- Transfer about 10 root tips onto a slide (SuperFrost®Plus, Menzel), squash with the tip of a needle and coverwith a drop of 60% acetic acid for ca. 1min.
- After adding 800µl Carnoy´s solution for 2 min, submerge the slide in 70% Ethanol, air-dry it and add 60µl DAPI solution (1µg/ml).
- Cover the slide with a cover slip and store it for 10 min in the dark.
- Wash off the cover slip and DAPI with H_2O_{dest}, air-dry, cover it with 10% glycerol and a new cover slip.
- Analyze the sample under the epifluorescence microscope. For each line and generation between 5 and 20 metaphases were analyzed. In tetraploid metraphase figures, the absence of 1 or 2 chromosomes was tolerated since the large member of chromosomes often covers single ones.

2.2.5 Flow cytometry analysis

Flow cytometry was essentially performed as described (e.g. Henry et al., 2005) using the high-resolution kit from Partec. Briefly, leaves were chopped with a sharp razor blade in 0.3ml of nuclei extraction buffer (solution A of the Partec kit) and filtered through 20µm or 30µm Cell Trics filters (Partec). The flow through was combined with 1.2ml of DAPI solution (solution B of the Partec kit) and analysed in a PAS II flow cytometer (Partec) equipped with a HBO lamp for UV excitation. The PAS II FlowCytometer distributes the measured particles according to their fluorescence intensity into 1024 different channels. Routinely, several thousand particles were measured (from channel V, at the gain value 350, L-L intensity 6, flow speed at 2) per leaf and measurements were often repeated. The peak positions of the 2C, 4C, 8C, 16C and 32C nuclei were compared between diploid and tetraploid plantswhose ploidy was already known or had been assessed by chromosome counts ofmetaphases. Due to endopolyploidy, flow cytometry measurements from a diploid *Arabidopsis* plant not only exhibit a 2C peak but also further peaks up to 32C. Consequently, plants with a higher basic ploidy level e.g. tetraploids lack the 2C peak. However, flow cytometry is not sensitive enough to discriminate between certain euploids and aneuploids, e.g. 4C versus 4C+1. Bearing this in mind, the chromosome number of plants was assessed by counting metaphase chromosomes of root tips.

2.2.6 Microarray

Fig. 2 shows the microarray experiment procedure. Schematrically total RNA was isolated and purified by conventional methods using QIAGEN columns. The quality of RNA used for microarray was controlled with the Bioanalyzer (Agilent). Only RNA probes without detectable degradation and impurities were further processed (as indicated by the RIN value). The RNA was then processed with the steps shown in Fig. 2.

cDNA synthesis (Agilent Low RNA Input Linear Amplification Kit)

During this step, the total RNA was reverse transcribed into double stranded cDNA

- Prepare spike in (Agilent RNA Spike-In Kit) dilutions:
  For the total RNA 500 ng, dilute spike in A or B (Always label Spike A Mix with cyanine 3 and Spike B Mix with cyanine 5) through three serial dilution steps: 1:20, 1:40, 1:4
  The first dilution can be kept at -70 ºC for up to 1 month. The second and third dilutions must be discarded after used.

- Mix the following components:Total RNA 8,3µl (500ng) diluted spike A (or B) 2µl, T7 promoter primer 1,2µl
- Incubate the mixture in 65 ºC for 10 minutes then place the reactions on ice for 5 min.
- Mix the following components as a master mix, mix gently by pipetting:

Component	volume (µl) per reaction
5X First Strand	4
0.1 M DTT	2
10 mM dNTP mix	1
MMLV-RT	1
RNaseOut	0.5

- Add 8.5µl of the master mix to each RNA reaction tube, mix gently, then transfer the tube into 40ºC thermo block and incubate for 2 hours.
- Incubate at 65ºC for 15 min to denature the MMLV-RT enzyme.
- Spin the reaction tube briefly and keep them on ice.

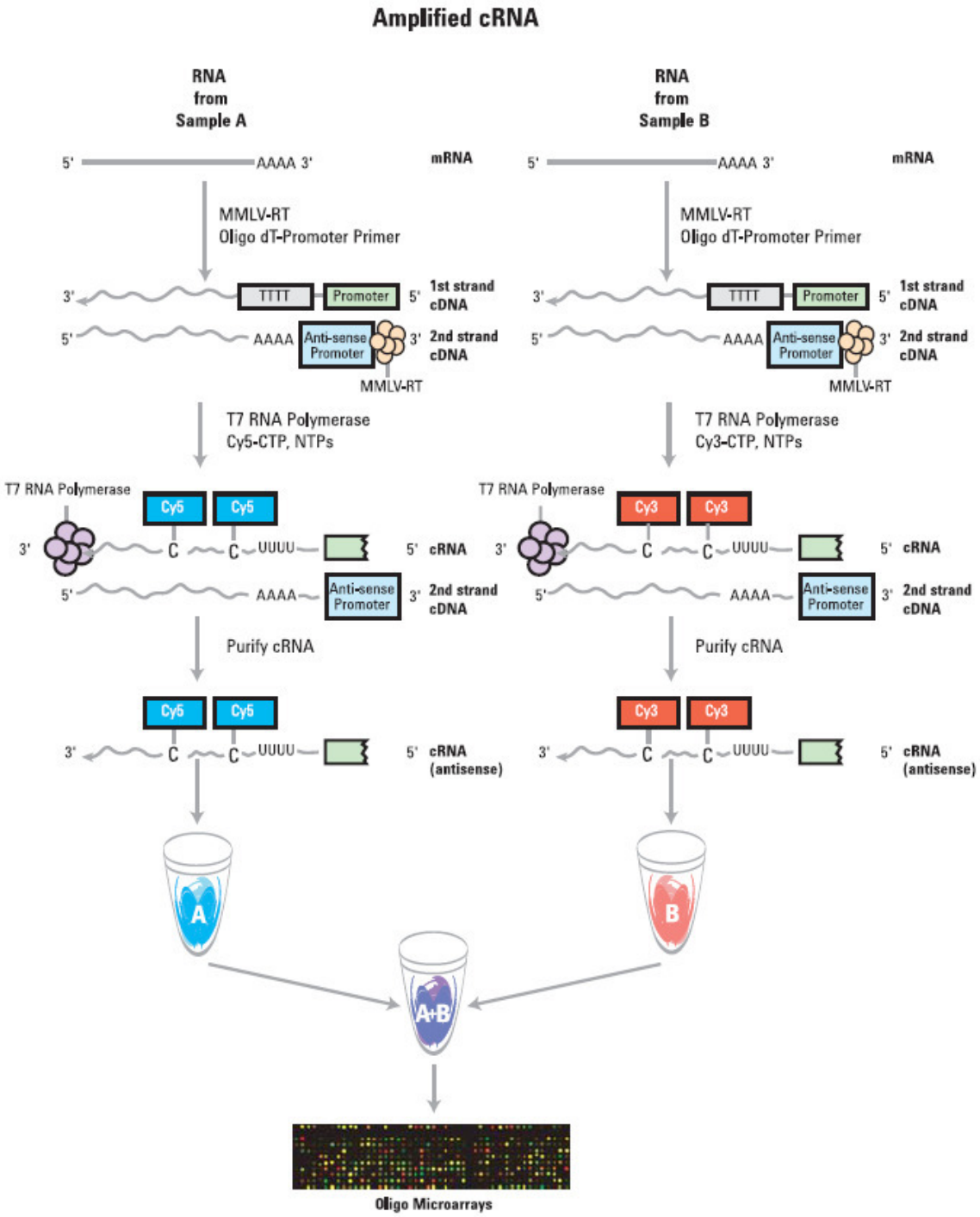


Fig. 2: Micorarray experiments procedure (Agilent)

cRNA synthesis (Agilent Low RNA Input Linear Amplification Kit)

During this step the cDNA was transcripted to cRNA, the cRNA was simultaneously labeled with Cy3 or Cy5 and amplified as well.

- Prewarm the 50% PEG solution at 40ºC for 1 min, keep at room temperature until needed.
- Mix the following components gently as master mix

Components	volume (µl) per reaction
Nuclease-free water	15.3
4X Transcription Buffer	20
0.1M DTT	6
NTP mix	8
50% PEG	6.4
RNaseOUT	0.5
Inorganic pyrophosphatase	0.6
T7 RNA Polymerase	0.8
Cyanine 3-CTP or cyanine 5-CTP	2.4

- Add 60 µl of the master mix to each reaction tube from cDNA synthesis, mix gently by pipetting.
- Incubate samples at 40ºC for 2 hours.

Purify the labeled amplified cRNA

In this step, the Qiagen RNeasy mini spin columns were used according to the supplier's description of the RNeasy Plant Mini Kit. The cRNA was purified and diluted in 30µl H_2O_{dest}.

Quantify the cRNA

The purified cRNA was quantified by using NanoDrop ND-1000 UV-VIS Spectrophotometer (Peqlab). The cRNA concentration, RNA absorbance ration (260nm/280nm), and Cy3 or Cy5 dye concentration (pmol/µl) were determined. The cRNA yield and the specific activity were calculated in the following way:
(*Concentration of cRNA*) * 30 µL (*elution volume*) / 1000 = µg of cRNA
(*Concentration of Cy3 or Cy5*) / (*Concentration of cRNA*) * 1000 = pmol Cy3 per µg
The experiment can be proceeded if the yield is more than 825 ng, and the specific activity is more than 8 pmol Cy3 or Cy5 per µg cRNA.

Hybridization(Gene Expression Hybridization Kit)

- Mix the following components for each microarray

Components	volume per array
cyanine 3-labeled, linearly amplified cRNA	825 ng
cyanine 5-labeled, linearly amplified cRNA	825 ng
10X Blocking Agent	11 µL

Nuclease-free water bring volume to 52.8 µL

25X Fragmentation Buffer 2.2 µL

- Incubate the reaction at 60 ºC for exactly 30 min to fragment RNA.
- Add 55µl 2x Hybridization buffer into each reaction tube to stop the fragmentation reaction.
- Mix well by careful pipetting. Take care to avoid introducing bubbles.
- Load a clean gasket slide into the Agilent SureHyb chamber base with the label facing up and aligned with the rectangular section of the chamber base.
- Spin the hybridization sample briefly, loading 100µl sample onto each gasket well. Slowly dispense the volume in a drag and dispense manner, being sure not to touch the gasket well.
- Place the "active side" of the microarray (Agilent, 4X44K) down onto the gasket slide. Verify that the sandwich-pair is properly aligned.
- Place the SureHyb chamber cover onto the sandwiched slides and slide the clamp assembly onto both pieces.
- Hand –tighten the clamp onto the chamber.
- Vertically rotate the assembled chamber to wet the gasket and assess the mobility of the bubbles.
- Place the assembled slide chamber in the rotisserie of the hybridization oven set to 65ºC, and rotate at 10 rpm.
- Hybridize at 65ºC for 17 hours.

Washing

- Prewarm the gene expression buffer 2 (Agilent) at 37 ºC overnight.
- Wash the dishes once with acetonitrile and 2 times with Milli-Q water.
- Disassemble the slide sandwich pair in Buffer1 (Agilent). Place the microarray slide in a slide rack and put the rack into a dish which has been filled with buffer1 and with a stir bar inside. Place the dish onto a magnetic stir plate and stir the buffer 1 for 1 min.
- Wash the slide in the warmed buffer 2 for 1 min in the same way as described in the last step.
- Wash the slide in acetonitile for 1 min in the way described before.
- Wash the slide in stabilization buffer (Agilent) for 1 min.
- Leave the slide in the rack in dark until it is dried.

Scan the slides

Scanner (Agilent) setting:	
Scan region	61x 21.6mm
Scan resolution	5 µm
Dye channel	Red & Green
Green PMT	XDR Hi 100%
	XDR Lo 10%
Red PMT	XDR Hi 100%
	XDR Lo 10%

2.2.7 RT-PCR and qRT-PCR

The standardized set of rules and steps are followed in RT- and qRT-PCR analysis respectively, which essentially follow the recommendations of Czechowski et al. (2005). Exceptions are for instance the selection of *ACT2* as reference for the sake of comparison with *Arabidopsis* allopolyploid analysis (Wang et al., 2006b). Where applicable the same standards were followed for RNA isolation used in microarray analysis. These included: the analysis of at least three biological replicates, isolation of RNA with the same tools (Promega kits), storage at -70ºC, RNA integrity analysis with the Bioanalyser 2100 (Agilent), DNAse treatment of RNA, primer design with the same set of selection parameters, design of primer pairs, which covered exon-exon junctions (care of known splice variants was taken if necessary), assessment of the positive control gene (*ACT2*; reference gene in qRT-PCR) in all isolated RNA batches, negative control reactions (water and buffer) for all tests, parallel set-up of control and test reactions and melting curve analysis to exclude the production of multiple amplicons.

RT-PCR was performed with the TaqMan®–kit (Roche) essentially as described (Treml et al., 2005). Briefly, the following RT-PCR mix was generated: 2µg/1µl total RNA, 1,0µl 10X Taq-Man RT buffer, 2,2µl 25 mM Mg_2Cl, 2,0µl deoxyNTPs mixture, 0,5µl oligo dT-primers, 0,2µl RNAse-inhibitor, 0,25µl MultiScribe RT (50u/µl) and H_2O was added to give a total volume of 10µl. The mix was incubated for 20 min at 25ºC and then for 45 min at 48ºC, followed by 5 min at 95ºC and10 min 4ºC. If necessary, the mix was stored at -20/-70ºC. For conventional RT-PCR 2 µl of the first strand cDNA was mixed with 2 µl 10X PCR-buffer, 2 µl dNTPs (10mM), 2 µl forward primer (10mM), 2 µl reverse primer (10mM), 0.4µl/1u TaqPolymerase and 9.6 µl H_2O to give a final volume of 20 µl. Between 30-40 cycles were performed (each 94˚C for 15 sec, 60˚C for 30sec, 72˚C for 1min) followed by 72˚C for 5min final extension. Estimation of frequency of (ecotype) specific transcript variants were performed by sequencing RT-PCR amplified fragments and comparison of the sequence results with sequence

results of calibrated mixtures of L*er*-0 and Col-0 and/or by determination of ecotype assignment of clones of RT-PCR material.

For qRT-PCR appropriate fragments of genes, i.e. assessed for non-overlapping regions with other homologous genes, were cloned using the pGEM-cloning system (Promega). QRT-PCR was performed in the Light-Cycler (Roche) using the SYBR Green dye method as indicated by the supplier (Roche). The cDNA from the reverstranscription reaction (as described before) was diluted as 1:4. Then 2.5µl of the dilution was used in the qRT-PCR. First pipette the reaction master mix, including 0.5µl forward primer (10mM), 0.5µl reverse primer (10mM), 1.5µl H_2O (from the kit), 5µl enzyme master mix (from the kit) for each reaction, into each well in the plate or into each reaction capillary, then add 2.5µl diluted cDNA or different diluted *ACT2* plasmids and mix well. Centrifuge the plate or the capillaries at 2000 rpm for 1 min to get everything in the bottom and put the plate or the capillaries into the Light cycler (Roche). The corresponding primers are given in the material. 10 times, 100 times, 1000 times diluted *ACT2* plasmids were used for the calibration reference curve in the same running. The cDNA of *ACT2* was also used as a control.

2.2.8 Amino acid extraction and GC MS measurement

The amino acids were extracted from 20 day old plant leaves with EZfaast kit (Phenomenex).

- Freeze the leaf material in liquid nitrogen and homogenize it until it becomes powder. Measure 50mg frozen powder and add 200µl H_2O_{dest} inside.
- Vortex it and try to keep all the leaf powder in the solution. Then, incubate the solution in the ultrasonic bath (250 kHz) for 10 min.
- Vortex again, and centrifuge for 10 min at 134000rpm and at 4˚C.
- Transfer 100µl supernatant into the glass vial and add 100µl reagent 1. Vortex briefly.
- Attach a sorbent tip to a 1.5ml syringe; pass the solution in the sample preparation vial through the sorbent tip by slowly pulling back the syringe piston.
- Pipette 200µl reagent 2 (washing solution) into the sample preparation vial and slowly pass the solution through the sorbent tip and into the syringe barrel.
- Detach the sorbent tip and discard the liquid in syringe.
- Pipette 200µl Eluting Medium into the sample.

- Pull back the piston of a 0,6ml syringe halfway up the barrel and attach the sorbent tip.
- Wet the sorbent with eluting medium, stop when the liquid reaches the filter plug in the sorbent tip.
- Eject the liquid and sorbent out of the tip, repeat until all the sorbent particles in the tip are expelled into the sample preparation vial.
- Using Dialamatic Microdispenser transfer 50µl reagent 4, vortex for 5 sec, stand for 1 min, then vortex again and stand for 1 min again.
- Add 100µl reagent 5 and vortex briefly and stand for 1 min.
- Transfer the upper organic layer into the autosampler vial.
- Dry the sample with nitrogen flow and dissolve the sample in 100µl reagent 6. Transfer the sample into an insert and place the insert in the same autosampler vial.

The samples then were analyzed with GC MS by Quirin Sinz in Prof. Dr. Wilfried Schwab' s lab.

3. Results

3.1 Generation of *Arabidopsis* tetraploid lines

The generation of polyploids involved three major steps: induction with colchicine, identification of candidate lines (sectors) and assessment of polyploid lines by analyzing the progeny of candidate lines. The rationale of the procedure was based on the straight forward verification of the architecture of trichome cells as a morphological surface "marker", which discriminated between diploid and polyploid *Arabidopsis* sectors and lines respectively (Fig. 3). Polyploid sectors had more multibranch trichomes than diploid sectors, a phenomenon, which has also been shown in the presented work (Perazza et al., 1999). The treated parental generations were examined according to this. Plants with polyploidy sectors were always counted as potential parents for polyploid lines, because as observed, if a polyploid sector developed reproductive organs, it regularly produced polyploid progeny.

3.1.1 Induction via Colchicine treatment

The initial tests were started by Prof.Dr.Ramon Torres Ruiz and Kristina Haage using high concentrations of colchicine and either extreme durations of exposition to this agent or high numbers of treatments. In these cases, most seedlings became necrotic and died. The tolerance to colchicine varied between ecotypes. With this extreme treatment, no other than Col-0 plants with polyploid sectors could be obtained. Diverse ecotypes were also treated with the "one-drop method" (see Mat. and Meth.) and concentrations of 0.5% and 0.1% colchicines respectively. Variable results were obtained. It turned out that the one-drop treatment was sufficient for polyploidy induction, however the concentration of the used colchicine was ecotype dependent. This method was continued to be used for Col-0 and L*er*-0 ecotypes to generate stable tetraploid lines for further transcriptome analysis. Col-0 ecotype was quite resistant to colchincine, most induced plants survived but L*er*-0 ecotype was quite sensitive and only a few plants survived after low concentration (0.1%) colchicine treatment. 30 Col-0 seedlings were induced with 0,5% colchicine. 25 lines survived and among these 25 lines, 10 lines had polyploidy sectors. 200 L*er*-0 seedlings were induced with 0,1% colchichine . Only 53 lines survived and among these 53 lines, 11 lines had polyploid sectors. Only the latter were followed to the next generation for further assessment (Fig. 3).

In addition, seven other ecotypes were also induced with the one-drop method with 0,1% colchicine, these were Bor-1, Bur-0, Ct-1, L*er*-1, Nd-1, Pro-0 and Ts-1.

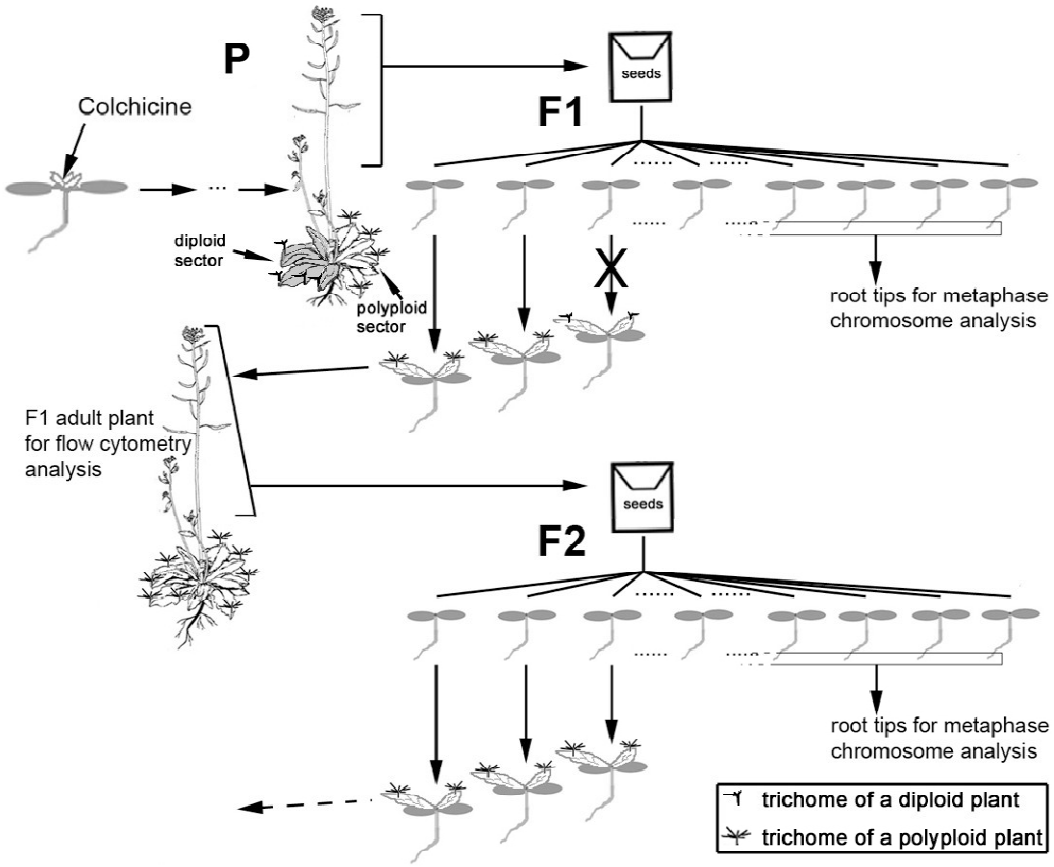


Fig. 3: Selection and assessment scheme for polypoid lines in consecutive generations (from Yu et al., 2009).
After the induction with colchicines polyploid sectors could be selected according to the trichome morphology. The selected plants gave the F1 seeds and the plants growing from the F1 seeds were used for metaphase chromosome analysis, trichome morphology assessment and flow cytometry assessment. Those tetraploid plants confirmed by all the methods were followed to the next generation.

3.1.2 Identification of polyploidy candidate lines

After the colchicine treatment, trichome branch number was used as a marker for polyploid candidate selection as described in Fig. 3 (and Mat. And Meth.). The trichome structure was ecotype and polyploidy level dependent. Different ecotype plants had different distribution of multi branch trichomes. This was verified by counting trichome branches in the diploid and the polyploid lines generated. For example, Zürich diploid plants had more one or two branch trichomes and less three or four branch trichomes than Col-0 diploid plants and L*er*-0 diploid plants. Within the same ecotype, the higher the ploidy level of the plant was the bigger portion of trichomes with more branches the plant had. For example L*er*-0 tetraploid plants had much more four branch trichomes than L*er*-0 diploid plants (Fig. 4). Therefore, the

induced plants could be characterized by comparing the trichome morphology of different ploidies within the same and between different ecotypes (Fig. 4).

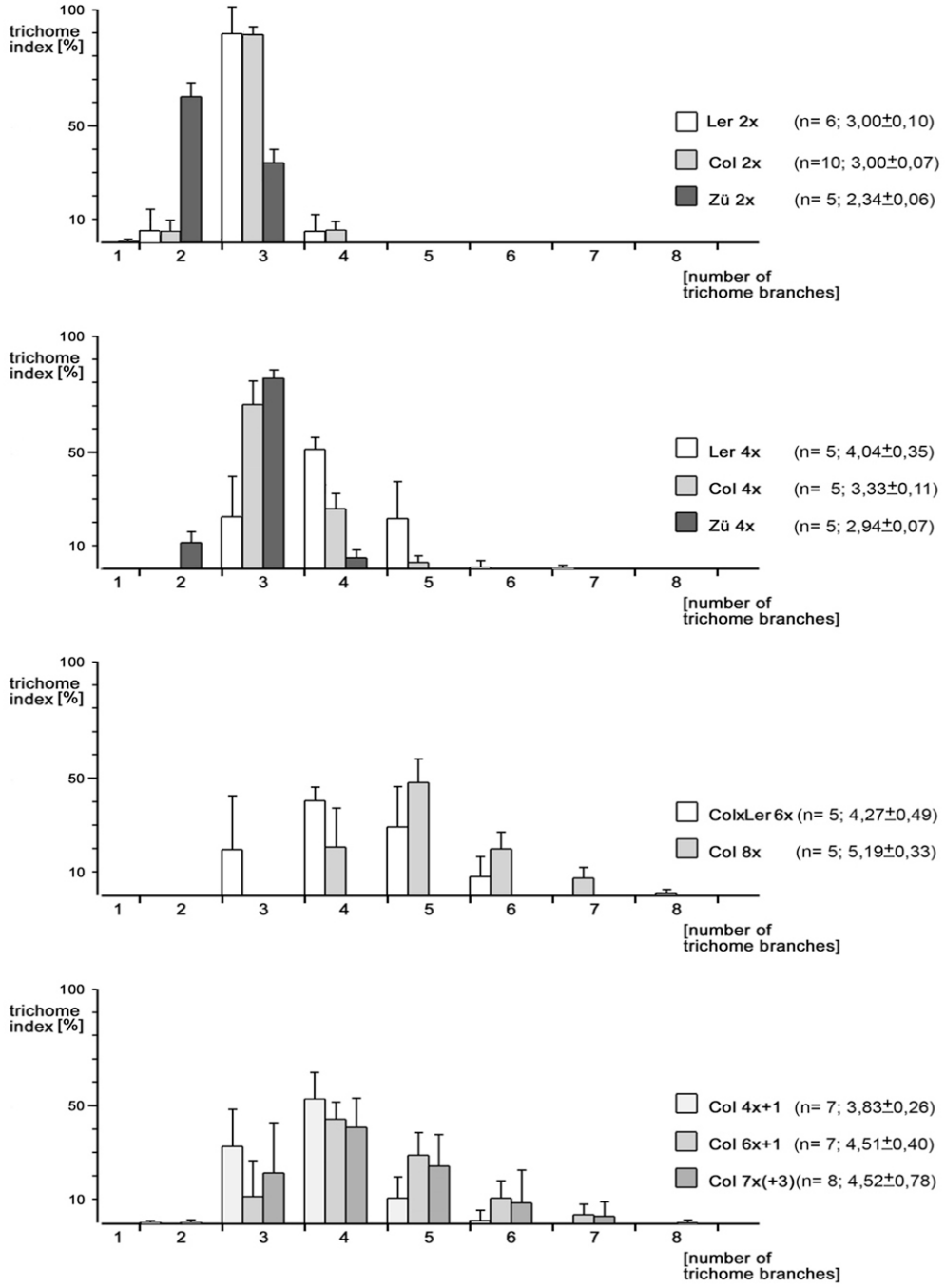


Fig. 4: The relation of basic polyploidy and branch numbers in *Arabidopsis* trichomes (from Yu et al., 2009).
The percentage of the trichomes with a certain number of branches is shown for L*er*-0, Col-0 and Zürich ecotypes; and for diploids, tetraploids, hexaploids, octoploids and aneuploids.

The elevated portion of multi branch trichome actually indicated the increase of the cell size of these trichome cells. Other epidermal cells (for example stomata cells) were also enlarged in polyploid in comparison to diploid plants (Fig. 5).

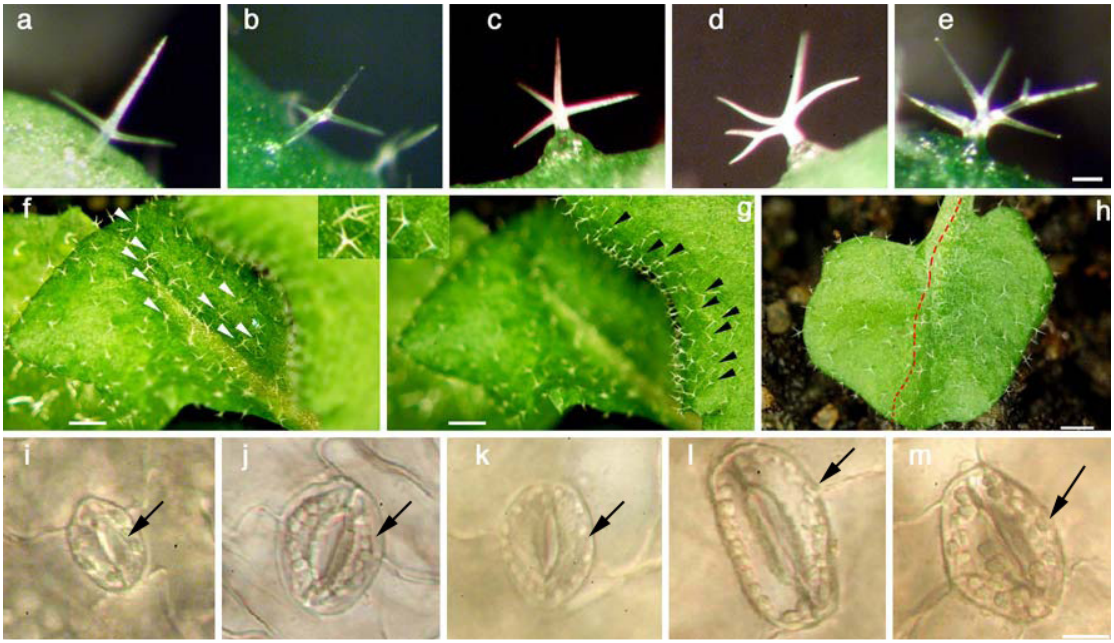


Fig. 5: Cell size and tissue size effects of polyploidy (from Yu et al., 2009).
Shown is the trichome morphology of diploid Col-0 (a), diploid L*er*-0 (b), tetraploid Col-0 (c), octoploid Col-0 (d) and aneuploid Col-0 (e) plants. Sectors of colchicine treated plants as visualized by trichome morphology and leaf morphology (f-h); f) polyploid trichomes (white arrowheads) on a leaf of a sectored plant ecotype Yo-0 (inset shows a magnification of such trichomes), g) a neighbored leaf with diploid trichomes (black arrowheads; inset shows a magnification); h) sectored leaf of a treated CIBC-5 ecotype, diploid sector (left) separated by a stippled line from the polyploid sector (right). Sizes of stomata from diploid Col-0 (i), tetraploid Col-0 (j), octoploid Col-0 (k) and two different (4x+1 and 6x+2) aneuploid Col-0 plants respectively (l+m). Note the elevated numbers of chloroplasts (shown with the arrows) in the stomata of plants with higher ploidies. Scale bars: 0.1mm in e) was the same for a-e); 1mm in f-h); 10µm in m) was the same for i-m).

Similarly, the sizes of microspores, seeds and nuclei were also bigger in the polyploid lines as compared to in the diploid lines (Fig. 6, Fig. 7).

All these size' enlargements in polyploid lines could be used for polyploid candidate selection after induction with colchicine, however, trichome morphology assessment was the most convenient and fastest way for this purpose.

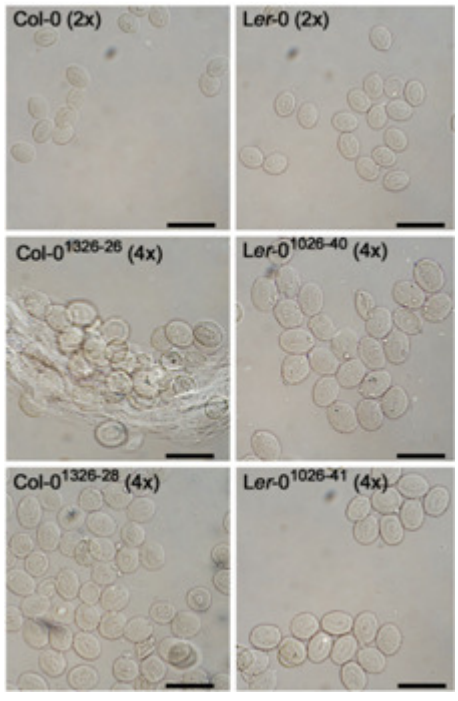


Fig. 6: Size of microspores in tetraploid lines (from Yu et al., 2009).
Comparison between diploid and tetraploid lines, Col-0 (2x), Col-0 (4x) lines 26 and 28, L*er*-0 (2x), L*er-0* (4x) lines 40 and 41.Scale bars: 50µm.

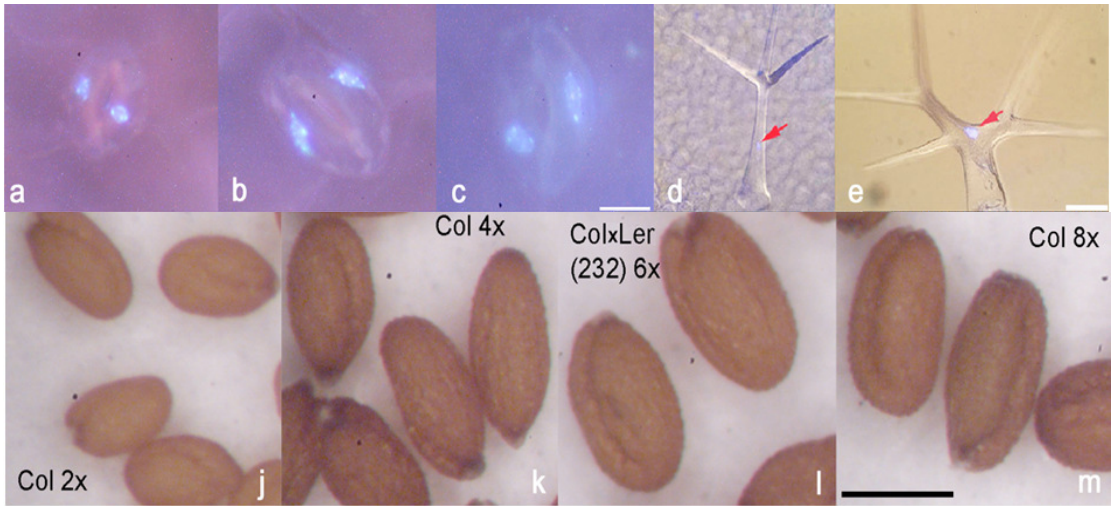


Fig. 7: Additional size features in polyploids (from Yu et al., 2009).
DAPI stained nuclei in stomata (a-c) and trichomes (d+e): a) diploid Col-0, b) tetraploid Col-0, c) octoploid Col-0, d) diploid L*er*-0 and e) aneuploid Col-0. Sizes of seeds from different polyploids as indicated (j-m). The seeds from the hexaploid Col-0 L*er*-0 recombinant inbred line 232 (l).Scale bars: in c) 10µm was the same in a+b); in e) 100µm was the same in d); in m) 50µm was the same in j-l).

3.1.3 Assessment of the polyploid candidate lines

After polyploidy candidate lines had been identified, they had to be assessed in the next generation by metaphase chromosome counting (Fig. 8) and flow cytometry analysis (Fig. 9).

The assessment of the following generations showed that most tetraploid lines were quite stable. They stayed in tetraploid state in the consecutive generations. However, the octoploid and hexaploid lines were often not stable as demonstrated by different ploidy levels in the following generations (Table1 and Appendix Table1).

For chromosome staining and counting, at least 10 seeds were used to produce seedlings, whose root tips were used as material. Seeds from the induced plant (the first generation) produced plants with different ploidy levels (seeds from the polyploidy sector giving polyploid plants and seeds from the diploid sector giving diploid plants). Therefore, the results from the chromosome staining also gave chromosome pictures with different ploidy levels in the first generation. From the second generation onwards, the tetraploid lines only produced tetraploid seeds (Table 1).

For the flow cytometry method the leaves from one plant were used in one sample and at least three plants for each line were measured. The pedigree of tetraploid lines which had been assessed by flow cytometry, trichome analysis and chromosome counting was used for subsequent analyses. In particular four Col-0 tetraploid lines and four L*er*-0 tetraploid lines were followed for three generations and

were Col-0$^{13\text{-}2\text{-}6\text{-}12}$, Col-0$^{13\text{-}2\text{-}6\text{-}19}$, Col-0$^{13\text{-}2\text{-}6\text{-}26}$, Col-0$^{13\text{-}2\text{-}6\text{-}28}$, L*er*-0$^{10\text{-}2\text{-}6\text{-}10}$, L*er*-0$^{13\text{-}2\text{-}6\text{-}40}$, L*er*-0$^{10\text{-}2\text{-}6\text{-}41}$, L*er*-0$^{31\text{-}1\text{-}6\text{-}6}$ respectively (lines in blue in Table 1). Both the chromosome counting method and the flow cytometry method showed that they stayed as tetraploids in all three generations (Table 1, Appendix Fig. 1, 2, 3). Microarray experiments were carried out using these lines as material. Table 1 shows the summary of the assessments of the induced lines. The flow cytometry and chromosome figures for the lines used for microarray and for additional seven tetraploid ecotypes used for qRT-PCR analysis are included in Appendix Fig. 1, 2, 3, and 4.

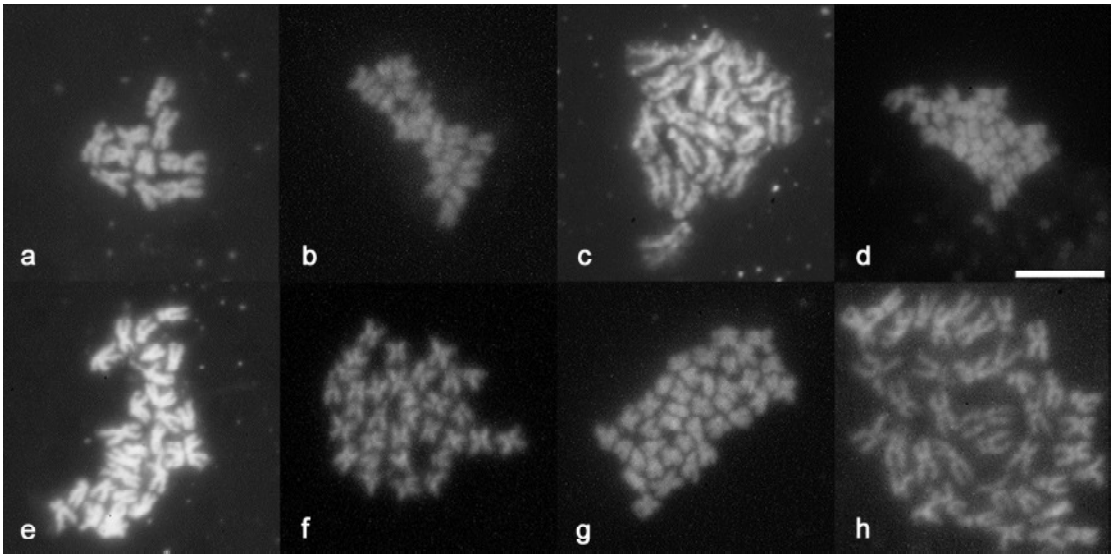


Fig. 8: Metaphase chromosomes of dividing root tip cells (from Yu et al., 2009).
Shown are metaphase chromosomes of different ploidies from the Col-0 ecotype (except the hexaploid, which was of mixed Col-0 x L*er*-0 background): a) 2x, b and c) 4x, d) 4x+1, e) 6x, f) 6x+2, g) 7x [+3] and h) 8x. The chromosomes in a, c, e and h had not reached full condensation. Scale bar in d) is the same for a-h): 10 µm. b), d), f), and g) were from Haage (2005).

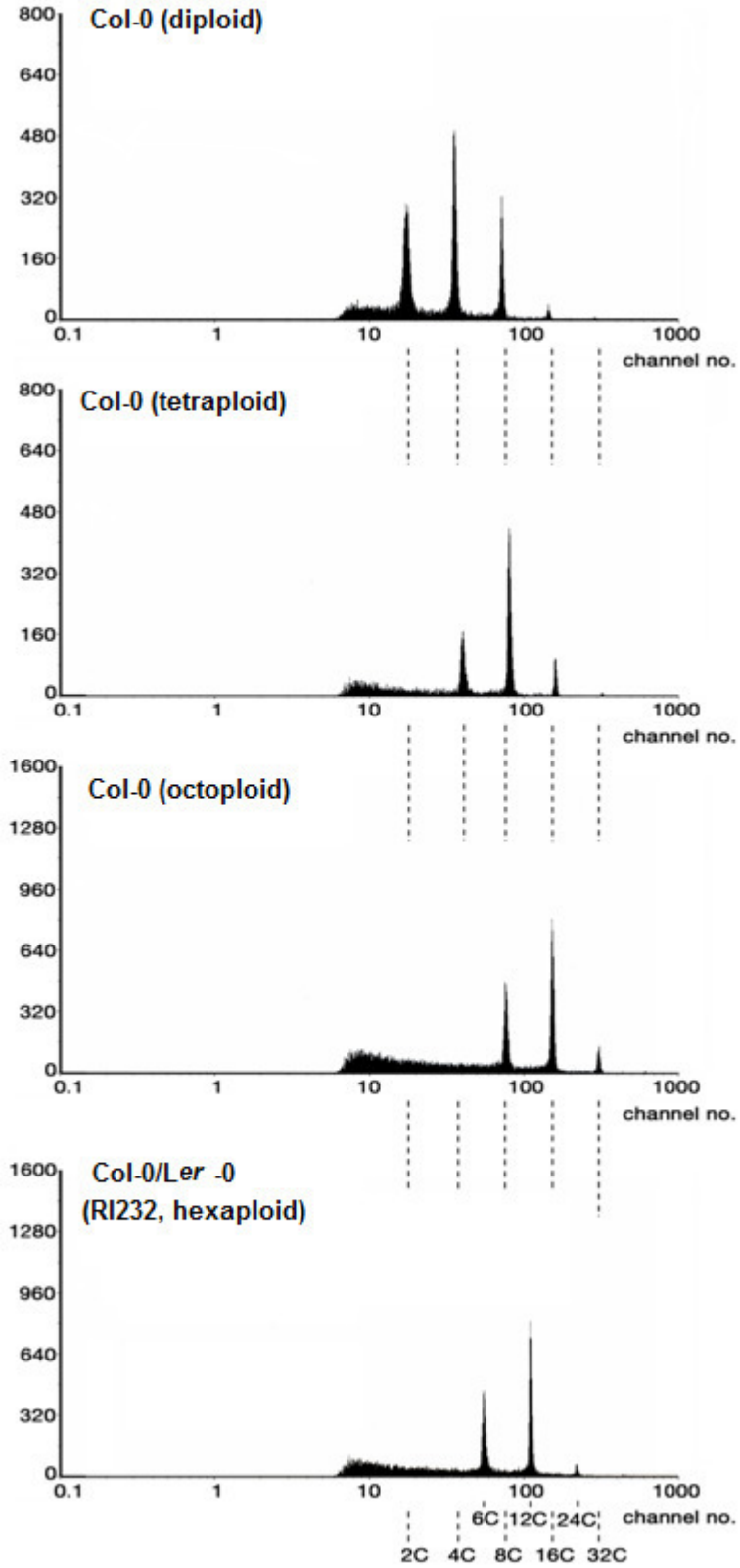


Fig. 9: Flow cytometry of *Arabidopsis* polyploids (from Yu et al., 2009).
The number of measured particle counts (nuclei) versus the channel of the PartecII analyser was given. Note the absence of the 2C peak in the tetraploid and the absence of the 2C+4C peaks in the octoploid plant respectively. Note also the position of the peaks of the hexaploid plant.

Table 1 Summary of the assessment of the induced lines in three generations

lines	First Generation		Second Generation		Third Generation	
Col-0	Chromoso--me staining	Flow Cytometry analysis	Chromoso--me staining	Flow Cytometry analysis	Chromoso--me staining	Flow Cytometry analysis
13-2-6-12 Col-0	20>>10	tetraploid	.20.	tetraploid	.20.	tetraploid
13-2-6-15 Col-0	20>>10	tetraploid	22/20	tetraploid		
13-2-6-18 Col-0	20>>10	diploid				
13-2-6-19 Col-0	.20.	tetraploid	.20.	tetraploid	.20.	tetraploid
13-2-6-26 Col-0	.20.	tetraploid	.20.	tetraploid	.20.	tetraploid
13-2-6-28 Col-0	.20.	tetraploid	.20.	tetraploid	.20.	tetraploid
4n Col-0	.20.	tetraploid	.20.	tetraploid		
Pcol31-1-5-1	40>>20>10	octaploid	.20.	tetraploid	.20.	tetraploid
Pcol31-1-5-2	15>>10	tetraploid	.10.	diploid		
Pcol 31-1-5-3	.40.	octaploid				
ColP9A			.20.	tetraploid	.20.	tetraploid
L*er*-0 0.1%						
10-2-6-5 L*er*-0	20>>10	tetraploid	24/.20.	tetraploid		
10-2-6-10 L*er*-0	20>>10	tetraploid	.20.	tetraploid	.20.	tetraploid
10-2-6-19 L*er*-0	10>>20	diploid				
10-2-6-40 L*er*-0	20>>10	tetraploid	.20.	tetraploid	.20.	tetraploid
10-2-6-41 L*er*-0	20>>10	tetraploid	.20.	tetraploid	.20.	tetraploid
10-2-6-27 L*er*-0	10>>20	diploid				
31-1-6-1 L*er*-0	30/20/10	hexaploid				
31-1-6-2 L*er*-0	20/10	pentaploid				
31-1-6-6 L*er*-0	.20.	tetraploid	.20.	tetraploid	.20.	tetraploid

The root tips of at least 10 seedlings growing from the seeds of the polyploid candidate lines were used in the chromosome staining method. The leaves from one progeny of the candidate line were used in the flow cytometry method. “20>>10” means more cells appeared in tetraploid level than in diploid level. “.20.” means all the cells showed a tetraploid level (for detail see Appendix Table 1). The lines in blue color were used in transcriptome analysis. Line ColP9A originated from the very first inductionexperiments with high colchicine concentrations and turned out to be a stable tetraploid but was not further used.

3.2 Gene expression analysis of tetraploid lines

Col-0 is a widely used ecotype for research. It is completely sequenced and the commercial microarrays based on its genome are available. L*er*-0 ecotype was used as a control experiment in the allopolyploid transcriptomic comparison (Wang et al., 2006 a). Col-0xL*er*-0 recombinant inbred lines are also available for mapping (thirty lines were converted into tetraploids in our lab). Therefore, these two ecotypes were chosen for microarray analysis to compare the transcriptome pattern between diploid and tetraploid plants using seedlings and leaves as material. The plants were grown

within constant light. For seedling experiments, approximately 100 seedlings per microarray analysis were harvested on the fifth day when root tips had penetrated the seed coat. Seeds that did not show emerging roots or which appeared to have germinated prematurely were discarded (<1%). For leaf experiments, the sixth to eighth leaves were harvested for microarray analysis as soon as the bud of the 11th leaf was visible. For all the lines ploidy grades were reassessed before starting transcriptomics. The quality of total RNA was monitored by using the Agilent 2100 bioanalyzer. RNA without detectable degradation and impurities was reverse transcribed into cDNA, and the cDNA was transcribed into cRNA and labeled with Cy3 or Cy5. The cRNA quantity and Cy3/Cy5-labeling efficiency was determined by using a NanoDrop spectrophotometer (Peqlab). cRNA was purified by using RNeasy Plant Mini Kit (Qiagen). cRNA then was fragmentated and applied onto the Agilent *Arabidopsis* 60-mer OligoMicroarray (4 × 44K platform) for hybridization. The hybridized microarray then was scanned and the data were analyzed by Georg Haberer and Klaus F. X. Mayer (Helmholtz Institute, Munich) and Thomas Rattei (TU München).

Although the tetraploid plants exhibited, mainly size differences on cellular and tissue level in comparison to diploids, the overall morphological shapes at rosette stage were similar. Fig. 10 shows the tetraploid and diploid lines used in the microarray analysis and the corresponding chromosome pictures. Preceding microarray analysis, lines were subjected to flow cytometry. Fig. 11 shows the flow cytometry results. Seedlings and the 6^{th} to 8^{th} rosette leaves respectively, were used as material. All the tetraploid plants showed no 2C peak in the results. Together, flow cytometry and chromosome counts were done for more than three times for each line. The additional figures for both methods are included in Appendix Fig. 1, 2 and 3 respectively. The flow cytometry results for the seven additional ecotypes are included in Appendix Fig. 4.

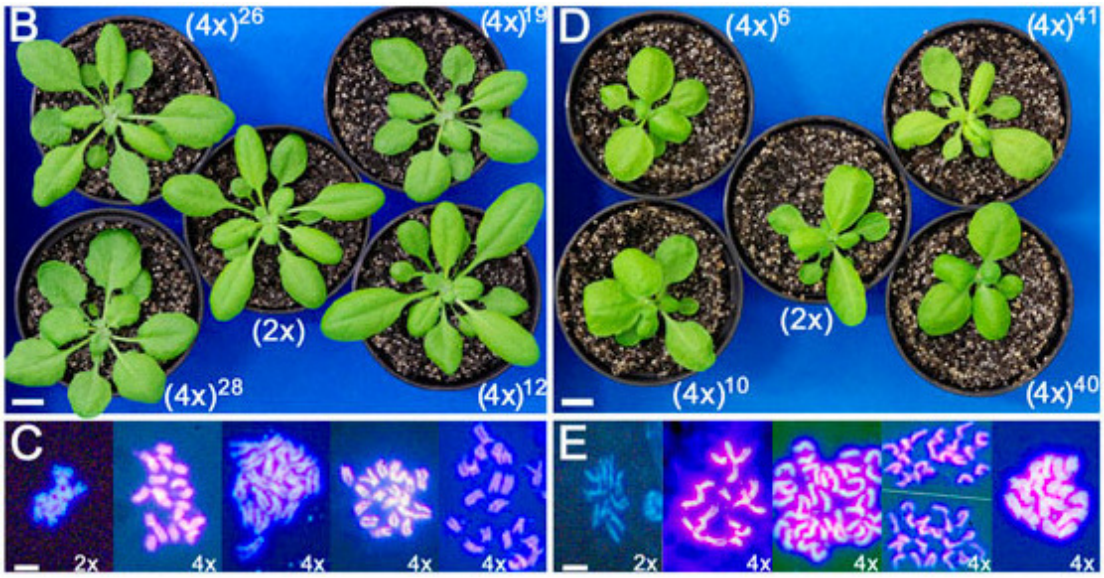


Fig. 10: Morphology and cytology of tetraploid *Arabidopsis thaliana* lines
a) Diploid and tetraploid Col-0 lines 1326-12, -19, -26 and -28 respectively. b) Diploid and tetraploid L*er*-0 lines 3116-6, 1026-10, -40 and -41 respectively. c) Representative mitotic chromosome figures of Col-0 (2x) and the Col-0 (4x) lines 1326-12, -19, -26 and -28 respectively (from left to right). d) Representative mitotic chromosome figures of L*er*-0 (2x) and the L*er*-0 (4x) lines 3116-6, 1026-10, -40 and -41 respectively (from left to right). Note the different mitotic stages; early meta- to late anaphase. Scale bars: 1 cm in a) and b); 10 µm in all subsections of c and d.

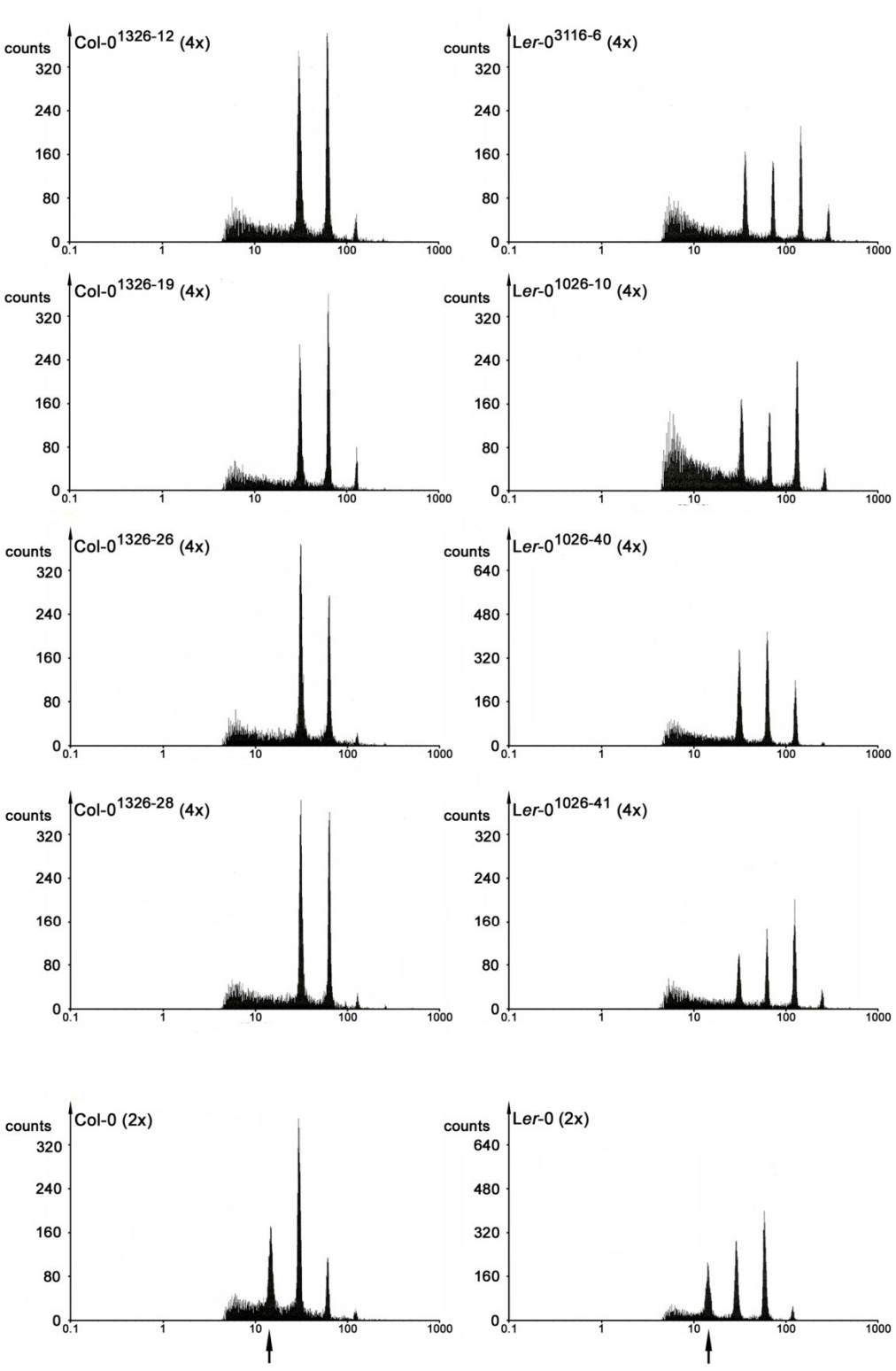


Fig. 11: Flow cytometry analysis of the tetraploid *Arabidopsis thaliana* ecotype lines with Col-0 and L*er*-0 background.
Diploid lines (2x) and tetraploid lines (4x) are indicated. Arrows point to the 2C peak in diploids.

3.2.1 Microarray experiments with seedlings:

Seven experiments (Table 2) were conducted in order to compare the transcriptome pattern difference between diploid and tetraploid of Col-0 (experiment 1 and 7) or diploid and tetraploid of L*er*-0 (experiment 5) ecotype; between consecutive generations of tetraploids (experiment 3 and 4); between Col-0 and L*er*-0 ecotypes of diploid (experiment 6) and tetraploid (experiment 2). Each experiment used four biological replicates. In the comparison of Col-0 diploid with Col-0 tetraploid seedlings, two experiments were conducted (experiment 1 and 7 in Table 2) as dye swap experiments. This means that the labels of diploid and tetraploid materials were reversed in these experiments. In the experiment 1 Col-0 diploid was labeled with Cy3 and Col-0 tetraploid was labeled with Cy5. In the experiment 7 Col-0 diploid was labeled with Cy5 and Col-0 tetraploid was labeled with Cy3.

Table 2: The microarray experiment schedule for different comparisons in seedlings

Experiment number	Experiment discription	First replicate	Second replicate	Third replicate	Fourth replicate
1	Col diploid vs.Col tetraploid C3	Col2n vs. Col12 C3	Col2n vs. Col19 C3	Col2n vs. Col26 C3	Col2n vs. Col28 C3
2	L*er*tetraploid C3 vs. Col tetraploid C3	L*er*6 C3 vs. Col12 C3	L*er*10 C3 vs. Col19 C3	L*er*40 C3 vs. Col26 C3	L*er*41 C3 vs. Col28 C3
3	L*er*tetraploid C2 vs. L*er*tetraploid C3	L*er*6 C2 vs. L*er*6 C3	L*er*10 C2 vs. L*er*10 C3	L*er*40 C2 vs. L*er*40 C3	L*er*41 C2 vs. L*er*41 C3
4	Col tetraploid C2 vs. Col tetraploid C3	Col12 C2 vs. Col12 C3	Col19 C2 vs. Col19 C3	Col26 C2 vs. Col26 C3	Col28 C2 vs. Col28 C3
5	L*er* diploid vs. L*er*tetraploidC3	L*er* 2n vs. L*er*6 C3	L*er* 2n vs. L*er*10 C3	L*er* 2n vs. L*er*40 C3	L*er* 2n vs. L*er*41 C3
6	L*er*diploid vs. Col diploid	L*er* 2n vs. Col2n	L*er* 2n vs. Col 2n	L*er*2n vs. Col 2n	L*er* 2n vs. Col 2n
7	Col tetraploidC3 vs. Col diploid	Col12 C3 vs. Col 2n	Col19 C3 vs. Col 2n	Col26 C3 vs. Col 2n	Col28 C3 vs. Col 2n

C2:the second generation; C3:the third generation; Col2n: Col-0 diploid; Col12: Col-0 13-2-6-12; Col19: Col-0 13-2-6-19; Col26: Col-0 13-2-6-26; Col28: Col-0 13-2-6-28; L*er*6: L*er*-0 31-1-6-6; L*er*10: L*er*-0 10-2-6-10; L*er*40: L*er*-0 10-2-6-40; L*er*41: L*er*-0 10-2-6-41.

3.2.1.1 Comparison between ecotypes (experiment 2 and 6 in Table 2):

For the transcriptome pattern difference between Col-0 and L*er*-0 ecotypes, 860 genes (representing 3.37% of the annotated genes) were found to be differentially expressed between diploid Col-0 seedlings and diploid L*er*-0 seedlings. 219 and 641

of these were over and underexpressed in Col-0 ecotype respectively (Appendix Table 2). This lies in the range which had been previously determined for a similar comparison of Col-0 versus L*er*-0 flowers using Affimetrix microarrays (Schmid et al., 2003), revealing that 553 genes and 408 genes respectively were under- or over-expressed between these two ecotypes. The comparison between four tetraploid L*er*-0 and Col-0 lines respectively revealed 348 differentially expressed genes. 74 and 274 of these were over- and under-expressed in Col-0 respectively (Appendix Table 3).

For the diploid ecotype comparison, 534 genes of the 860 genes were assigned to known Tair Gene Ontologies covering the category “biological process”, and for the tetraploid ecotype comparison, 179 genes of the 348 genes were assigned to this category (Collaboration with T. Rattei, TU München). The “biological process” included multicellular organismal process, reproduction, response to stimulus, cellular process, multi-organism process, biological regulation, metabolic process. For the diploid comparison, 603 genes of the 860 genes were assigned to the category “molecular function”, and for the tetraploid ecotype comparison, 234 genes of the 348 genes were assigned to this category (Collaboration with T. Rattei). The “molecular functions” included catalytic activity, transporter activity, enzyme regulator activity, transcription regulator activity, structural molecule activity, molecular transducer activity, binding, antioxidant activity, electron carrier activity.

3.2.1.2 Comparisons between tetraploid Col-0 seedlings and diploid Col-0 seedlings (experiment 1 and 7 in Table 2):

Experiments 1 and 7 in Table 2 were the dye swap experiments comparing the transcriptome difference in seedlings between Col-0 diploid lines and four independently generated tetraploid Col-0 lines (Yu et al., 2009). Evaluating the data set only from the experiment 1, 22 differentially expressed genes were detected (Appendix Table 4) and evaluating the data set only from the experiment 7, 64 genes were detected (Appendix Table 5). 13 genes overlapped between these two evaluations (Table 3). The statistical power of the combined data sets of these two experiments (4 data sets for each experiment) allowed the detection of a higher number of differently expressed genes (Collaboration with G. Habererand K. F. X. Mayer, Helmholtz Institute, Munich). 589 genes were found differentially expressed (ca. 2.3% of the *A. thaliana* transcriptome). 476 genes of the 589 genes displayed a fold change difference (FC) of at least 1.5 (Appendix Table 6). This FC has been selected as threshold value in the analysis of synthetic allotetraploid *Arabidopsis suecica* lines (Wang et al., 2006b). 166 of the 589 genes displayed at least a two-fold

expression change difference (Appendix Table 6). The reliability of the combined contrast analysis was shown first, by positively qRT-PCR evaluated genes (see Fig. 14), those were absent in the evaluation of one (At5g60250) or the other (At2g25460, At5g60250, At5g64310) experimental sub-series (see below) and second by the very high overlap of the differentially expressed gene groups in both separated series (21 of 22 genes and 55 of 64 genes for experiment 1 and 7 respectively, see Table 3). A similar result was obtained in the leaf analysis (see below). Table 4 shows the top 30 genes in this 589 gene list, according to the fold change and the significance value, the p value. 7 genes (in bold) of them were confirmed by qRT-PCR.

Analysis of Gene Ontology representations indicated a coverage of most important functional groups of biological processes and molecular functions (Table 5) and were extended by a search for significant enrichments of groups, which uncovered under- and over-representation of biological processes such as those related to photosynthesis, membrane, oxidoreductase activity, response to temperature stimulus (Appendix Table 7). This search did not imply that a particular gene is unambiguously involved in the name-giving pathway. It rather indicated a predominance of a pathway, which reflected a characteristic physiology and gene activity of the assayed material. This search was complemented with a deeper comparative in silico analysis based on term-supported matching which delivered a striking enrichment of genes related to the particular developmental programs (Fig. 12, Yu et al., 2010). Thus, 43 of the top-most genes covered ethylene-, stress-, senescence- and defense-related processes respectively (Table 6) with adjusted p-values far below 0.05. Further conspicuous gene groups concerned photosynthesis/chlorophyll (27+37), biosynthesis of sugars and cell wall (112), genes related to metal ions (39), ATPases (14), calcium (12) and 25 transcription factors including eight of the NAM-family (Fig. 12).

3.2.1.3 Comparison between tetraploid L*er*-0 seedlings and diploid L*er*-0 seedlings (Experiment 5 in Table 2)

The same seedling analysis with the ecotype L*er*-0 (NW20) used four independently generated tetraploid L*er*-0 plant lines. The result indicated only 9 genes, which did not overlap with those found in the differentially expressed genes of the Col-0 seedlings analysis. The 9 genes were At3g26520.1 TIP2 (Tonoplast intrinsic protein 2), At4g39090.1 RD19 (responsive to dehydration 19), At4g05050 UBQ11 (Ubiquitin 11), At2g39730 RCA (Rubisco activase), At5g60390 elongation factor 1-alpha, At5g48180.1 Kelch repeat containing protein, At3g16460 Jacalin lectin family protein, At5g43570.1 Serine protease inhibitor, At2g33790.1 pollen Ole e 1 allergen and

extension family protein. All of them were down-regulated in tetraploid lines.Their identity did not indicate a conspicuous functional clustering. At3g26520.1 TIP2 was also differentially expressed between Col-0 diploid leaves and Col-0 tetraploid leaves (see below), but it was up-regulated in the Col-0 tetraploid leaves. At5g60390 was also detected to be differentially expressed between L*er*-0 diploid leaves and L*er*-0 tetraploid leaves and was up-regulated in the L*er*-0 tetraploid leaves (See below).

3.2.1.4 Comparison between consecutive generations of tetraploid L*er*-0 seedlings and of tetraploid Col-0 seedlings (Experiment 3 and 4 in Table 2)

For the comparison of the consecutive generations of the tetraploid lines, the seedlings of the 2nd and 3rd generations after induction were used for comparing both Col-0 and L*er*-0 lines. Cytological analysis of the chromosomal stability of the lines analyzed had been found to be very stable during at least three consecutive generations (Table1, Yu et al., 2009). The results from the microarray data also showed an almost complete identity between both generations and in both ecotypes. L*er*-0 (4x) did not reveal any difference while Col-0 (4x) showed differences in only 6 genes (Appendix Table 8). These were At3g54890 LHCA1 (Light Harvesting Complex Gene 1), At5g54270.1 LHCB3 (Light-harvesting chlorophyll binding protein3), At5g56080.1 (putative Nicotineamine synthase), At4g25100 FSD1 (Iron superoxide dismutase 1), At2g47015.1 microRNA-gene and At1g13080 ATMRP3 ATPase. The first five were over- and the latter under-expressed respectively.

Table 3: Overlap of the differentially expressed genes between the different seedling experimental evaluations.

		L2-C2	L4-C4	C2-C4 Combined data sets	C2-C4	C2-C4	L2-L4
Overlap of the result of the evaluations	**Experiment**	**6 in Table 2**	**2 in Table 2**	**1+7 in Table 2**	**7 in Table 2**	**1 in Table 2**	**5 in Table 2**
Experiment		860	348	589	64	22	9
6 in Table 2	860	860	224	37	7	1	0
2 in Table 2	348	224	348	15	4	2	0
1+7 in Table 2	589	37	15	589	55	21	0
7 in Table 2	64	7	4	55	64	13	1
1 in Table 2	22	1	2	21	13	22	0
5 in Table 2	9	0	0	0	1	0	9

Loess-holm substracted significance value $p<0.05$, L2: L*er*-0 diploid, C2: Col-0 diploid, L4: L*er*-0 tetraploid, C4: Col-0 tetraploid

Table 4: Top 30 differentially expressed genes in seedlings – diploid vs. tetraploid Col-0 (two times four biological replicates, Cy3/Cy5-dye swap).

GeneID	Description	logFC[2]	P<0.05[1]
AT4G13420.1	HAK5 (High affinity K+ transporter 5); potassium ion transporter	1,7778	5,11E-08
AT5G57760.1	unknownprotein	1,6590	3,68E-07
AT1G53480.1	unknownprotein	4,6430	5,58E-07
AT2G20800.1	NDB4 (NAD(P)H DEHYDROGENASE B4); NADH dehydrogenase	-1,4789	8,01E-07
AT5G01380.1	transcriptionfactor	-1,4724	1,08E-06
AT1G53490.1	DNA binding	1,1692	2,31E-06
AT5G09570.1	unknownprotein	-2,9358	2,76E-06
AT5G64870.1	unknownprotein	-1,9268	5,21E-06
AT1G29395.1	COR414-TM1 (cold regulated 414 thylakoid membrane 1)	1,1199	5,36E-06
AT2G18193.1	AAA-type ATPase family protein	-1,2257	6,13E-06
AT5G62520	SRO5 (SIMILAR TO RCD ONE 5); NAD+ ADP-ribosyltransferase	-1,2269	7,55E-06
AT4G28850.1	xyloglucan:xyloglucosyltransferase, putative / xyloglucanendotransglycosylase, putative / endo-xyloglucantransferase, putative	2,6847	1,09E-05
AT2G19190.1	FRK1 (FLG22-INDUCED RECEPTOR-LIKE KINASE 1); kinase	1,6762	1,16E-05
AT1G18140.1	LAC1 (Laccase 1); copper ion binding / oxidoreductase	1,1724	1,46E-05
AT5G42380.1	calmodulin-relatedprotein, putative	-1,3647	1,50E-05
AT4G19680	IRT2 (iron-responsivetransporter 2); ironiontransporter/ zinciontransporter	1,2251	1,63E-05
AT1G73830.1	BEE3 (BR ENHANCED EXPRESSION 3); DNA binding / transcription factor	1,3919	2,13E-05
AT4G14060.1	major latex protein-related / MLP-related	1,3302	2,13E-05
AT2G23150.1	NRAMP3 (NRAMP metaliontransporter 3); manganeseiontransporter/ metaliontransporter	-0,7791	2,20E-05
AT1G80660.1	AHA9 (*Arabidopsis* H(+)-ATPase 9); hydrogen-exporting ATPase, phosphorylative mechanism	-2,1788	2,24E-05
AT5G52940.1	unknownprotein	-1,3623	2,42E-05
AT5G48010	pentacyclictriterpenesynthase, putative	0,9256	2,97E-05
AT1G70440.1	SRO3 (SIMILAR TO RCD ONE 3); NAD+ ADP-ribosyltransferase	-2,0297	3,63E-05
AT1G28760.1	unknownprotein	-1,4336	3,77E-05
AT5G54710.1	ankyrinrepeatfamilyprotein	0,9805	3,83E-05
AT5G19220.1	ADG2 (ADPG PYROPHOSPHORYLASE 2); glucose-1-phosphate adenylyltransferase	1,0658	4,18E-05
AT1G17960.1	threonyl-tRNAsynthetase, putative / threonine--tRNA ligase, putative	-0,8739	4,20E-05
AT1G80670.1	transducin family protein / WD-40 repeat family protein	-1,1553	5,29E-05
AT3G60160.1	ATMRP9 (*Arabidopsis thaliana* multidrug resistance-associated protein 9)	-0,9247	5,34E-05
AT5G39530.1	unknownprotein	0,8820	5,65E-05

AGI codes in bold indicated genes assessed by qRT-PCR.[1] loess-holm substracted P-value; [2] Fold change $\log_2$

Table 5: GeneOntologies of the differentially expressed genes in seedlings of tetraploid Col-0 plants (In collaboration with T. Rattei and M. Matthes, TU München)

GOs - biologicalprocesses		
Col-0 2n vs. Col-0 tetraploid seedlings		
Goid	**#genes**	**GO description**
GO:0032502	7	developmentalprocess
GO:0032501	10	multicellularorganismalprocess
GO:0000003	7	reproduction
GO:0050896	68	responsetostimulus
GO:0009987	15	cellularprocess
GO:0051704	12	multi-organismprocess
GO:0065007	55	biologicalregulation
GO:0008152	161	metabolicprocess
GO:0051179	43	localization
GOs - molecularfunctions		
Col-0 2n vs. Col-0 tetraploid seedlings		
Goid	**#genes**	**GO description**
GO:0003824	178	catalytic activity
GO:0005215	32	transporter activity
GO:0030234	4	enzymeregulatoractivity
GO:0030528	7	transcriptionregulatoractivity
GO:0005198	5	structuralmoleculeactivity
GO:0060089	7	moleculartransduceractivity
GO:0005488	161	binding
GO:0016209	8	antioxidantactivity
GO:0009055	3	electroncarrieractivity

Table 6: Ethylene-related differential gene expression program in Col-0 4x seedlings

GeneID	**Description**	**P<0.05[1]**	**logFC[2]**
Ethylenesynthesis, perception, response			
AT5G37990.1	S-adenosylmethionine-depend. met.-transfer.	0.01533	0.59527
AT1G62380.1	ACO2 (ACC OXIDASE 2)	0.00498	0.66891
AT3G61400.1	2-oxoglutarate-depend. dioxygenase, putative	0.00219	-0.96541
AT5G43450.1	2-oxoglutarate-depend. dioxygenase, putative	0.00016	0.75241
AT1G04310.1	ERS2 (ETHYLENE RESPONSE SENSOR 2)	0.00425	0.57871
AT1G06160.1	ethylene-responsivefactor, putative	0.03924	0.57534
AT1G25470	AP2 domain-containing TF, putative	0.00435	-0.60953
AT1G43160.1	RAP2.6 (related to AP2 6); DNA binding / TF	0.00142	0.97463
AT2G31230.1	ATERF15 (ETHYLENE-RESPONSIVE TF)	0.02281	0.73769
AT2G38340.1	AP2 domain-containing TF, putative (DRE2B)	0.00587	-1.06908
AT2G44940.1	AP2 domain-containing TF TINY, putative	0.03193	0.48035
AT2G47520.1	AP2 domain-containing TF, putative	0.00012	-1.14349
AT4G34410.1	AP2 domain-containing TF, putative	0.00050	-1.51963
AT5G57760.1	unknownprotein, (MPK4 dep.,ethylenerelat.)	0.00000	1.65902
senescence/stress/defence			
AT2G40340	AP2 domain-containing TF, putative	0.00742	-1.1458
AT3G11020	DRE B2B (DRE-bindingprotein 2B), TF	0.00070	-1.4454
AT1G17020.1	SRG1 (SENESCENCE-RELATED GENE 1)	0.00976	-0.67608
AT1G22160.1	senescence-associated protein-related	0.02382	0.46378
AT1G51820.1	leucine-rich repeat protein kinase, putative	0.00017	0.92829
AT1G69490.1	NAP (NAC-LIKE, ACTIVATED BY AP3/PI);TF	0.00166	1.31951
AT2G19190.1	FRK1(FLG22-IND. RECEPT.-LIKE KINASE1)	0.00001	1.67618
AT2G29290.1	tropinonereductase/dehydrogenase, putative	0.01168	0.72892
AT4G30270.1	MERI5B (MERISTEM-5); hydrolase	0.01424	-0.69498
AT4G30430.1	TET9 (TETRASPANIN9)	0.00121	-1.56025
AT4G35770	SEN1 (DARK INDUCIBLE 1)	0.00218	1.46476
AT5G23030.1	TET12 (TETRASPANIN12)	0.02888	1.03791

AT5G41080	glycerophosphoryldiesterphosphodiesterase	0.00497	0.90135
AT1G29390	COR314-TM2 (cold regul.314 thyl. membr.2)	0.01515	0.55457
AT1G29395.1	COR414-TM1 (cold regul. 414 thyl. membr.1)	0.00001	1.11990
AT1G35720.1	ANNAT1 (ANNEXIN ARABID.1);calcium bind.	0.00033	0.84941
AT1G73330.1	ATDR4 (A. thaliana drought-repressed 4)	0.02699	0.60120
AT2G25940.1	ALPHA-VACUOLAR PROCESS. ENZYME)	0.01315	-0.45956
AT3G51910.1	AT-HSFA7A (A. thaliana heat shock TF A7A);	0.01067	1.32320
AT4G03430.1	STA1 (STABILIZED1); RNA splicing factor,	0.00479	-0.72075
AT4G12040	zinc finger (AN1-like) family protein	0.00966	-0.55338
AT4G37930.1	SHM1 (SERINE HYDROXYMET.-TRANSF.1)	0.00131	0.98870
AT5G38710.1	proline oxidase, putat./osmotic stress-respd.	0.00044	-0.70265
AT5G62520	SRO5, NAD+ ADP-ribosyltransferase	0.00001	-1.22689
AT1G55020.1	LOX1 (Lipoxygenase 1); lipoxygenase	0.00185	-0.80797
AT4G37930.1	SHM1 (SERINE HYDROXYMET.-TRANSF.1)	0.00131	0.98870
AT5G48657	defense protein-related	0.00914	0.73791
AT4G22214.1	unknownprotein	0.00045	2.43343
AT4G22217.1	unknownprotein	0.00378	2.07322
AT2G43535.1	trypsininhibitor, putative	0.01544	0.56456

TF: transcription factor; [1] loess-holm substracted P-value; [2] Fold change $\log_2$

3.2.2 Microarray experiments with leaves:

Five experiments were performed to compare the diploid and tetraploid leaf transcriptome difference (Table 7). For Col-0 ecotype, three experimental series were performed (total of 12 microarray experiments). One was to compare the four independently induced tetraploid lines with diploid lines (experiment 8 in Table 7). The other two were to compare separately four individual sibling lines from the same induced line of Col-$0^{1326\text{-}28}$ or Col-$0^{1326\text{-}26}$ with the diploid lines to exclude the individual difference (Experiment 9 and 10 in Table 7). For L*er*-0 ecotype, the same experiments were performed independently two times to have 8 microarray replicates for the statistical evaluation (Experiment 11 and 12 in Table 7).

Table 7: The microarray experiment schedule for different comparisons in leaves

Experiment number	Experiment discription	First replicate	Second replicate	Third replicate	Fourth replicate
8	Col diploid vs. Col tetraploid C3	Col2nvs. Col12 C3	Col2n vs. Col19 C3	Col2n vs. Col26 C3	Col2n vs. Col28 C3
9	Col diploid vs. Col tetraploid C3 Individual line	Col2n vs. Col26-1C3	Col2n vs. Col26-2 C3	Col2n vs. Col26-3 C3	Col2n vs. Col26-4 C3
10	Col diploidvs. Col tetraploid C3 Individual line	Col2n vs. Col28-1 C3	Col2n vs. Col28-2 C3	Col2n vs. Col28-3 C3	Col2n vs. Col28-4C3
11	L*er*diplolidvs. L*er*tetraploid C3	L*er* 2n vs. L*er*6 C3	L*er* 2n vs. L*er*10 C3	L*er* 2n vs. L*er*40 C3	L*er* 2n vs. L*er*41 C3
12	L*er* diploid vs. L*er*tetraploidC3	L*er* 2n vs. L*er*6 C3	L*er* 2n vs. L*er*10 C3	L*er* 2n vs. L*er*40 C3	L*er* 2n vs. L*er*41 C3

C3:the third generation; Col2n: Col-0 diploid; Col12: Col-0 13-2-6-12; Col19: Col-0 13-2-6-19; Col26: Col-0 13-2-6-26; Col28: Col-0 13-2-6-28; L*er*6: L*er*-0 31-1-6-6; L*er*10: L*er*-0 10-2-6-10; L*er*40: L*er*-0 10-2-6-40; L*er*41: L*er*-0 10-2-6-41

3.2.2.1 Comparison between Col-0 diploid leaves and Col-0 tetraploid leaves (Experiment 8, 9, 10 in Table 7)

Similar to the situation in seedlings, the evaluation of the combined data sets of the experiment 8, 9 and 10 in Table 7 (Col-0 diploid lines compared to four independently induced Col-0 tetraploid lines and to four individual plants from the pedigree of tetraploid lines Col-$0^{1326\text{-}26}$ and Col-$0^{1326\text{-}28}$ respectively), which together gave 12 biological replicates, resulted in higher gene expression alteration numbers than the evaluation of the separated experimental sub-series based on four biological replicates (Collaboration with Georg Haberer, Helmholtz institute Munich). 289 genes were found differentially expressed in the combined evaluation (Appendix Table 9). These gene expression alterations were distinct to those of tetraploid seedlings. 192 of these 289 genes were more than 1.5 x up- and 55 were more than 0.67 x down-regulated, and 42 genes exhibited more subtle changes. The evaluation of the data from the experiment 8, 9, 10 separately in Table 7 gave 60, 201, 19 differentially expressed genes respectively (Appendix Table 10, 11 and 12). 5 of the 60 genes, 30 of the 201 genes and 4 of the 19 genes overlapped with the results of the combined evaluation (Table 8). The reliability of the result from the combined evaluation was proven by qRT-PCR (see below) for the genes At2g40610, At4g32280, At5g18010 and At5g48900, which were (partly) absent in experimental sub-series but were clearly highlighted in the combined contrast analysis (Appendix Table 9).

The in silico scan of the genes obtained in tetraploid Col-0 indicated that the leaf had shifted its programs completely including a strong reduction of the seedling pattern (Fig. 12, Appendix Table 13). Only 13 genes (FC>1.5) overlapped with the differential expressed gene list of the Col-0 diploid tetraploid seedlings comparison. The “cell wall/sugar program” (87 vs. 26 genes) were extensively reduced by more than 70% and exhibited only two overlaps (At1g22400 and At4g30270). No NAM transcription factors were found. The leaf shut down the ethylene/stress program (9 genes) and initiated an auxin synthesis and signalling program (20 genes; FC>2.0 for all; Table 9) with predominance of *IAA* (*Indoleacetic Acid induced*) and *SAUR*-like genes (*Short Auxin Up-regulated RNAs*, Gil et al., 1994). The former are known as antagonists of auxin responsive factors (ARFs) (Ulmasov et al., 1999). The latter, whose function is still unknown, were organized into two clusters on chromosome 1 and 5 respectively. According to the microarray data, not all members of the clusters were over-expressed. *SAUR* gene cluster at the fifth chromosome (designated At5g180-c) comprised six highly homologous copies, At5g18010, At5g18020, At5g18080, At5g18030, At5g18050, At5g18060. The six genes were dispersed in a region of 20

kb withsome copies <2 kb apart. Over-expression of At5g180-c was predominantly caused by At5g18010 according to the following test. The cDNA fragment was amplified from both diploid leaves and tetraploid leaves using the primer pair which can amplify all the cluster genes resulting in a mixture of the cDNA fragments including all the transcribed cluster genes. Then the cDNA fragments mixture was cloned and the clones were sequenced. The sequences of the clones were compared with the sequences of the cluster genes. At5g18010 displayed the highest, namely three fold over representation within cDNA clones from tetraploid (12/58 clones) vs. clones from diploid tissue (2/29 clones).

Table 10 shows the top 30 differentially expressed genes according to the fold change and the significance. 10 of them were confirmed by qRT-PCR (see below).

3.2.2.2 Comparison between L*er*-0 diploid leaves and L*er*-0 tetraploid leaves: (Experiment 11 and 12 in Table 7)

Similar to the L*er*-0 seedling analysis, the evaluation of the combined data sets of experiment 11 and 12 in Table 7, which were based on 8 biological replicates, revealed only 22 (FC ≥1.5) differentially expressed genes (Appendix Table 14) with 18 up- and 4 down-regulated (five additional genes exhibited more subtle changes) genes. At5g60390, which encoded elongation factor 1-alpha, was up-regulated in tetraploid L*er*-0 leaves (compared to diploid L*er*-0 leaves) and down-regulated in tetraploid Ler-0 seedlings (compared to diploid L*er*-0 seedlings).

Table 8: Overlap of the differentially expressed genes between the different leaf experimental evaluations.

		C2-C4 L	C2-C4 L	C2-C4 L	C2-C4 L
Overlap of the result of the evaluations	experiment	8+9+10 in Table 7	8 in Table 7	9 in Table 7	10 in Table 7
experiment		289	60	201	19
8+9+10 in Table 7	289	289	5	30	4
8 in Table 7	60	5	60	6	3
9 in Table 7	201	30	6	201	12
10 in Table 7	19	4	3	12	19

loess-holm substracted significance value p<0.05, L2: L*er*-0 diploid, C2: Col-0 diploid, L4: L*er*-0 tetraploid, C4: Col-0 tetraploid

Table 9: Auxin-related differential gene expression program in Col-0 4x leaves.

GeneID	Description	P<0.05[1]	logFC[2]
	***IAA* and ARFs**		
AT4G32280.1	*IAA*29 (indoleacetic acid-induced prot. 29); TF	0.00002	3.47777
AT1G15580.1	*IAA*5 (indoleacetic acid-induced prot. 5); TF	0.00009	2.11704
AT3G62100.1	*IAA*30 (indoleacetic acid-induced prot. 30); TF	0.00028	1.05211
AT3G23030.1	*IAA*2 (indoleacetic acid-induced prot. 2); TF	0.00734	0.98331
AT4G14550.1	*IAA*14 (SOLITARY ROOT); transcription factor	0.00318	1.91286
AT3G15540.1	*IAA*19 (indoleacetic acid-induced prot. 19); TF	0.02768	1.20338
AT2G46530	ARF11 (AUXIN RESPONSE FACTOR 11); TF	0.01600	1.14890
	AUX and PINs		
AT2G38120.1	AUX1 (AUXIN RESIST.1); amino acid transp.	0.00002	1.52125
AT1G23080	PIN7 (PIN-FORMED 7)	0.00222	1.49477
	AUXIN responsive genes		
AT5G18010.1	auxin-responsive protein, putative	0.00015	2.63509
AT5G18050.1	auxin-responsive protein, putative	0.00019	2.64581
AT5G18080.1	auxin-responsive protein, putative	0.00031	1.99739
AT5G18030.1	auxin-responsive protein, putative	0.00558	2.18128
AT5G18020.1	auxin-responsive protein, putative	0.00863	1.85869
AT1G29460.1	auxin-responsive protein, putative	0.01181	2.28179
AT1G29500.1	auxin-responsive protein, putative	0.01850	1.73619
AT3G59070.1	auxin-responsive protein, putative	0.66339	-0.37350
AUXIN synthesis			
AT4G13260.1	YUC2 (YUCCA2); monooxygen./ oxidoreduct.	0.00122	1.13341
AT4G28720.1	flavin-containing monooxygen. family protein /	0.00061	1.31848
AT1G04180.1	flavin-containing monooxygen. family protein /	0.00008	2.13776
others			
AT1G50280.1	phototropic-responsive NPH3 family protein	0.00254	1.28201

TF: transcription factor; [1] loess-holm substracted P-value; [2] Fold change $\log_2$

Table 10: Top 30 differentially expressed genes in leaves – diploid vs. tetraploid Col-0 (three times four biological replicates).

GeneID	Description	logFC[2]	P<0.05[1]
AT1G53480.1	unknown protein	4,714	0
AT5G66580.1	unknown protein	2,303	3,93E-06
AT1G53490.1	DNA binding	1,007	1,11E-05
AT2G38120.1	AUX1 (AUXIN RESISTANT 1); amino acid permease/ transporter	1,521	1,60E-05
AT4G32280.1	*IAA*29 (indoleacetic acid-induced protein 29); transcription factor	3,478	1,80E-05
AT5G44210.1	ATERF-9/ATERF9/ERF9 (ERF domain protein 9); DNA binding / transcription factor/ transcriptional repressor	1,818	2,02E-05
AT1G09250.1	transcription factor	1,438	2,02E-05
AT5G39860.1	PRE1 (PACLOBUTRAZOL RESISTANCE1); DNA binding / transcription factor	2,709	2,67E-05
AT5G48900.1	pectate lyase family protein	2,269	7,07E-05
AT1G04180.1	flavin-containing monooxygenase family protein / FMO family protein	2,138	8,44E-05
AT1G15580.1	*IAA*5 (indoleacetic acid-induced protein 5); transcription factor	2,117	9,29E-05
AT5G43810.1	ZLL (ZWILLE)	0,870	9,94E-05
AT5G64770.1	unknown protein	1,396	1,01E-04
AT1G06080.1	ADS1 (DELTA 9 DESATURASE 1); oxidoreductase	4,160	1,14E-04
AT2G40610.1	ATEXPA8 (*ARABIDOPSIS THALIANA* EXPANSIN A8)	3,339	1,15E-04
AT5G18010.1	auxin-responsive protein, putative	2,635	1,45E-04
AT1G51820.1	leucine-rich repeat protein kinase, putative	-2,739	1,51E-04
AT2G32870.1	meprin and TRAF homology domain-containing protein / MATH domain-containing protein	1,530	1,78E-04
AT5G18050.1	auxin-responsive protein, putative	2,646	1,92E-04
AT1G26945.1	transcription regulator	2,381	2,05E-04
AT3G62100.1	*IAA*30 (indoleacetic acid-induced protein 30); transcription factor	1,052	2,76E-04

AT1G02350.1	protoporphyrinogen oxidase-related	1,841	2,76E-04
AT5G49100.1	unknown protein	1,459	2,87E-04
AT4G19810.1	glycosyl hydrolase family 18 protein	-0,854	2,88E-04
AT5G18080.1	auxin-responsive protein, putative	1,997	3,08E-04
AT3G21330.1	basic helix-loop-helix (bHLH) family protein	2,122	3,25E-04
AT5G03180.1	zinc finger (C3HC4-type RING finger) family protein	0,941	4,67E-04
AT5G50335.1	unknown protein	2,100	5,06E-04
AT5G02540.1	short-chain dehydrogenase/reductase (SDR) family protein	2,677	5,58E-04
AT5G66590.1	allergen V5/Tpx-1-related family protein	2,246	5,81E-04

AGI codes in bold indicated genes assessed by qRT-PCR.[1] loess-holm substracted P-value; [2] Fold change $\log_2$

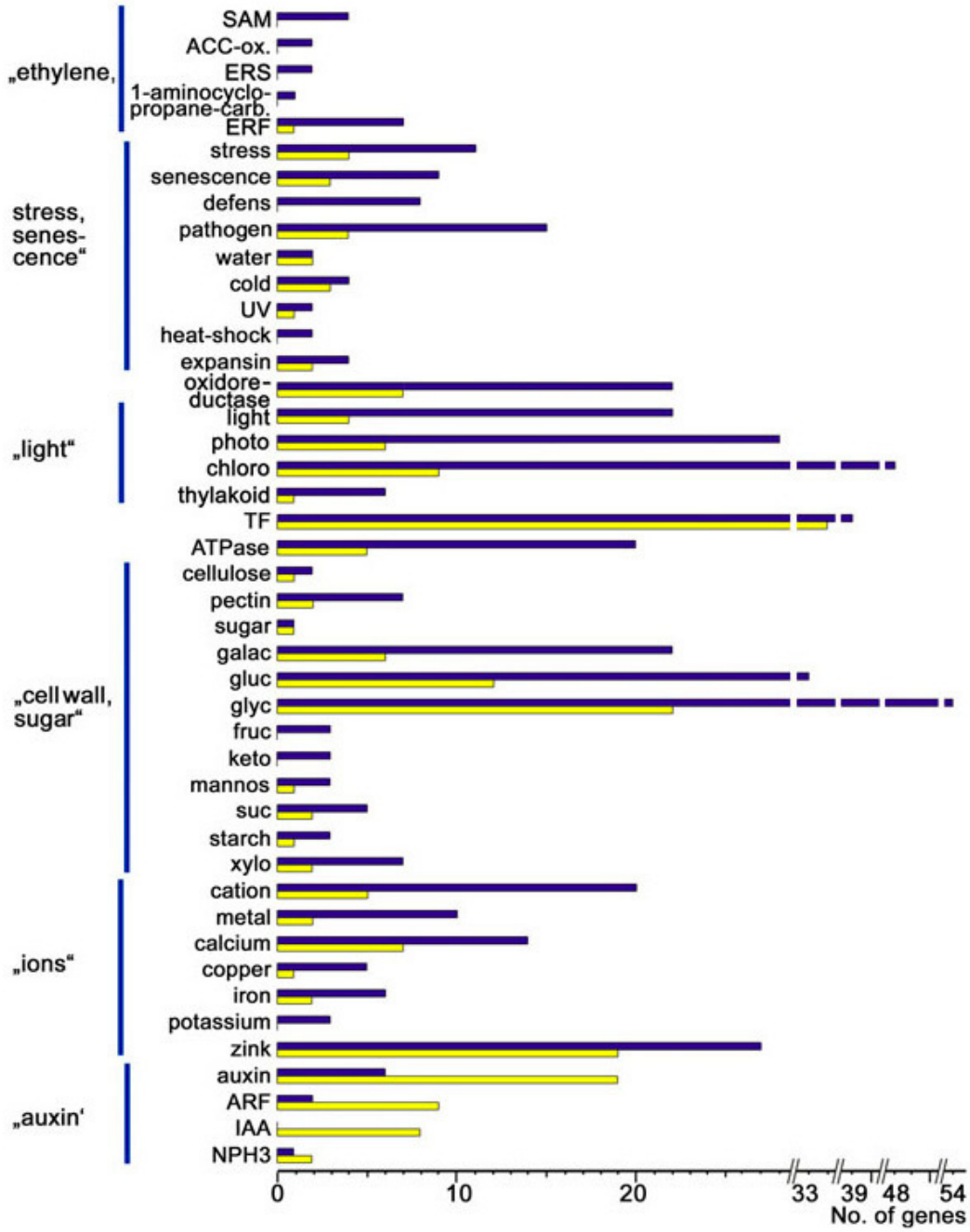


Fig. 12: In silico term-supported GO-representation of genes with altered expression in Col-0 tetraploids as found in TAIR linked to indicated functional terms (collaboration with M. Matthes and T. Rattei) (from Yu et al., 2010).

Blue and yellow bars indicated detected genes found in seedlings and leaves respectively. Bottom designations indicated key processes or functional areas respectively covered by the selected functional terms.

3.2.3 RT-PCR results

To verify the microarray data, several genes from the differential expression list (appendix table 6, Col-0 diploid seedlings versus tetraploid seedlings, and appendix table 9, Col-0 diploid leaf versus tetraploid leaf) were tested with RT-PCR. More than 50% of them were consistent with the microarray results. Fig. 13a shows the RT-PCR results for some genes in the seedling differential expression list. The results are consistent with the microarray data. Three of the genes (At2g23150, At2g18193, At2g20800) were down-regulated in the tetraploid Col-0 seedlings compared to in the diploid seedlings and one gene (At5g57760) was up-regulated. The bands of Act2 show that the same amount of total RNA from diploid and tetraploid seedlings had been applied in the experiment. At2g23150 encodes a member of the Nramp2 metal transporter family, which localized in vacuolar membrane. At2g18193 encodes AAA-type ATPase family protein, functioning in ATP binding. At2g20800 encodes NAD(P)H dehydrogenase, which is located in extrinsic to mitochondrial inner membrane, in mitochondrion, and in plastid. At5g57760 is associated with ethylene induced activity of a defense pathway controlled by MPK4 (MAP Kinase4). Fig. 13 b shows examples for the genes from the Col-0 leaf differential expression list. Three of them (At5g48900, At5g39860, At5g66580) clearly showed the consistent differential expression with the microarray data (up-regulated in tetraploid). The gene At2g38120 does not clearly show differential expression in the result. At5g48900 encodes a member of pectate lyase family protein, which is located in endomembrane system. At5g39860 encodes a transcription factor which is involved in unidimensional cell growth and responds to brassinosteroid stimulus. At5g66580 encodes an unknown protein, which is involved in N-terminal protein myristoylation. RT-PCR was just a preliminary test to give a fast impression of the quality of the microarray data. Most genes tested by this method were then tested with qRT-PCR to get more accurate results. In the RT-PCR experiment, mostly only one plant for one sample was used to test the differential expression, however, in the subsequent qRT-PCR experiments, at least three biological replicates were used for each sample (more close to the amount of replicates used in microarray compared to the amount of replicates in RT-PCR) to exclude the variation among different plants in the same ploidy level.

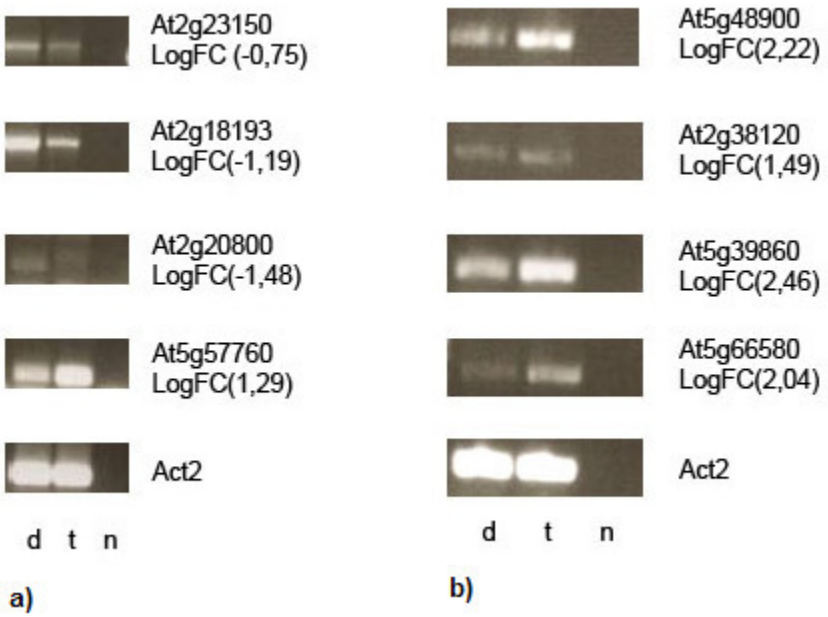


Fig. 13: RT-PCR results of the genes in the Col-0 differential expression list a) for seedlings (see Appendix Table 6); b) for leaves (see Appendix Table 9).
Three successive gel pockets were loaded with the RT-PCRs using the same gene primer pairs. In the first pocket (d) the diploid cDNA was used as the template and in the second (t) the tetraploid cDNA was used. The third pocket (n) was a negative control, in which the template had been omitted. Each gene locus and the expression $\log_2$ fold change in the microarray results are indicated.

3.2.4 qRT-PCR results

qRT-PCR experiments of selected microarray-detected genes were performed in order to test whether Col-0 tetraploids modulate physiological programs during development. The selected genes covered different functional categories from both seedlings and leaves respectively (at least three biological replicates; Fig. 14). The differentially expressed genes between Col-0 diploid seedlings and Col-0 tetraploid seedlings assessed by qRT-PCR were: genes encoding proteins related to (putative) ion binding/transport or using ions as cofactors such as HAK5 (At4g13420, potassium, showing highest absolute over-expression), At2g25460 (Calcium), At5g60250 (Zink) and LAC1 (At1g18140; copper as cofactor); a gene encoding an ethylene related protein At5g57760 (associated with ethylene induced activity of a defense pathway controlled by MPK4); genes encoding proteins related to arabinogalactan, pectin and cellulose/cell wall such as AGP1 (At5g64310, arabinogalactanprotein), and At2g25460 which is also capable of binding calcium; finally a gene encoding a transcription factor (At5g01380). The differentially expressed genes between Col-0 diploid leaves and Col-0 tetraploid leaves assessed by qRT-PCR were: an ethylene related gene At5g66590 (an allergen); genes encoding proteins related to arabinogalactan, pectin and cellulose/cell wall such as At5g48900 (a pectate lyase) and ATEXPA8 (At2g40610, an expansin); a gene

encoding a transcription factor At3g21330 (a HEC2-like transcription factor); genes encoding auxin-related functions like *IAA*29 (At4g32280) and a complete "*SAUR*-like" gene cluster comprising six members At5g180-c (At5g18010, At5g18020, At5g18030, At5g18050, At5g18060, At5g18080); and At5g12050 which encodes a protein of unknown function with similarity to helicases. Together with the microarray data these results confirmed the shift of physiologically relevant gene expression alterations between seedlings and growing leaves.

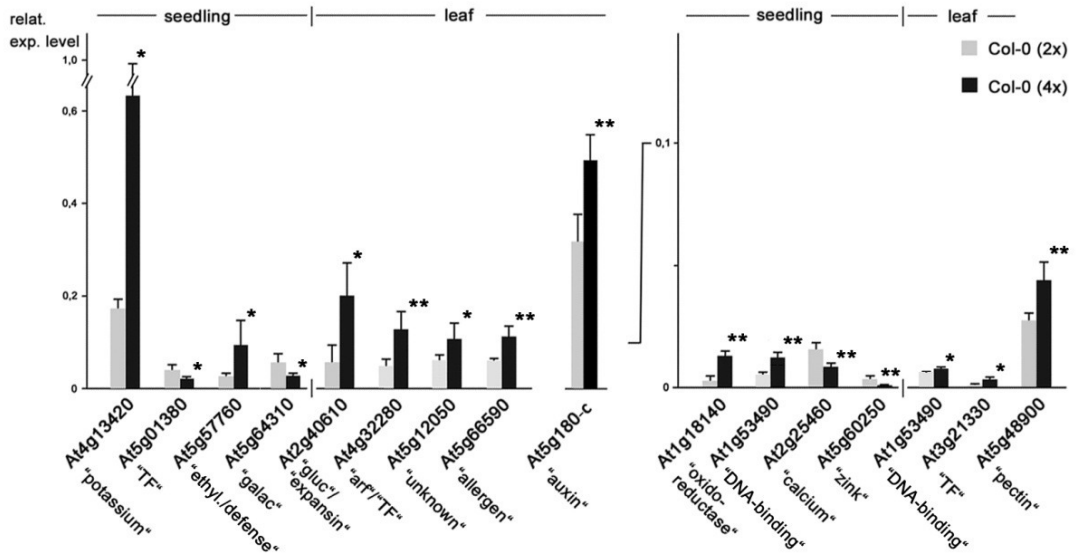


Fig. 14: Altered expression of selected genes in A. thaliana Col-0 autotetraploids as shown by qRT-PCR of genes representing disparate functional categories.
At5g180-c indicated qRT-PCR of a complete "*SAUR*-like" gene cluster comprising six members: At5g18010-30, At5g18050-60 and At5g18080. One tailed t-test: $p \leq 0.05$ (*), $p \leq 0.01$ (**), $p \leq 0.0005$ (***); bars with s.d.

3.2.5 Gene expression alterations in tetraploids of other *A. thaliana* ecotypes than Col-0 and L*er*-0

The extreme gene expression alteration difference between Col-0 and L*er*-0 indicated that the gene expression alteration by polyploidy might depend on the ecotype background. The qRT-PCR measurements of the expression of At4g32280 and At5g180-c cluster in leaves of seven other diploid and tetraploid ecotypes respectively (Bor-1, Bur-0, Ct-1, L*er*-1, Nd-1, Pro-0, Ts-1, see 3.1.3) were performed (Fig. 15). They showed the variability of the gene expression alteration by polyploidy among different ecotypes. The absolute expressions of these genes were higher in ecotype Pro-0 compared to the other ecotypes and were significantly higher in the tetraploid Pro-0 than in the diploid Pro-0. At4g32280 was significantly up-regulated in tetraploid Bor-1, tetraploid Bur-0, tetraploid Nd-1 and tetraploid Pro-0 compared to their diploids. At5g180-c cluster was significantly up-regulated in tetraploid Bor-1,

tetraploid Nd-1 and tetraploid Pro-0, and was down-regulated in tetraploid Ts-1 compared to their diploids.

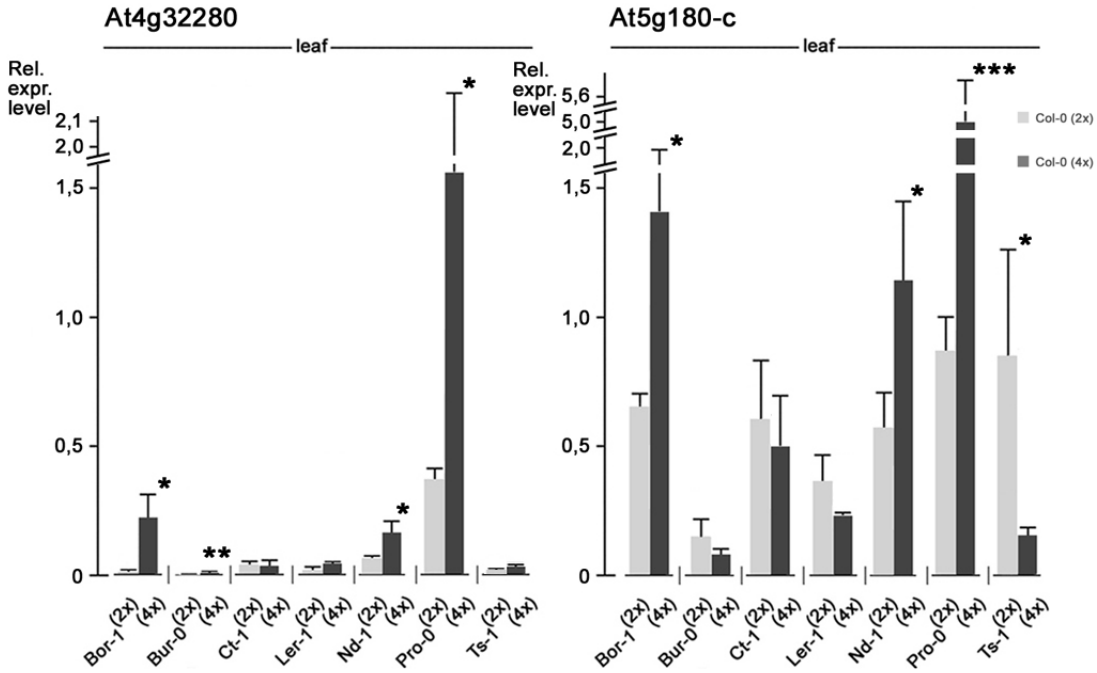


Fig. 15: qRT-PCR of At4g32280 and At5g180-c in different diploid vs. tetraploid ecotypes.

Left: relative expression of At4g32280; right: relative expression of the "*SAUR*-like" gene cluster At5g 180-c, which represented six members At5g18010-30, At5g18050-60 and At5g18080 amplified with an universal primer pair which amplifies all copies, ploidies and tissure are indicated. One tailed t-test: $p \leq 0.05$ (*), $p \leq 0.01$ (**), $p \leq 0.0005$ (***); bars with s.d.

3.3 *MRD1*, a valuable marker for monitoring polyploidy in *Arabidopsis thaliana* Col-0

3.3.1 *MRD1* up-regulated in Col-0 autopolyploid tissues

MRD1 was first detected as being down-regulated in the mutant *mto1-1*, in which there was an over-accumulation of soluble methionine (up to 40 fold; Goto and Naito, 2002). *MRD1* was conspicuous because it ranked almost in all differential gene expression experiments in the top group of altered genes in Col-0 autopolyploids (Appendix Table 3, 4, 5, 6, 9, 10, 11, 12). The qRT-PCR results confirmed that *MRD1* displayed a weak basic expression in diploid Col-0, diploid L*er*-0 and was significantly up-regulated in Col-0 tetraploid seedlings, leaves and also in triploid leaves but not altered in L*er*-0 tetraploid lines (Fig. 16 a). *MRD1* had been found to be (weakly) expressed throughout the adult plant development i.e., in seedlings, young rosette leaves, old rosettes, and siliques (Goto and Naito, 2002).

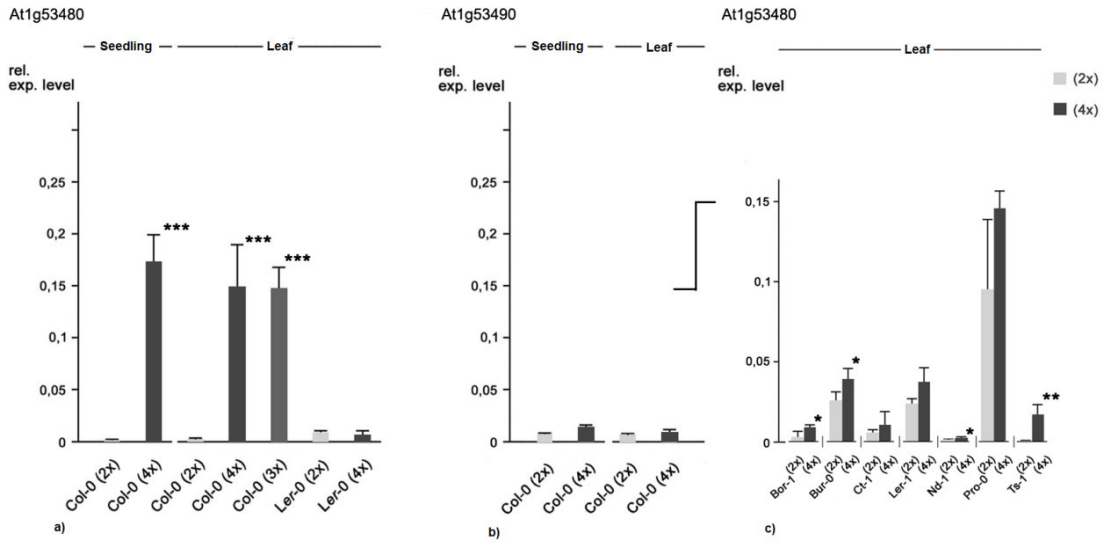


Fig. 16: Real time qRT-PCR analyses of *MRD1* (At1g53480) and At1g53490 in different backgrounds.
a) Analysis of *MRD1* (At1g53480) in different ploidies of Col-0 and L*er*-0 background; b) Analysis of At1g53490 in different ploidies of Col-0 background in seedling and leaf tissue; c) Assessment of altered *MRD1* expression in di- and tetraploid leaves of seven additional ecotypes.

MRD1 (At1g53480) is closely neighbored to another gene At1g53490 (Fig. 18), which also appeared as over-expressed in seedling and leaf tissue but to a much lower level (Fig. 16b). In fact, the TAIR annotation displays these genes as natural anti-sense pair and they appear to be species-specific loci of unknown function, because truncated copies were only found in *A. lyrata* among all plant sequence compilations. However, if At1g53490 interferes with *MRD1* transcription (and vice versa), it obviously does not appreciably affect the tetraploid *MRD1* expression as shown by qRT-PCR (Fig. 16) and northern blot analysis (Fig. 17). The same amounts of mRNA from diploid and tetraploid Col-0 leaves and L*er*-0 leaves were applied in the northern blot experiment. Obviously, At1g53480 (*MRD1*) band was detected in the Col-0 tetraploid leaves, no obvious band was detected in other lines. This indicated that *MRD1* is up-regulated in tetraploid Col-0 lines. For At1g53490, there was no obvious band in all the lines, which indicated low expression in these lines (Fig. 17). The size of the detected *MRD1* band in the northern blot did not fit with the indicated splicing in the current annotation but was consistent with the result of the northern blot in the work of Goto and Naito (2002). Available T-DNA insertions of both genes did not exhibit conspicuous morphological phenotypes. All plants tested (progeny of 6-8 siblings per insertion line) produced normal viable seedlings. The analysis identified this transcription unit is more easily activated in Col-0 polyploid. *MRD1* expression was also analyzed in seven other ecotypes: Bor-1, Bur-0, Ct-1, Ler-1, Nd-1, Pro-0, Ts-1 (Fig. 16c). It was significantly up-regulated in tetraploid lines of Bor-1, Bur-0,

Nd-1, Ts-1. Especially in Ts-1, it was 20 fold up-regulated. The expression of *MRD1* in other four ecotypes was not significantly different between diploid and tetraploid lines. The variable absolute and relative expression differences between different ecotypes again demonstrated the impact of natural variation for this gene expression alteration in autopolyploids.

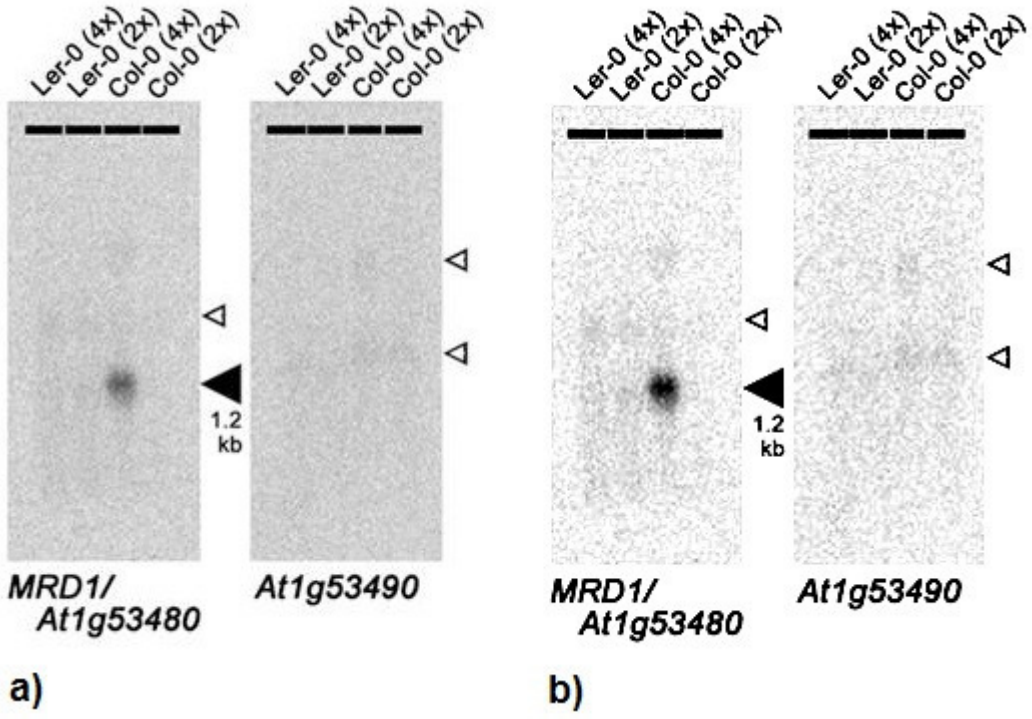


Fig. 17: Northern blots with At1g53480 and At1g53490 probes:
Isolated mRNA from diploid and tetraploid L*er*-0 and Col-0 respectively was used and hybridized with probes as indicated in Fig. 18. Size of approximate transcript length is shown in kb. Open arrow heads indicate weak bands probably including homologous gene copies. a) the original photo; b) the photo with adjusted contract and brightness.

The Genomic region of At1g53480 (*MRD1*) and At1g53490 differed between L*er*-0 and Col-0 in several nucleotide sites (Fig. 18). There were 9 nucleotide sites difference between L*er*-0 and Col-0 in the genomic region of *MRD1* and only 4 nucleotide sites difference in the region of At1g53490 (shown in red, Fig. 18).

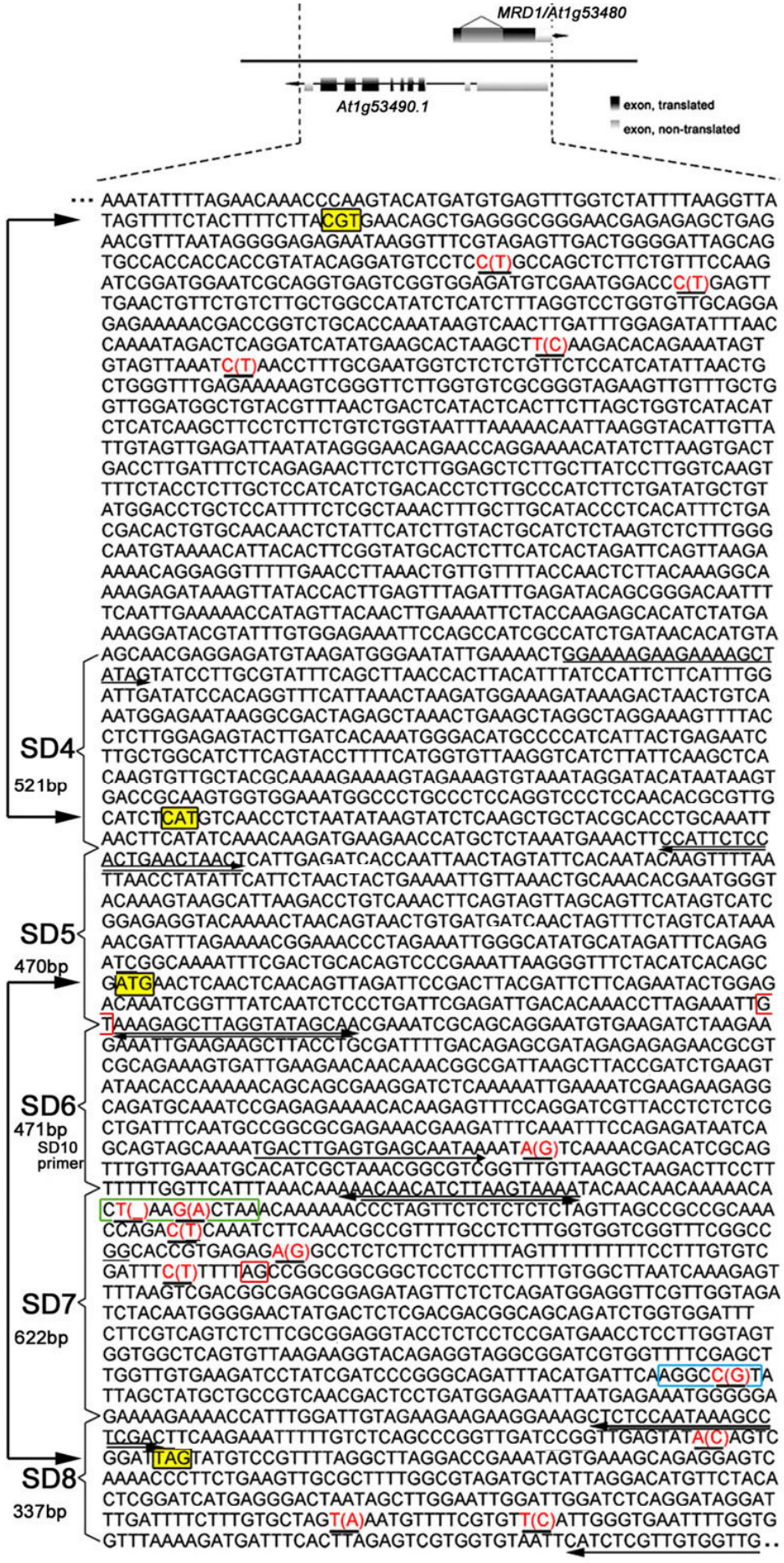

MRD1/At1g53480
At1g53490.1
exon, translated
exon, non-translated
SD4
521bp
SD5
470bp
SD6
471bp
SD10
primer
SD7
622bp
SD8
337bp

Fig. 18: Sequence polymorphisms between L*er*-0 vs. Col-0 in the genomic region of At1g53480 and At1g53490 (from Yu et al., 2010).
Initiation and stop codons are (yellow) colored and boxed. Polymorphisms are underlined and in red (Col-0 in parentheses). Linked arrowheads on the right indicate the probes used for Northern blots (from ATG to STOP codon). Regions SD4 to SD8, their sizes given, indicated on the left, were analyzed with respect to DNA methylation (primers used for amplification are indicated by arrows beneath the sequence). The triangular lines above *MRD1*/At1g53480 indicate an intron annotation in TAIR (splice sites boxed in red) not found in this study.

3.3.2 Methylation anaylsis of *MRD1* genomic region in Col-0, L*er*-0 diploid and tetraploid lines

The enzyme McrBC was used to verify the methylation state of the genomic regions of *MRD1* in Col-0, L*er*-0 diploid and tetraploid lines (see Mat. and Meth. and Fig. 19). McrBC only digests methylated DNA. The methylation state of the DNA could therefore be shown by the result of the PCR amplification after McrBC digestion, i.e. failure of the PCR amplification indicated that the DNA was methylated. The result indicates that the genomic region covering SD7 and SD8 (shown in Fig. 18) of *MRD1* was highly methylated in diploid Col-0 and L*er*-0 leaves (absence of bands in the samples with McrBC digestion in Fig. 19). In L*er*-0 tetraploid leaves this methylation state was unchanged, however, in Col-0 tetraploid leaves, this region was highly demethylated (bands occurred in the samples with McrBC digestion). In the SD6 region (see Fig. 18) which was closer to the 5' end of *MRD1*, both L*er*-0 and Col-0 indicated low methylation in diploid as well as tetraploid lines. This indicated that the over-expression of this gene in Col-0 tetraploid is correlated with the complete demethylation of part of its transcribed region SD7 SD8 involving epigenetic mechanism. Note however, that partial demethylation of tetraploid L*er*-$0^{1026\text{-}41}$ does not lead to high over-expression. (Fig. 19)

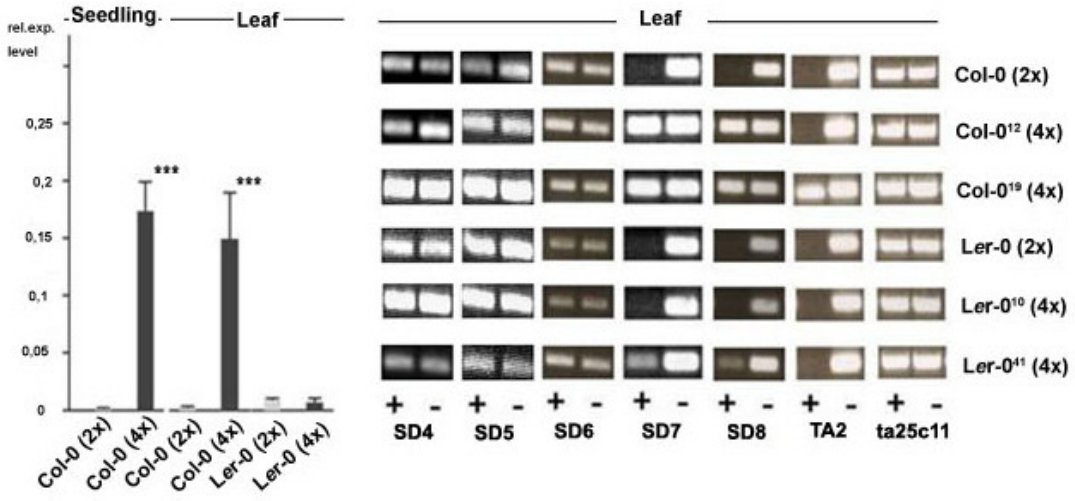


Fig. 19: Expression of *MRD1* (left) and the methylation state (right) of the SD4, SD5, SD6, SD7, SD8 regions (see Fig. 18) in *MRD1* in Col-0 and L*er*-0 diploid and tetraploid lines.

Left: shows the qRT-PCR results of *MRD1* in different ploidies of Col-0 and L*er*-0 background and right: shows the correspondent methylation state of *MRD1* in different background.This methylation was verified by digestion of genomic DNA by McrBC, followed by PCR amplification using primers specific for SD4, SD5, SD6, SD7 or SD8 region (shown in Fig. 18). Failure to amplify a product after digestion by McrBC indicated that this region was methylated. "+" indicates that the DNA used here was digested with McrBC, and "-" indicates that the DNA used here was not digested with McrBC. Control primers from an unmethylated region (ta25c11, Vaughn et. al., 2007) and a methylated transposon (TA2, Vaughn et. al., 2007) indicated complete digestion and amplification in each case except in Col-0^{19}, a small part of DNA here was not completely digested with McrBC.

3.3.3 Over-expression of *MRD1* in Col-0 tetraploid is inherited in Col-0 tetraploid L*er*-0 tetraploid hybrid

The striking difference between Col-0 and L*er*-0 tetraploids allowed to test for *MRD1* expression in tetraploid hybrids and thus for heritability of the tetraploid Col-0 response. Real time qRT-PCR showed that the expressions of *MRD1* in Col-0/L*er*-0 tetraploid hybrids were up-regulated too, compared to the diploids. This showed that the capability of sensing and responding to tetraploidy by Col-0 was transmitted to the hybrid. McrBC-Analysis demonstrated that this was accompanied by maintained demethylation in this locus at least in part of the gene copies (Fig. 20a, b). However, the methylated copies can still be present since they were already digested and could not be amplified by PCR, therefore the presence of them will not affect the results in Fig. 20 b.

RT-PCR amplified transcripts of the tetraploid hybridwere sequenced. The sequencing results showed that both *MRD1*^{Col} and *MRD1*^{Ler} alleles were transcribed at almost the same level (Fig. 20c; Fig. 21). The activation of the L*er*-0 *MRD1* allele in tetraploid hybrids was also confirmed by the analysis of cDNA clones. To this end, transcripts from tetraploid hybrids were reverse transcribed and amplified by *MRD1* specific primer pairs (SD10F, SD7R see material and methods). The resulting

sequences were cloned and sequenced. Around 11 clones showed L*er*-0 specific sequence and 14 clones showed Col-0 specific sequence. Both the clone sequence results and the sequencing profiles indicated that the L*er*-0 allele is significantly expressed (as well as the Col-0 allele) in the tetraploid hybrids.

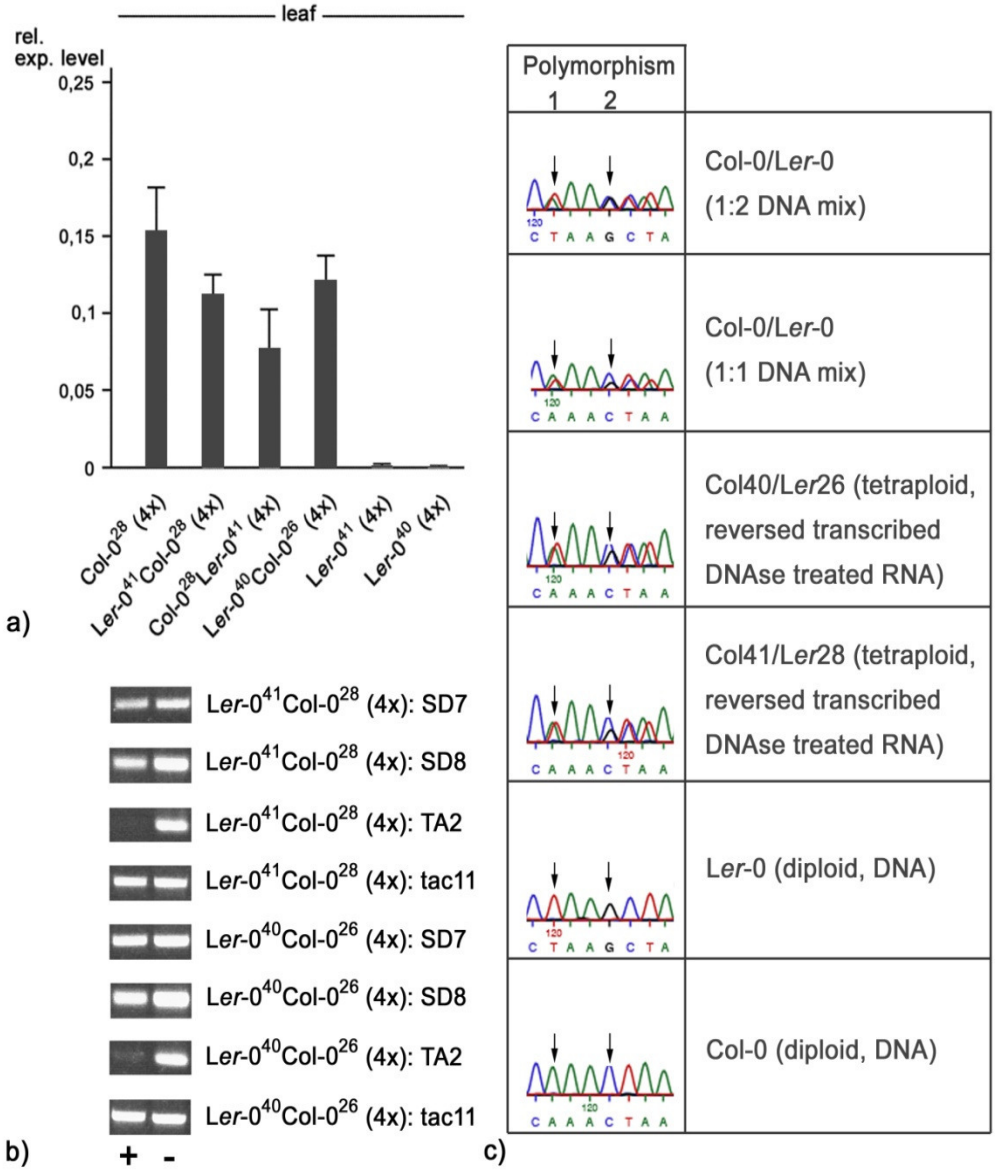


Fig. 20: Inheritance of *MRD1*/At1g53480 over-expression in tetraploid *Arabidopsis thaliana* F1-hybrids.

a) Real time qRT-PCR in different hybrid tetraploid Col-0/L*er*-0 combinations (leaf tissue). b) Methylation analysis of At1g53480 segments (as in Fig. 18; leaf tissue). "+" indicates with McrBC digestion, "-" indicates without McrBC digestion. The transposons TA2 and tac11 were used as positive and negative control respectively. c) Sequence reactions of RT-PCR amplified transcripts identifying *MRD1*^{Col} and *MRD1*^{Ler} alleles in tetraploid Col-0/L*er*-0-hybrids. Reactions of mixtures, tetraploid hybrids and diploids are given. Arrows point to positions of sequence polymorphisms in *MRD1*/At1g53480. The 2nd polymorphism is caused by a deletion. Index numbers refer to the tetraploid lines as in a). This sequence is the green-boxed sequence in Fig. 18.

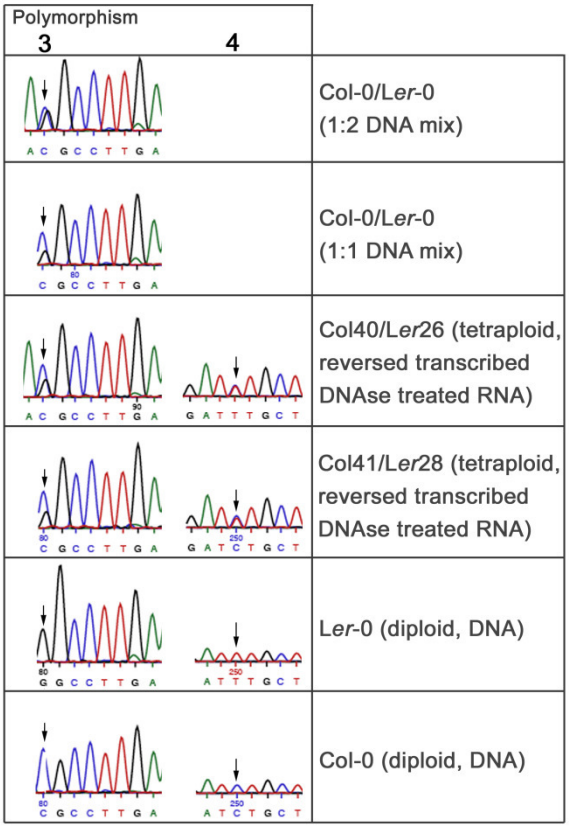


Fig. 21: Inheritance of At1g53480 over-expression in tetraploid *Arabidopsis thaliana* hybrids.
Identification of At1g53480^{Col} and At1g53480^{Ler} alleles in tetraploid Col-0/L*er*-0-hybrids by sequencing RT-PCR amplified transcripts. Given are sequence reactions of different ecotype mixtures, tetraploid Col-0/L*er*-0 hybrids (index numbers refer to hybrids in Fig. 20) and diploid L*er*-0 and Col-0 plants. Arrows point to positions of a third and a fourth sequence polymorphism in the putative coding region of the gene. The left sequence is the reverse sequence in the blue box in Fig. 18.

3.4 Gene expression analysis of triploid lines

3.4.1 Microarray analysis

Microarrays were performed to compare the transcriptome pattern difference in leaves between Col-0 diploid lines and Col-0 triploid lines with eight biological replicates. The Col-0 triploid lines were generated by crossing Col-0 diploid lines with Col-0 tetraploid lines. The assessment of the ploidy level with flow cytometry of the used triploid lines is shown in Fig. 22.

The comparison produced significant differences between diploid and triploid Col-0 leaves. However, the number of differentially expressed genes was much lower than in the comparison between diploid and tetraploid Col-0 leaves. Thus, At3g25640, At2g37870, At2g38530, At3g62820, At1g64360, At5g51170, At5g61890, At4g39675, and At3g61490 showed strong up-regulation (FC>1E+04, $p<0.05$) in triploid leaves compared to the diploid leaves. These genes were unchanged in tetraploid leaves

(Table11). However, apart from these, very few additional differences (only four more genes, $p \leq 0.05$) were found (Table11). Only three of these genes, At1g06350, At2g28870 and At1g53480 could be found in the di- vs. tetraploid differential group.

Table 11 Differential expressed genes between Col-0 diploid and Col-0 triploid leaves

Gene	description	Fold change	P value < 0.05
AT1G53480.1	Unknown, *MRD1*	4,33E+04	2,814E-07
AT3G25640.1	unknown protein	1,33E+04	2,060E-03
AT2G37870.1	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein	1,58E+04	7,978E-03
AT2G38530.1	LTP2 (LIPID TRANSFER PROTEIN 2); lipid binding	2,44E+04	1,145E-02
AT5G44720.2	molybdenum cofactor sulfurase family protein	7,29E-02	2,442E-02
AT3G62820.1	invertase/pectin methylesterase inhibitor family protein	1,44E+04	2,850E-02
AT2G28870.1	unknown protein	6,86E-02	2,896E-02
AT1G06350.1	fatty acid desaturase family protein	7,85E-02	3,082E-02
AT1G64360.1	unknown protein	1,21E+04	3,132E-02
AT5G51170	unknown protein	1,01E+04	4,237E-02
AT5G61890.1	AP2 domain-containing transcription factor family protein	1,62E+04	4,529E-02
AT4G39675.1	unknown protein	1,26E+04	4,529E-02
AT3G61490	glycoside hydrolase family 28 protein / polygalacturonase (pectinase) family protein	1,23E+04	4,720E-02

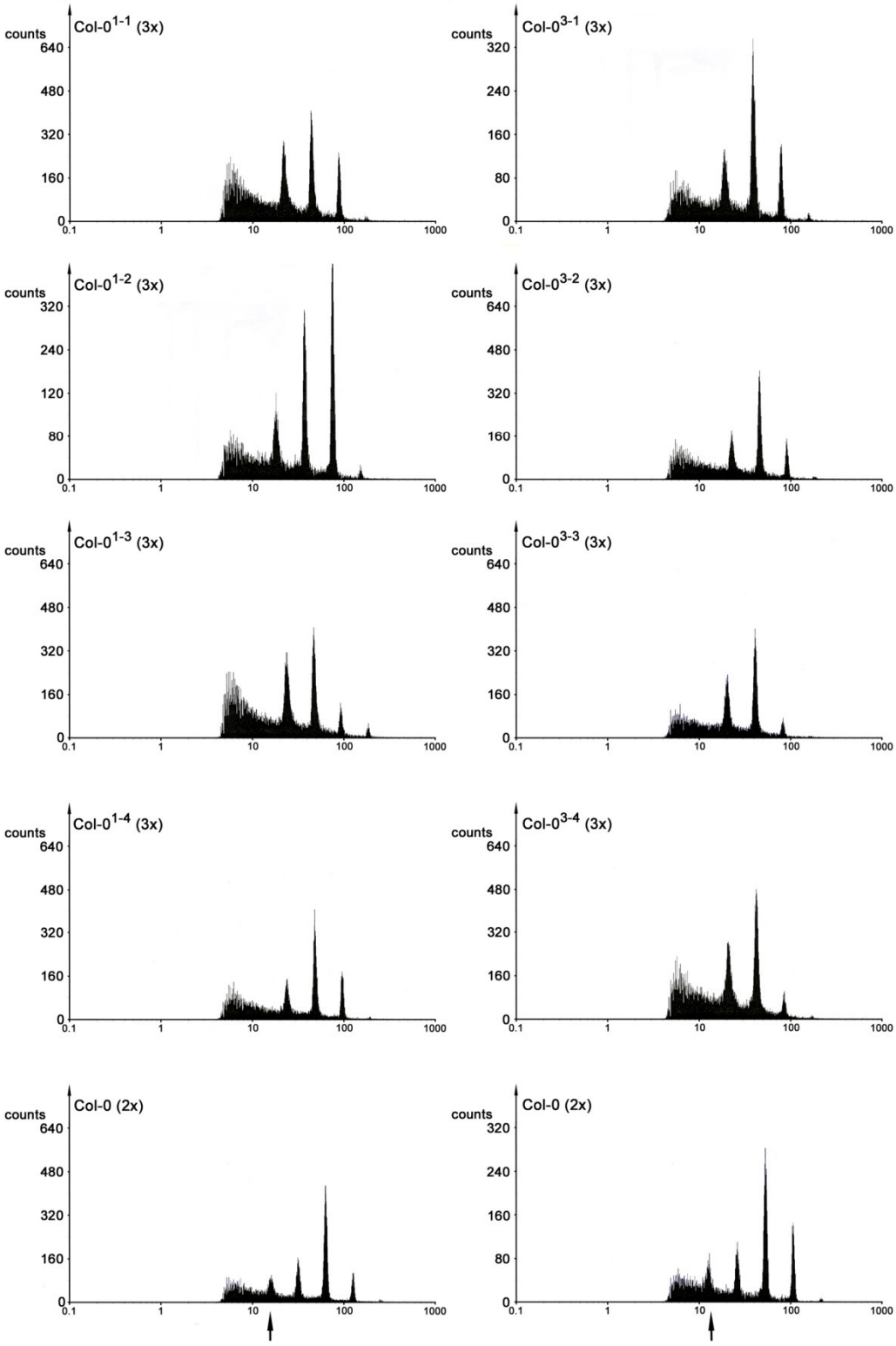


Fig. 22: Flow cytometry analysis of the Col-0 triploid lines
Diploid lines (2x) and triploid lines (3x) are indicated. Arrows point to the 2C peak in diploids.The first peak (3C peak) in triploid lines occurs in the position between 2C and 4C peaks in diploid lines.

3.4.2 Expression of *MRD1* in triploid lines

The microarray results indicated that *MRD1* was over-expressed in Col-0 triploid lines. Real time qRT-PCR (Fig. 23) confirmed this result. Interestingly, the triploid hybrid lines resulting from the crossing of Col-0 tetraploid with L*er*-0 diploid showed highly over-expression of this gene as well (Fig. 24a, the strong bands), but the triploid lines from crossing of L*er*-0 tetraploid with Col-0 diploid showed no over-expression (Fig. 24a, only very weak basal bands). Correspondingly, the DNA in the *MRD1* over-expressing triploid hybrid lines showed demethylation, while the DNA from the other triploid lines (from crossings between Col-0 or L*er*-0 diploid with L*er*-0 tetraploid) still exhibited more methylation (Fig. 24b). For details of McrBC experiments see 3.3.2. The conditions of the reaction with enzyme McrBC are critical. In some experiments the DNA was not digested completely by the enzyme. This can be controlled by using TA2, which is methylated (Vaughn et. al., 2007). In Fig. 24b the weak At1g53480-SD7 bands in the McrBC digested DNA samples were due to the incomplete digestion since there are also corresponding weak TA2 bands in the digested samples.

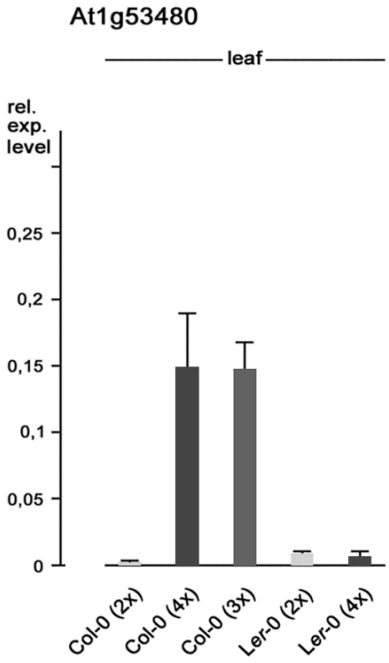


Fig. 23: Expression of *MRD1* in different polyploidy level leaves

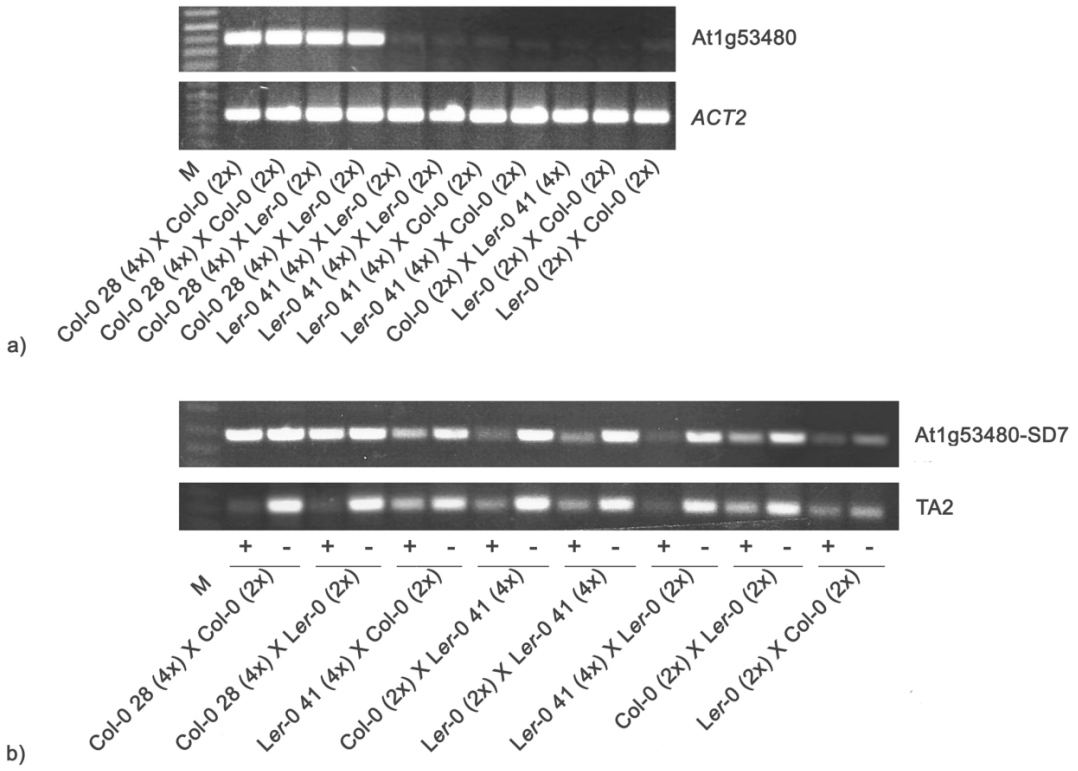


Fig. 24: Expression and methylation state of *MRD1* in different triploid hybrids.
a) *MRD1* is up-regulated in the triploid hybrids originating from the cross between Col-0 tetraploid and Col-0 or L*er*-0 diploid. *b) MRD1* genomic SD7 region (see Fig. 18) is strongly demethylated in the triploid hybrids which over-express *MRD1*. The mehylation profile was analyzed with McrBC enzyme digestion followed by PCR amplification see Fig. 19 and Material and Methods. "+" indicates the DNA used here was digested with McrBC, and "-" indicates that the DNA used here was without the digestion of McrBC. TA2 again as control to see whether the DNA was digested by McrBC completely or not.

3.5 Physiologic effects of tetraploids (Amino acid analysis)

To see the physiologic consequence of the ploidy change, the free amino acid amount in the diploid and tetraploid Col-0 plants was measured and compared. Eighteen amino acids were compared between diploid and tetraploid leaves (see Material and Methods). Three of them showed significantly different contents in diploid vs. tetraploid leaves.

In tetraploid Col-0 leaves there were much more methionin and tryptophan (Fig. 25).

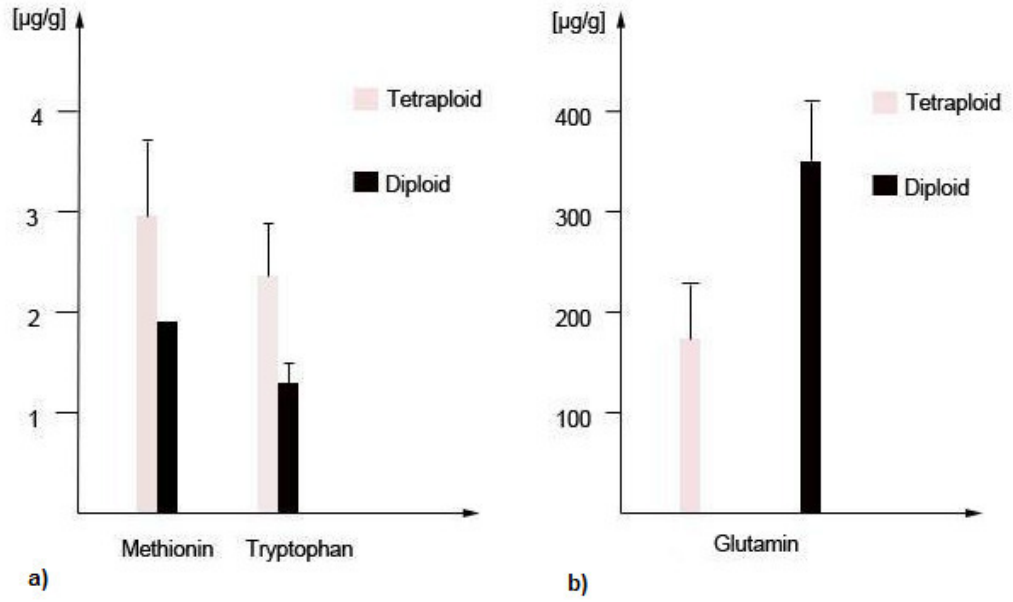


Fig. 25: Tetraploid leaves had more Methionin and Tryptophan and less Glutamine than diploid leaves.

There was also a significant difference between Col-0 diploid and Col-0 tetraploid leaves with respect to the Glutamine content. Tetraploid leaves had much less glutamine. Additional amino acids showed a different content in diploid and tetraploid leaves, however, the calculated content values among the biological replicates showed big variations, resulting in overlapping standard deviation of diploids vs. tetraploids. Additional experiments are needed to confirm these differences. All the amino acids contents measured in this experiment are shown in the Appendix Table 15.

3.6 Seed weight measurement of Recombinant Inbred Lines (RILs)

Seeds from the RILs were ordered from the stock center NASC. They were generated from a cross between the *Arabidopsis* ecotypes Columbia (Col-0) and Landsberg *erecta* (L*er*-0) (Alonso-Blanco et al., 1998) with Columbia as the male parent. The seeds from the heterozygous (F1) individuals were self-crossed over 8 generations to establish nearly isogenic lines. 30 lines with the highest frequency of recombination over the five chromosomes were converted into tetraploids by Kristina Haage (Haage, 2005) and 26 of them were confirmed as tetraploid lines with flow cytometry (Appendix Fig. 5). To find a polyploidy effect, different traits (trichome tip

numbers, size of the leave, size of the stem, ramification, flower, *erecta* phenotype) were scored for these 26 lines for both diploids and tetraploids.

Over 1000 markers data along the five chromosomes can be downloaded from the NASC website. These markers had been mapped using 100 RILs. The markers show whether a certain region of the chromosome from a RIL is of L*er* or Col background respectively. When several RILs have abnormal phenotypes and a certain region of the chromosome among these RILs have the common marker whereas the other RILs have the other marker then this region is linked to this phenotype. This is a very rough mapping. The aim of this effort is to establish a RI -tetraploid population, which allows the mapping of tetraploid affected traits (see Discussion). This population could help to find the possible linked gentic region to the abnormal phenotypes in tetraploid plants.

Most RI tetraploid lines showed normal diploid phenotype. However, some tetraploid plants of RI5 (11plants of 20 plants, 11/20), RI35 (2/10), RI190 (6/15), RI194 (4/10), RI302 (6/13), RI303 (3/10) and RI332 (4/15) showed strange flowers, which means that they had either more than four petals or more than four sepals or more than six stamens or sometimes one of the stamens fused with the carpel. Some tetraploid plants of RI4, RI13, RI115, RI191, RI231, RI232, RI 245, RI267, RI288, RI295, RI303, RI358 and RI370 showed also strange flowers but in rare cases. There was one region between the marker “LRRPK” and marker “CDS111”, where these RILs all showed solely Col-0 markers, other RILs with normal flowers showed solely L*er*-0 markers. This region may link to the strange flowers in the RI tetraploid plants. The flowers from the pure Col-0 tetraploids are normal. This indicates the strange flowers might be pleiotropically caused in combinantion by loci Col-0 and loci L*er*-0.

In addition, seed weight was scored. The goal was to find the difference between diploid seed weight and tetraploid seed weight in order to map possible QTLs (Quantitative trait locus). To this aim seed weight of both diploids and tetraploids of these 26 RILs were measured in 2007 and 2008 for two consecutive generations. The difference of the diploid seed weights between these two parental ecotypes (Col-0 0,020mg and L*er*-0 0,016mg) enabled a tentative QTL mapping with these original diploid RILs (see Mat. and Meth.). The tetraploid seed weights of these two ecotypes were also different (Col-0 0,030mg and L*er*-0 0,032mg).

The plants of the third generation of the 26 tetraploid RILs were used for seed weight measurement in 2007. The diploid RILs’ seed weight was also measured. Fig. 26a shows the results. Almost in all the tetraploid RILs, the seed weight was higher than the seed weight in the correspondent diploid RILs. There was one tetraploid RIL (RI231) which an exceptionally high seed weight. The seed weight of this line was

even higher than the seed weight from a hexaploid plant (see Fig. 26a), but lower than the seed weight from an octoploid plant. The seed weight from the pure Col-0 or L*er*-0 tetraploid and diploid lines are also shown in Fig. 26a. Note that the tetraploid L*er*-0 seed weight was higher than the tetraploid Col-0 seed weight. However, in diploids, the situation was reversed. QTL anaylsis was performed in collaboration with the Lehrstuhl für Pflanzen Züchtung (Prof. Schön). No QTL was found in the diploid RILs, and one QTL was found in the tetraploid RILs with LOD value (Logarithm of Odds, significance) of 2.01. Although, this QTL analysis was very rough, it indicated the possibility to perform the QTL mapping with the tetraploid RIL population.

In 2008, the fourth generation of the tetraploid RILs was used for the seed weight measurement and the diploid RILs were measured at the same time. Fig. 26 b shows the results. The seed weight from the tetraploid RILs was higher than the seed weight from the diploid RILs. The seed weight of tetraploid and diploid lines were in the same range as the seed weight value obtained in 2007. However, some individual tetraploid RILs showed quite different values compared to 2007. For example, RI231 tetraploid line which had extremely heavy seeds in the third generation showed normal seed weight like other tetraploid lines in the fourth generation (Fig. 26).

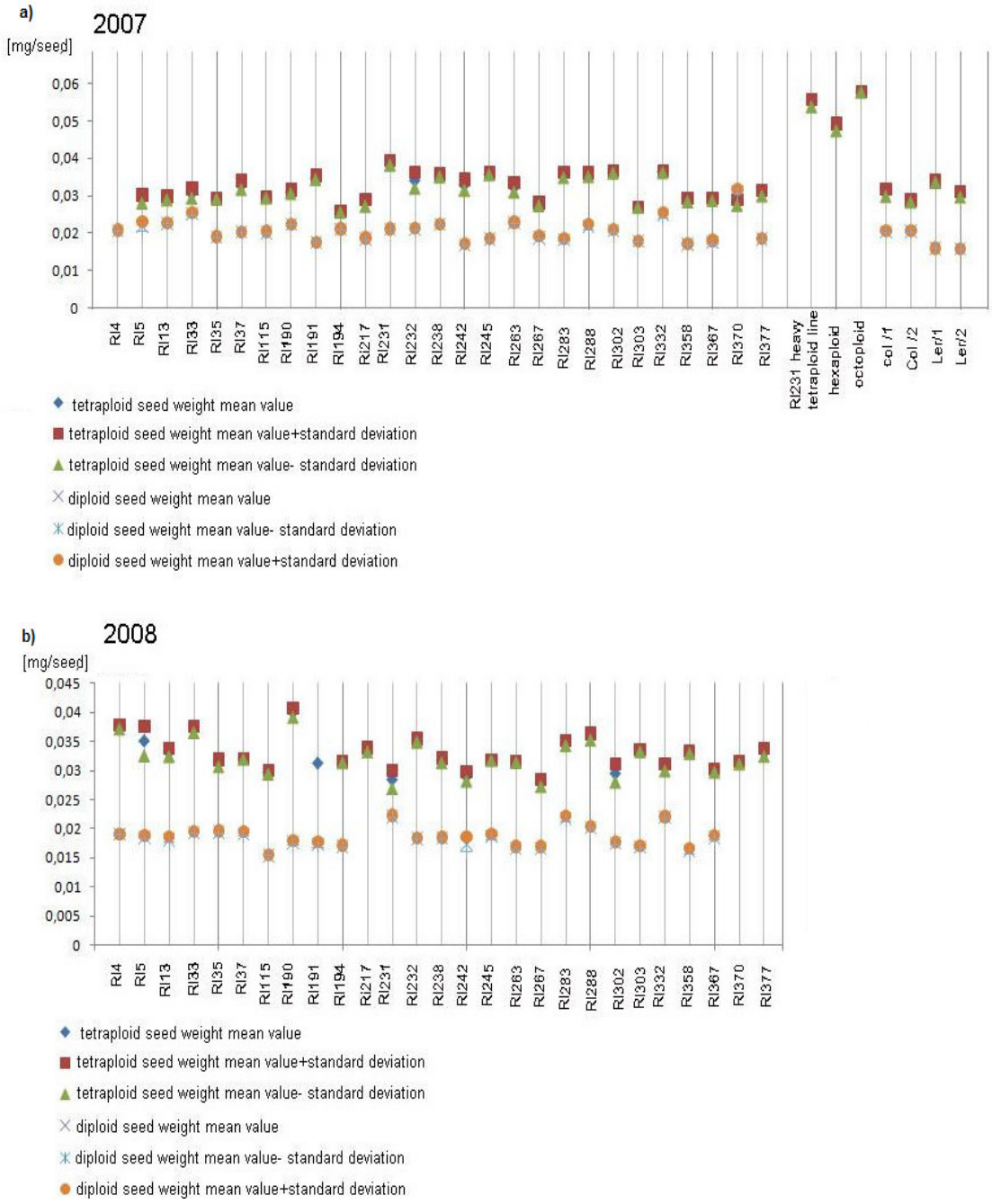


Fig. 26: Seed weight of diploid and tetraploid (the third generation after induction) RILs: Blue spots show the mean value of the seed weight of three measurements (in each measurement more 100 seeds were used). The pink and yellow spots show the mean value plus and minus standard deviation of the three measurements. On the right side, one extreme tetraploid line has seeds even heavier than the seeds from one hexaploid line. The corresponding plant has less seeds than the other plants. Seeds from one octoploid RIL and from pure Col-0 or L*er* diploid, tetraploid lines were also measured.

The number of the RILs used here is not sufficient to enable a reasonable QTL mapping, although the lines used were the most recombinant lines among the 100 RI RILs of the Lister and Dean population (Table 12, Lister and Dean, 1993). More RILs are needed to be converted into tetraploid for this attempt. Additional neo-tetraploid

RILs were generated and are included in table 12. Five of these lines were confirmed as tetraploids with flow cytometry analysis (Appendix Fig. 6).

Table 12: RILs arranged according to the recombinant rank

RI Line	Recombinant events	Recombinant rank	RI Line	Recombinant events	Recombinant rank
RI 190	207	1	RI 34	138	34
RI194	202	2	RI14	136	35
RI 295	199	3	RI167	134	36
RI284	197	4	RI84	134	37
RI238	197	5	RI33	134	38
RI191	197	6	RI113	133	39
RI4	197	7	RI25	133	40
RI302	196	8	RI30	132	41
RI245	195	9	RI182	131	42
RI232	192	10	RI123	129	43
RI263	190	11	RI62	129	44
RI303	187	12	RI266	128	45
RI370	186	13	RI397	127	46
RI288	183	14	RI237 *	127	47
RI180	181	15	RI79	125	48
RI35	181	16	RI54	125	49
RI332	180	17	RI131	123	50
RI 283	178	18	RI214	122	51
RI267	176	19	RI279	121	52
RI59	171	20	RI359	120	53
RI217	165	21	RI377	118	54
RI37	165	22	RI173	117	55
RI367	164	23	RI13	117	56
RI166	153	24	RI5	117	57
RI52	149	25	RI358	115	58

RI356	147	26	RI71	113	59
RI46 *	147	27	RI67	110	60
RI29 *	146	28	RI349	108	61
RI231	142	29	RI242	108	62
RI179	142	30	RI115	104	63
RI177 *	141	31			
RI36	139	32			
RI107 *	138	33			

The rank is calculated according to the number of recombinations in these RILs; lines in red are the 26 RILs used for seed weight measurements. Lines in blue are the newly generated tetraploid lines. Lines with * are the newly generated tetraploid lines confirmed with flow cytometry analysis.

4. Discussion

4.1 Colchicine effects and induction of polyploidy

Colchicine is a highly poisonous secondary metabolite from *Colchicum autumnale* (and other Colchiceae). Colchicine binds in an equimolar and poorly reversible manner to soluble nonpolymerized tubulin with high activation energy (Bhattacharyya et al., 2008), forming a tubulin-colchicine complex (Niel and Scherrmann 2006; Wilson et al., 1999; Ravelli et al., 2004). Heterodimers of α- and β-tubulin form dynamic polymers termed microtubules that can elongate and contract as filaments to change structure and function of the cytoskeleton, exemplified by the interphase microtubule network and the mitotic spindle (Wilson et al., 1999; Ravelli et al., 2004). Microtubules are involved not only in cell division, but also in signal transduction, regulation of gene expression, migration, and secretion (Wilson et al., 1999). At low concentrations, colchicine arrests growth of microtubules and, at higher concentrations, colchicine promotes depolymerization of the microtubules (Ravelli et al., 2004). Colchicine binds not only to tubulin but also to leukocyte membrane proteins that might provide sites for attachment of microtubules and consequent cytoskeletal reorganization (Borron et al., 1996).

Colchicine's toxicity is an extension of its mechanism of action - binding to tubulin and disrupting the microtubular network. As a result, affected cells experience impaired protein assembly, decreased endocytosis and exocytosis, altered cell morphology, decreased cellular motility, arrest of mitosis (Niel, 2006).

Whether colchicine has mutagenic effects in addition to its toxic effects in *Arabidopsis* is not clear. In the experiments performed in this work, none of the mutants that most frequently occur in mutagenic screens were detected (i.e. the albino, crème and fusca p henotypes; Jürgens et al., 1995; Misera et al., 1994). Similarly, an independent study of over 40 colchicine induced *Brassica* allopolyploids did not provide any indication of a mutagenic effect of colchicine (Lukens et al., 2006). There is also no evidence for long term genomic instabilities caused by colchicine. In fact, these could be caused by other factors, such as scaling incompatibilities of cell volume versus intracellular structures (e.g. spindle geometry; Storchova et al., 2006).

The most important aspect of our colchicine treatments pertains to colchicine effects on the epidermal (L1) and sub-epidermal (L2, L3) layers. Following treatment, the apex can develop in three different ways. First, it might produce detectable diploid and polyploid sectors. Second, diploid cells overgrow polyploid cells and most of the plant turns diploid again and produces diploid progeny. Conversely, polyploid cells

overgrow diploid cells; i. e. polyploid progenies are mainly produced. In any case, it is a crucial prerequisite of our strategy that the altered polyploidy of the epidermis (L1 layer) on the treated plant reflects an altered polyploidy of inner cell layers (L2 and L3 layer respectively) because these are the cells that give rise to generative cells. This is not self-evident because L1 and the inner layers (L2, L3) are developmentally separated from early on (Takada and Jürgens, 2007). In fact, we found that once a polyploid sector had been identified on a plant, the same plant regularly delivered polyploid progeny (Table1). This can be explained by two possible effects, which have been observed previously (Dawe and Freeling, 1991; Tilney-Basset, 1986 and references therein). The first is that colchicine intruded into epidermal and subepidermal layers in most of the treatments. The second is that cells from one tissue "invaded" the adjacent lineage.

The sensitivity to colchicine is ecotype dependent. Col-0 was the most "resistant" ecotype among the nine ecotypes (Col-0, Ler-0, Bor-1, Bur-0, Ct-1, Ler-1, Nd-1, Pro-0 and Ts-1). There can be two possible reasons for this. The first is that Col-0 ecotype may have better detoxification ability. In medication, colchicine has been used to treat gout for more than 2000 years, and pseudogout and Familial Mediterranean Fever (FMF) for several decades (Terkeltaub, 2008). Adenosine triphosphate-binding cassette subfamily B member 1 (ABCB1) transports colchicine out of cells. Increased transporting function results in the resistance to colchicine. Several single nucleotide polymorphisms of ABCB1 were identified with the potential to influence expression and quantitative transporter function, thus to influence the sensitivity to Colchicine (Sauna et al., 2007). The sensitivity difference to colchicine among different ecotypes may relate to the polymorphisms of ABCB1 region among these ecotypes. Secondly, Col-0 may have a better ability to degrade cyclin B and survive the C-mitosis. Kung et al. (1990) have shown that the ability of a cell type to survive C-mitosis after being treated with colchicines is positively correlated with its ability to degrade cyclin B during the prolonged mitotic period.

Polyploidy led to an increase in cell size. The trichome cells in polyploids were not only enlarged but also had an increase in the number of branches. This phenomenon made it easy to identify the polyploid sectors in plant which survived colchicine treatment. The trichome branch number in the leaves was also ecotype dependent. Different ecotypes displayed different distribution of trichomes with particular numbers. Therefore, it was important to take the diploid plants from the same ecotype as the reference when assessing the trichome morphology of the survived plants. When the number of the trichomes with more branches of a certain sector elevated significantly, compared to the diploid plants, this sector was most probably

polyploid. Using the trichome morphology as the marker, the first polyploid selection step increased in speed and efficiency.

After colchicine has suppressed mitosis by depolymerising the tubulin spindle apparatus in a cell of the shoot apex, a mosaic plant will grow. In this plant the root, hypocotyls, cotyledons and possibly the first primary leaves will be diploid and the following growing parts from the colchicine affected cell in the shoot apex will be polyploid (at least those parts which originate from the affected cell). In the same plant, some cells in the shoot apex were affected by colchicine and gave polyploid sectors, while others were not affected and remained giving diploid sectors. Therefore, the seeds of the induced plants with polyploid sectors, were of mixed ploidy grades. In contrast, tetraploid seeds from the induced plants, once confirmed with flow cytometry and chromosome counting mostly gave tetraploid progeny in the following generations. The ploidy level was assessed in selected lines for three generations. Aneuploidy occurred quite rare in these cases (according to the chromosome counting results). The hexa- or octoploids were not as stable. Aneuploidy occurred more often in these lines and they often segregated to lower ploidy level (see Yu et al., 2009).

4.2 Discussion about the overlaps of the transcriptome alteration lists

For the seedling transcriptome comparisons, there were 860 genes differentially expressed between Col-0 diploid and L*er*-0 diploid; 348 genes between Col-0 tetraploid and L*er*-0 tetraploid. 224 genes overlapped between these two lists (Fig. 27). The big overlapping here itself already suggested that autopolyploidisation would not give much significant gene expression alterations, and it also suggested that the difference (860 vs. 348) of altered expression between Col-0 and L*er*-0 as diploids and tetraploids respectively, might be less than suggested due to statistical evaluation at the chosen reference significance value ($p<0.05$).

There were 589 genes differentially expressed between Col-0 diploid seedlings and Col-0 tetraploid seedlings; 9 genes between L*er*-0 diploid seedlings and L*er*-0 tetraploid seedlings (Fig. 27). There was no overlap between these two lists. This indicated that the transcriptome alteration is ecotype dependent (see below).

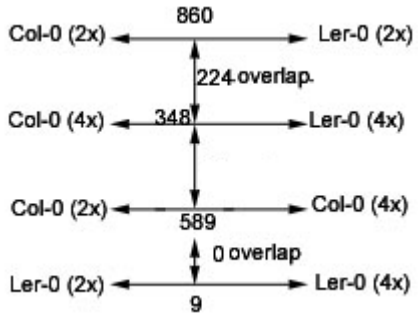


Fig. 27 Overlaps of the transcriptome alteration list in seedlings comparison. (see text)

For the leaf transcriptome comparisons, there were 27 genes differentially expressed between L*er*-0 diploid and L*er*-0 tetraploid; 289 genes between Col-0 diploid and Col-0 tetraploid; 15 genes between Col-0 diploid and Col-0 triploid. One gene (At5g60390, with oppositive alteration, encoding a calmodulin binding translation elongation factor) overlapped between the L*er*-0 diploid vs. tetraploid seedling comparison list and the correspondent leaf comparison list (Fig. 28a). 16 genes overlapped between the Col-0 diploid vs. tetraploid seedling comparison list and the correspondent leaf comparison list (Fig. 28b). Five of these 16 genes were conversely altered. At1g51820 (encoding a Leucin-rich repeat receptor kinase which is repressed by abscisic acid, Sanchez et al., 2004) and At1g22400 (encoding a UDP-glycosyltransferase) were up-regulated in tetraploid seedlings and down-regulated in tetraploid leaves; At1g04180 (encoding YUCCA9, a flavin monooxygenase involved in auxin biosynthesis ,Sairanen et al., 2012), At4g13260 (encoding YUCCA2, also a flavin monooxygenase involved in auxin biosynthesis, Zhao et al., 2001) and At1g06350 (enconding a fatty acid desaturase family protein) were down-regulated in tetraploid seedlings and up-regulated in tetraploid leaves. At1g06350 was also up-regulated in triploid leaves. The low overlapping number and the oppositive alteration direction indicate that the alteration is developmentally specific or tissue specific (see below). Six of the 16 overlapping genes were signaling pathway involved or signal responsive genes. They were At1g51820, At3g01490, At5g02760, At4g30270, At3g20820 and At4g30650. At1g51820 encodes a Leucin-rich repeat receptor kinase, which was repressed by abscisic acid, Sanchez et al., 2004 and was activated by Nep1 (necrosis and ethylene inducing peptide 1) like proteins (Qutob et al., 2006). At3g01490 also encodes a kinase and the relative phosphorylation is significantly increased after sucrose treatment (Nittyla et al., 2007). At5g02760 encodes a protein phosphatase and is a cold responsive signaling element (Lee et al., 2005). At4g30270 encodes a xyloglucan transferase, functions in the cell wall modification (Campbell and Braam, 1999), and responses to

brassinosteroids (Kauschmann et al., 1996). At3g20820 encodes a leucine rich repeat family protein which functions in cell wall signal transduction (Kobe and Kajava, 2001). At4g30650 encodes a low temperature / salt responsive protein and is up-regulated by abscisic acid (Sanchez et al., 2004). Among these six genes, only At4g30270 was down-regulated in both tetraploid seedlings and leaves. And the others except At1g51820 were all up-regulated in both tetraploid seedlings and leaves. Blanc and Wolfe have observed that those genes involved in signal transduction and transcription had been preferentially retained after the ancient polyploidy events (Blanc and Wolfe, 2004). This indicated that the alteration of these signaling pathway involved genes might important for the evolution of the polyploids and relate to giving the better chance for polyploids for surviving.

There were 88 genes differentially expressed between L*er*-0 diploid and tetraploid leaves in the microarray work of Wang (Wang et al., 2006a). There is no overlap between these 88 genes and the 27 genes in the data here. This maybe because the tetraploid plants used in this work were the plants induced with the colchicine, but the tetraploid plants Wang used were plants obtained by regeneration of roots and had a transgenic kanamycin resistance gene from a T-DNA (Luca Comai et al., 2000).

There were three genes overlapping between the Col-0 diploid-tetraploid comparison list and the diploid-triploid comparison list (Fig. 28b). These three genes include At1g53480 (*MRD1*). *MRD1*'s function, if any, remains unclear (Yu et al., 2010). Interestingly, truncated copies were only found in *A. lyrata* among all plant sequence compilations (Yu et al., 2010). The other two genes were At1g06350 (encodes fatty acid desaturase family protein, oxidoreductase function) and At2g28870 (encodes unknown protein, preferentially expressed in guard cells in comparison with mesophyll cells, Leonhardt et al., 2004). At1g06350 and At2g28870 were only slightly up-regulated while At1g53480 was highly up-regulated in Col-0 polyploid leaves. The low overlapping indicated that the transcriptome alteration differed a lot when the chromosome number became odd. This triploid case will be disscussed later seperately. Among these three genes, two were also altered between Col-0 diploid and tetraploid seedlings, At1g06350 (slightly down-regulated in tetraploid seedlings) and At1g53480 (highly up-regulated in tetraploid seedlings).

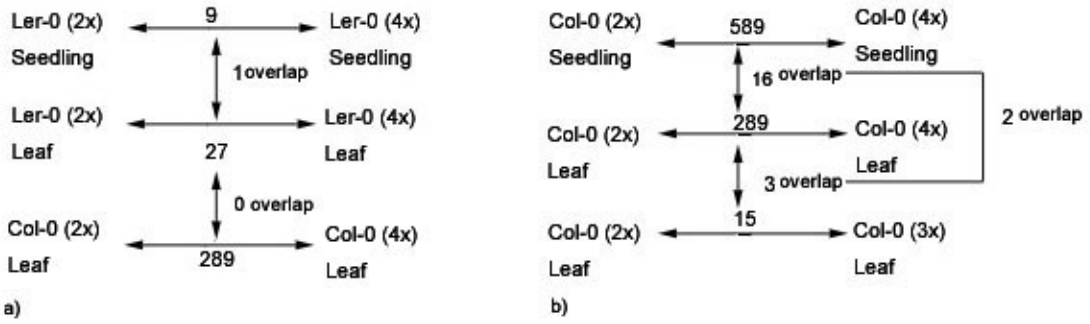


Fig. 28 Overlaps of the transcriptome alteration lists among seedling comparison and leaf comparison with respect to Col-0 and L*er*-0 (for details see text).

4.3 Alteration of transcriptome in *A. thaliana* autotetraploids depends on ecotype, i.e., genome composition.

It was generally expected that the uniform genomes of autopolyploids, in contrast to those of allopolyploids, should not exhibit significant gene expression alterations. This observation was supported by limited analysis. In the work of Wang, as a control experiment for *Arabidopsis* allopolyploid, L*er*-0 autotetraploid leaves only showed 88 differentially expressed genes in comparison with diploid leaves (Wang et al., 2006a). In potato leaves and root tips, subtle gene expression change (1.4%) was found between diploid and tetraploids (Stupar et al., 2007). In *citrus* leaves, the differential expression genes between diploid and autotetraploid were less than 1% and the maximum rate of gene expression change within a 2-fold range (Allario et al., 2011). In cabbage, proteomic analysis indicated similar few changes between autotetraploids and diploids (Albertin et al., 2005). In *Paulownia fortunei*, only 6.09% transcripts were significantly differentially expressed between diploid and autotetraploid (Zhang et al., 2014). The presented data here also shows subtle gene expression change between autotetraploid and diploid for the two ecotypes and two tissues. Although Col-0 altered around five hundred genes in two tissues, these five hundred genes had relatively low fold change (most within a 2-fold range), and anyway five hundred genes only made very minor ratio from 26,000 genes.

Compared to L*er*-0, which displayed an almost diploid expression profile, Col-0 showed relatively more gene expression change. This significant ecotype specific difference in gene expression alterations was also showed in other ecotypes. Limited analysis of other ecotypes with selected probes supported the notion that the response to tetraploidy was variable and depended on the genomic composition. In the Bor-1 and Nd-1 ecotypes, all three genes were up-regulated, whereas in other ecotypes, only two of the genes were altered in their expression. Ct-1 and L*er*-1 did

not show any response to all three genes. Whether this observation indicated variable degrees of response capability has to be further investigated.

There may be two reasons that the transcriptome alteration is ecotype dependent. The first might be the polymorphisms among the ecotypes; the second might be the epigenetic variability among the ecotypes. Col and L*er* gene methylation was highly divergent, consistent with the evolutionary distance between these two ecotypes. The other ecotypes also have high variable gene methylation pattern (Vaughn et al., 2007). It is not known whether such epigenetic modifications are perfectly passed over from di- to tetraploidy and equally to all four gene copies. A recent work suggested that this may not be the case and such modifications can be unstable over generations. DNA methylation in specific regions of the genome can fluctuate over relatively short timescales in *Arabidopsis*. Such sites can be considered as going through recurrent cycles of forward and reverse epimutation, which is very different from what is found at the level of the genome sequence, where reverse mutations are exceedingly rare (Becker et al., 2011). It should be mentioned that the ecotype specific gene expression alterations shown in this study were also clearly distinct from aneuploid syndromes (BirchLer et al., 2005; Huettel et al., 2008; Henry et al., 2007), because they occurred in *A. thaliana* autotetraploids, i.e., balanced euploids. In contrast, aneuploidy is an out-of-balance situation leading to extensive gene expression alterations in *Arabidopsis* (Huettel et al., 2008) and segregation distortion of loci such as *Arabidopsis SENSITIVE TO DOSAGE IMBALANCE* (*SDI*) (Henry et al., 2007).

4.4 Transcriptome alterations in autotetraploid *Arabidopsis* are developmentally specific.

The presented data suggested that gene expression alterations depended on the developmental stage. This was reflected by the low overlap (16 genes) between altered seedling and leaf transcriptomes and by different representation of GO groups (see Yu et al., 2010). Apparently, the Col-0 response was a general alteration or relaxation of gene expression control, which consequently covers genes of different stages and/or it allows an easier accessibility of genes for expression alteration at their corresponding stages of activity. The functional gene groups displayed by seedlings and leaves are well known from these tissues. Seedlings displayed a biphasic mode of ethylene related gene activity (Etheridgeet al., 2005) whereas any form of leaf organogenesis was tightly linked to localized auxin accumulation and auxin driven gene activities (Benkova et al., 2003; Treml et al.,

2005). Interestingly, neo-allopolyploid *A. suecica* also revealed a conspicuous alteration of ethylene/stress related genes (Wang et al., 2006a) showing partly similar reactions in both forms of polyploidy. However, they also revealed different gene expression alterations not observed in autopolyploids such as those considering heat shock genes. Since allopolyploid effects on gene regulation may be caused by genome doubling and intergenomic interactions. These heat shock genes may come from the combination of these two effects. 68% of non-additively expressed genes of allotetraploid (hybridization between tetraploids of *A.thaliana* and *A. arenosa*) found in Wang's work were differentially expressed between *A.thaliana* and *A. arenosa* these two parent species. 33 heat shock genes displayed expression differences from the midparent value in allopolyploid. 31 out of these 33 genes were highly expressed in *A. thaliana* compared to *A. arenosa.* Therefore, the difference between these two parent species leads to more gene expression alterations in allopolyploids compared to autopolyploids.

A recent work reported few genes to be different between diploids and autotetraploids when individual lines were analyzed. The comparison of the three paired lineages detected only minor differences when using the less stringent Benjamini and Hochberg FDR procedure. Using the 'per gene' variance assumption only At2g32210 (which is also present in the differential expressed gene list of this work, see Appendix Table 5) was detected as up-regulated (Pignatta et al., 2010). However, the material and the standards of analysis are different if compared to this work. First, the whole plant was taken for the RNA used in microarray in Pignatta' work. While in this work, specific tissues (leaves or seedlings) were taken for the RNA used in microarray analysis. Second, the plant used in Pignatta' work was older (4 weeks' old) than the plant used here (two weeks' old). Four week old plant already has some flowers, such tissue can have quite different gene expression alteration and this can buffer the leaf gene expression alterations. The tissue dependent gene alteration may be a reason that Pignatta et al. (2010) found much less altered genes. Notably, only three independent microarrays were analyzed. However, in this work, twelve microarrays' data were combined. Thus, the statistical power is very different between these two works. In fact, our initial analyses had also revealed that, with a limited comparison of only four independent replicates, the detected number of gene expression alteration found with twelve replicates would have been extremely reduced. In fact, more recent work shows that there are stable alterations in gene expression between diploids and tetraploids. Li and coworkers found a subtle but stable increase in the expression of *ICK (Cyclin dependent Kinase Inhibitor)* genes in tetraploid seedlings (Li et al., 2012), although only small number of selected genes

were analyzed. Among those genes, *ICK4* and *ICK2* also appeared in the analyses presented here (see Appendix Table 9). Del Pozo and Ramirez-Parra found 471 genes (very close to the number presented in this work) with altered expression between the leaves of tetraploid versus diploid plants. In plants subjected to drought even more (a total of 1360) genes showed different levels of expression (Del Pozo and Ramirez-Parra, 2014). In both works the specific tissues and stages instead of whole and old plants were analyzed. Together, these findings also support the results presented in this work.

In newly formed cotton allopolyploids, Adams et al. (2003) also found that alleles from the two hybridized genomes (designated A and D) differ in expression patterns among the tissues examined for 11 out of the 18 genes considered. In the most extreme case, *adhA* was entirely expressed from the A gene in carpels but from the D gene in petals and stamen.

4.5 The transcriptome alteration in triploids

The comparison of the response to tri- vs. tetraploidy indicated an increased sensitivity with higher gene doses in the latter because the gene expression alteration in Col-0 triploid leaf tissue is conspicuously low. This is partly similar to maize (Guo et al., 1996) but dissimilar to monoploid potato (Stupar et al., 2007). The fact that all except three genes are not included in the list of tetraploids is reminiscent to the B chromosome odd number effects (Guo et al., 1996; Jones and Rees, 1982).

Darlington and Upcott (1941) found that maize plants with odd numbers of Bs had more chiasmata than did plants with even B numbers. Jones and Rees (1982) explicitly named it the odd-even effect after finding that the between-cell variance in the number of chiasmata in rye plants with odd number of Bs was significantly higher than that observed in plants with even numbers of B chromosomes. A similar odd-even effect has been detected, in a number of plants and animals, for traits such as protein and RNA amounts, dry nuclear mass, exophenotypic characters and fitness related traits (e.g. fertility) (for review, see Jones and Rees, 1982). In general, an odd number of B chromosomes were more detrimental than even numbers. No clear explanation is available for the odd-even effect. There were two possible causes for this. One is that the equational division and lagging of the extra chromosomes, when univalents (odd Bs), could favour the appearance of abnormal meiotic products, and the formation of bivalents (even Bs), when there are two or more extra chromosomes, inhibits this process. The other cause is an increase in cell stress generated by odd numbers (Camacho et al., 2004).

However, the third copy of the normal chromosome is different from the B chromosome. The effect of triploidy is only similar to such odd even effect because in the microarray results here, nine genes show strong up-regulation in triploid leaves compared to diploid leaves and they were unchanged in tetraploid leaves. However, triploid is not more detrimental than even number ploidy. Triploids of *A. thaliana* are fertile, producing a swarm of different aneuploids (Henry et al., 2005). A recent work on human oocytes reprogramming somatic cells to a pluripotent state showed that reprogramming human cells using occytes is more feasible when the resultant cell is in the triploid state (without removing the oocyte genome) rather than in the diploid state (with removing the oocyte genome). The development of the latter one arrests at late cleavage stages in association with transcriptional abnormalities while the triploid cells develop to the blastocyst stage (Noggle et al., 2011).

The triploid plants were obtained by crossing the diploid with tetraploid plants. The alteration of the expression between diploid and triploid leaves may also result from the interaction between the one chromosome from the diploid and the two chromosomes from the tetraploid, since the epigenetic pattern (Yu et al., 2010) and the organization of the nucleus (more chromosomes are included) were changed when diploids were converted to tetraploids.

4.6 The *Arabidopsis thaliana* transcriptome alteration response to tetraploidy has a genetic basis and displays epigenetic phenomena.

The comparison of Col-0 vs. L*er*-0 tetraploids clearly showed that the transcriptome alteration response does not depend on the chromosome number per se but on the origin of the chromosomes. Furthermore, the alteration was completely transmitted through selfing to the next generation. The alteration of the expression of the gene *MRD1* in different ploidy levels especially showed this. The over-expression of *MRD1* was transmitted to tetraploid hybrids (Col-0 L*er*-0 tetraploid hybrids) and triploid hybrids (Col-0 tetraploid crossed with Col-0 diploid or L*er*-0 diploid) as well. Notably, in these cases, only two chromosome sets originate from the “responsive” Col-0 ecotype. Taken together, this result suggests that Col-0 but not L*er*-0 possesses one or more genetic factors that are capable of sensing the alteration of genome dosage and inducing gene expression alterations. Also, the analysis of other ecotypes shows that this ability depends at least partly on the genotype.

Possibly, the absence of *MRD1* over-expression in some tetraploids is due to mutation or the polymorphisms between these ecotypes. It is known that L*er*-0 originates from X-irradiated parents (NW20; TAIR). Many polymorphisms between

Col-0 and L*er*-0 are C to T transitions resulting in the CC composition alteration (see Fig. 16 in results). In both the CHH and CHG contexts, a cytosine immediately followed by another cytosine (CC context) has a significantly lower tendency to be methylated than a cytosine neighboring an adenine or thymine (Lister et al., 2008). Therefore, the slight polymorphisms in this region will result in methylation pattern difference and may also result in the demethylation capacity difference. Moreover, it is known that diploid Col-0 and L*er*-0 genomes possess variable DNA methylation patterns (Vaughn et al., 2007, in this work the fourth chromosome was tested). Although this natural epigenetic variability seems not to cause significant gene expression differences in diploids (Vaughn et al., 2007), it is not known whether this variability could contribute as such at the tetraploid level. At this point of discussion, it seems necessary to differentiate between sensing vs. induction vs. transmission/preservation. Although the sensing factors are not known, what they could sense can be speculated. Altered nuclear surface to volume ratios in tetraploids have been discussed as causative for gene expression regulatory changes (Comai, 2005; Misteli, 2007). Polyploids generally show increased nuclei, this implies an altered nuclear surface to volume ratio. According to the relationship between volume and surface of a sphere, doubling the genome is expected to double the volume that is occupied by chromatin, but cause only a 1.6-fold increase in the nuclear envelope surface. This difference, although apparently modest, can change the stoichiometry of the interaction between components of chromatin that are located at the nuclear periphery and envelope-bound proteins (Comai, 2005). Transcription is a very complex nuclear process. The transcription machinery is surprisingly dynamic and significantly determined by stochastic events (Misteli, 2007). These properties are based on the highly transient interaction of proteins with chromatin. Molecular crowding facilitates such stochastic interactions and makes them more efficient (Minton, 2000). Whether the tetraploid nuclear formation could change the molecular crowding degree of some nuclear molecules in some positions is unknown. The possible alteration of these two factors (altered nuclear surface to volume ratio and possible altered molecule crowding degree) could be sensed and induce the alteration of other factors like the small RNA biogenesis, epigenetic changing. These can further cause an alteration of the transcription of the gene.

The gene expression alteration of *MRD1* in various Col-0 vs. L*er*-0 tetraploids, Col-0/L*er*-0 tetraploid hybrids and Col-0/L*er*-0 triploid hybrids are strongly correlated with DNA (de)methylation. Several analyses of selected (trans) genes have demonstrated changes in gene expression between plants with altered ploidy grade (Riddle et al., 2010; Adams et al., 2004; Comai et al., 2000; Mittelsten et al., 2003; Wang et al.,

2004) and some of these have also uncovered a link to epigenetic phenomena, in particular DNA (de)methylation. Upon sensing a higher chromosome number in a nucleus, the induction of DNA methylation of selected genes could be caused by RNA directed DNA methylation (RdDM), which has been recently discovered in *Arabidopsis* (Teixeira et al., 2009; Zheng et al., 2008).

Small interfering RNAs (siRNAs) direct the cytosine methylation of DNA sequences that are complementary to the siRNAs. In plants, double stranded-RNAs (dsRNAs) generated by RNA-dependent RNA polymerase 2 (RDR2) serve as precursors for Dicer-like 3 dependent biogenesis of 24 nt siRNAs (Xie et al., 2004). Plant specific RNA polymerase IV (Pol IV) is presumed to generate the initial RNA transcripts that are substrates for RDR2 (Pikaard et al., 2008; Mosher et al., 2008). siRNAs are loaded onto an *Argonaute4* or *Argonaute6*-containing RISC (RNA induced silencing complex) that targets the de novo DNA methyltransferase DRM2 (Domain Rearranged Methyltransferase 2) to RdDM target loci (Wassenegger et al., 1994; Law and Jacobsen, 2010; Matzke et al, 2007). *DNA methyltransferase1* (*MET1*), which is a CG specific maintenance methyltransferase, also plays a role in de novo methylation (Aufsatz et al., 2004). Nascent RNA transcripts from the target loci are generated by another plant-specific RNA polymerase, Pol V, and these transcripts help recruit complementary siRNAs and the associated RdDM effector complex to the target loci in a transcription-coupled DNA methylation process (Mosher et al., 2008). Small RNA binding proteins such as ROS3 may direct target-specific DNA demethylation by the ROS1 family of DNA demethylases (Zheng et al., 2008). Chromatin remodeling enzymes and histone modifying enzymes also participate in DNA methylation and possibly demethylation (Chinnusamy and Zhu, 2009). Therefore, methylation and demethylation both are dynamic processes and can be mediated by small RNAs. CHH methylation cannot be sustained by the maintenance methyltransferases. It must occur de novo every time there is DNA replication, and guidance by siRNAs ensures preservation of specific CHHmethylation patterns in daughter cells (Zhang and Zhu, 2012).

The methylome project of diploid Col-0 has shown that *MRD1* is “body-methylated” and not “promoter-methylated” (Zhang et al., 2006). In addition, data from the Massively Parallel Signature Sequencing (MPSS) project indicated an accumulation of small RNAs in particular for the 3′ region of the gene (Lu et al., 2005; TAIR9 GBrowse). A part of these small RNAs are related to *Argonaute1*, which is a protein important for microRNA pathway and a large excess of this protein can interfere with the function of RISC (Vaucheret et al., 2004). Lister and coworkers found a similar correlation to the results presented here between methylation level and the

transcription level of *MRD1*. They found that the transcription of *MRD1* is highly up-regulated in *met1* mutants (deficient in CG maintenance DNA methylation, Saze et al., 2003) compared to the wild type Col-0 diploid. There is nearly no methylation of this gene in the mutant and the small RNA accumulation is also absent in the mutant but high in the wild type (Lister et al., 2008). So the high transcription in this gene is related to a low methylation level and a low small RNA accumulation. Interestingly, in the mutant *rdd* (ros 1-3 dm/2-1 dm/3-1 triple mutant, Penterman et al., 2007), where nearly all DNA demethylation is eliminated, the transcription level is similar as in the wild type but the small RNA accumulation is less if compared to wild type (Lister et al., 2008), this may indicate that for the demethylation of this region also a small RNA demanded. In this study, it is only clear that the low expression of this gene is accompanied with a high methylation level and small RNA accumulation in the diploid Col-0 plants. Whether the tetraploid Col-0 plants have less (or even none, like in the *met1* mutant) small RNA accumulation is unknown. It will also be interesting to test whether the small RNAs in L*er*-0 are the same ones as in Col-0 diploid. There can be a part of small RNA which is different between Col-0 and L*er*-0, namely those responsible for demethylation.

Basically, the study of *MRD1* indicates one epigenetic option for maintaining the observed transcriptome alterations. However, the observed alterations should not be assigned to DNA methylation alone. Epigenetic effects can be based on other DNA modifications. Furthermore, alteration of the DNA methylation pattern of one transcription factor/repressor could be sufficient to alter the expression of other genes without any further change of their methylation. Based on the sequence data of reversed transcribed *MRD1*-RNA, it is tempting to speculate that L*er*-0 *MRD1* displays higher transcriptional activation in the tetraploid hybrids and in the triploid hybrids than in diploids. Especially in triploids, the L*er*-0 *MRD1* originates from the diploid L*er*-0 not from the tetraploid L*er*-0 but is still activated upon “contacts” with the two copies from the tetraploid Col-0 lines. This activation could happen post fertilization unlike transcriptional reactivation of transposons in pollen (Slotkin et al., 2009). Alternatively, this gene could be activated during gametogenesis and then silenced upon fertilization. Then this silencing would be suppressed in tetraploid Col-0 and F1 Col-0/L*er*-0 hybrids because of the presence of chromosomes originating from tetraploid Col-0. The final effect resembles the opposite of paramutation of loci such as maize B-I (Chandler and Alleman, 2008). However, it is also possible that the dosage of a suppressor not present in Col-0, or the small RNAs (coming from the L*er*-0 side) corresponding for methylation maintenance are diluted in the hybrids. This observation is also complicated by the fact that a considerable part of L*er*-0

MRD1 is strongly methylated in the hybrids (Yu et al., 2010). In addition, there is always a basal level of *MRD1* transcription in the tissues tested regardless of the ploidy level. However, the existence of strongly methylated L*er*-0 *MRD1* in hybrids does not indicate the nonexistence of demehylated L*er*-0 *MRD1*. The HpaII and MspI digestion experiment in Yu et al. (2010) showed the methylated *MRD1* in the tetraploid hybrids were from L*er*-0 side, however this experiment is not able to detect the demethylated L*er*-0 *MRD1* in hybrids if it exists. It could be that the small RNAs normally accumulated in the L*er*-0 *MRD1* region are diluted by two chromosome copies from tetraploid Col-0, and not all the L*er*-0 *MRD1*'s methylation state can be maintened. In the meantime a part of the L*er*-0 *MRD1* could contact the small RNAs specific for the Col-0 demethylation. Thus, demethylation of a part of the L*er*-0 *MRD1* could occur and then might result in the up-regulation of the expression of L*er*-0 *MRD1* in hybrids. This demethylation might occur on the both *MRD1* L*er*-0 copies but only in a part of cells as shown in Fig. 29a. Alternatively, the demethylation might occur in one or the other of the two copies as shown in Fig. 29b. It could also be a mixture of the cases in Fig. 29a and b. The methylated *MRD1* L*er*-0 would give the observed result in the HpaII/MspI digestion experiment (Yu et al., 2010) and the demethylated *MRD1* L*er*-0 would be responsible for the up-regulated expression of *MRD1* L*er*-0 in hybrids detected in the sequencing results. This speculation is based on the factors existing and becoming active in Col-0 tetraploid and inducing the decrease of the small RNAs for the methylation maintenance and producing the small RNAs for demethylation. Similarly, the triploids coming from the crossing between tetraploid L*er*-0 and diploid Col-0 are not able to over-express *MRD1*, because the factors in diploid Col-0 are not active and the methylation state of this gene is maintained.

The speculated control mechanisms for *MRD1* expression are schown in Fig. 30. However, the reality may be far more complex. Whether there is also methylated *MRD1* L*er*-0 which is over-expressed in triploid hybrids needs to be tested and it remains to be determined whether paramutation-like phenomena are involved.

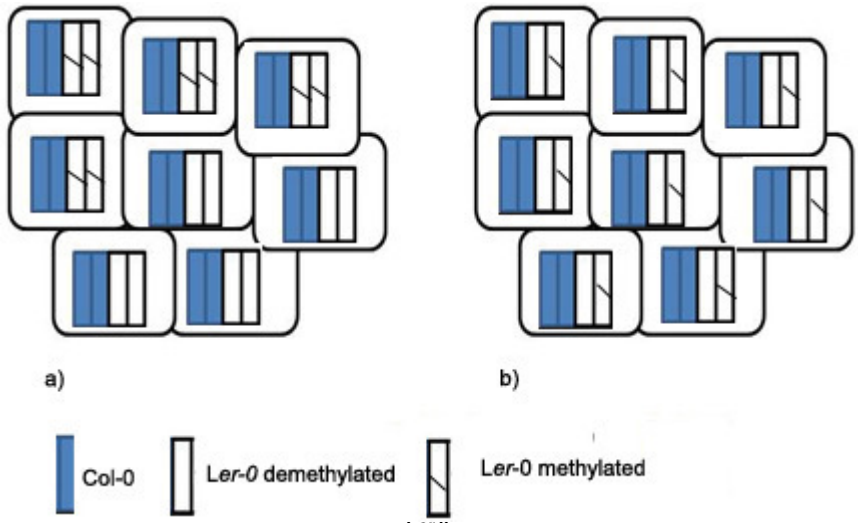


Fig. 29 Demethylation of *MRD1*$^{Ler\text{-}0}$ region models in tetraploid hybrids or in triploid hybrids.

a): A part of the cells harboring the demethylated *MRD1* on both L*er*-0 Chromosomes, the other cells harboring the methylated *MRD1* region on both L*er*-0 Chromosomes; b) Only one of the L*er*-0 copies is demethylated in the *MRD1* region.

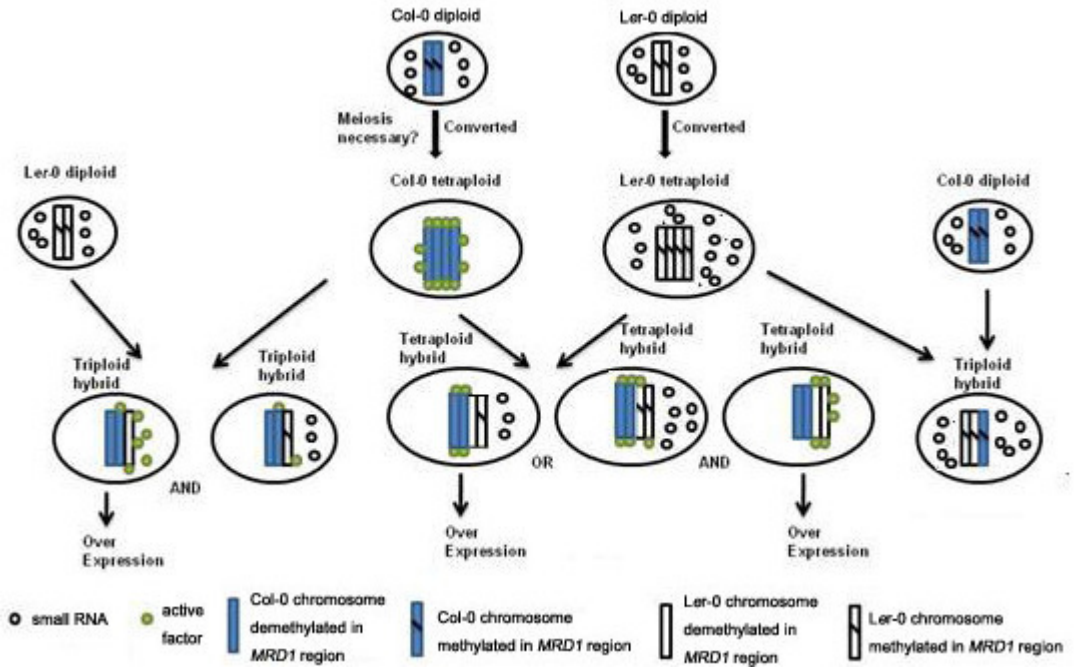


Fig. 30 Possible mechanism for *MRD1* expression in tetraploid hybrids and triploid hybrids

Small RNAs accumulate in the *MRD1* region in diploid plants. They correspond to the maintenance of the methylation state of this region. When converted to tetraploids, the active factors in Col-0 plants can sense the higher chromosome number and reduce small RNA synthesis and reduce methylation. In L*er*-0 such factors do not exist or are not active, so there is no response when converted into tetraploids. At this point, it is not clear whether it is necessary to pass meiosis to activate these factors. The activated factors and the small RNAs are inherited to the tetraploid hybrid and triploid hybrid but may be sometimes not evenly. The cells with less activated factors and more small RNAs can still maintain the methylation state. The cells with enough activated factors and diluted small RNAs (or even no small RNAs) can demethylate the L*er*-0 chromosome in *MRD1* region and over-express the gene. This may also explain why the expression of *MRD1* is lower in tetraploid hybrids than pure Col-0 tetraploid. The triploid hybrids from crossing between Col-0 diploid and L*er*-0 tetraploid can't demethylate the chromosomes in the *MRD1* region because the small RNAs inherited from both sites are sufficient for the methylation maintenance and no active factors are inherited.

4.7 Physiological effects of tetraploids

The presented results of this part (the amino acid content measurement), do require additional tests. The variation of the free amino acid content among the tetraploid individuals was sometimes considerable. More critically controlled growing conditions need to be applied so that all plants receive the same intensity of light (from all directions) and get the same amount of water. Alternatively plant culturing could be performed with a suitable randomization protocol. Since even limited bias of the growing conditions could influence the results considerably, it is likely that more differences will be found in Col-0 and in L*er*-0 plants as well, when the critical controlled conditions are applied and more biological replicates are used.

The increased amount of methionin in Col-0 tetraploid compared to Col-0 diploid and the over-expression of *MRD1* in Col-0 tetraploid make the story more complex, since *MRD1* was detected as been down-regulated in the mutant *mto1-1*, in which there was over accumulation of soluble methionine (Goto et al., 2002). This further points to the complex control of the expression of *MRD1* and might also indicate a complex relationship between the expression of *MRD1* and the free methionine amount. Perhaps the expression of *MRD1* needs to be beyond a certain level to result in higher free methionine content, otherwise the down-regulation of *MRD1* will increase the free methionine content.

The elevated amount of tryptophan in Col-0 tetraploid may relate to the over-expression of the auxin synthesis genes in tetraploids (see Table 9 in results chapter). Many auxin responsive genes are also up-regulated in tetraploids leaves. Tryptophan is an important precursor for auxin (Glawischnig et al., 2000). The over-expression of the auxin responsive genes indicates the elevated demand for auxin synthesis, therefore more auxin precursor is needed. It will be interesting to test whether there is higher auxin level in tetraploid Col-0 plants.

While writing this thesis, the metabolic content from seedlings was analyzed and compared between diploid and tetraploid by a master student, Mehmet Can Gülersönmez who discovered significant metabolic content difference between tetraploid seedlings and diploid seedlings (Gülersönmez, 2011). Tetraploids exhibited a lower level of amino acids, especially arginine, asparagine, glutamine, lysine, ornithine and putrescine were found in lower level in tetraploid seedlings. Glyceric acid, sucrose, xylose and gamma-tocopherol were found in higher levels in tetraploid seedlings. It should be noted, that seedlings are the more reliable material for such tests compared to leaves. It is easy to control and to get the same condition for each

sample. Tetraploid seedlings develop slower than the diploid seedlings (Li et al., 2012). This may be a reason for a reduced amino acid content in tetraploid seedlings. An induced tetraploid *Arabidopsis* line was found more tolerant to boron deficiency (Kasajima et al., 2010). Boron is an essential nutrient for plant growth and reproduction. The root elongation (both main root and lateral root) is much better in autotetraploids compared to in diploids under the boron deficient condition. Since Kasajima and his coworkers used Col-1 and Col-7 to test, it will be interesting to test whether other tetraploid ecotypes also have such tolerance and whether they are more tolerant to other environmental stresses compared to their diploids. There are many other tetraploid plants found to be more tolerant to certain stresses compared to their diploids. Induced tetraploid *Dioscorea zingiberensis* plants showed more heat resistance (Zhang et al., 2010). Autotetraploid turnip (*Brassica rapa L.*) exhibited a better adaptation to salinity stress (Meng et al., 2011). However, tetraploid citrus plants were more sensitive to salt stress than diploid plants under sufficient water supply (Mouhaya et al., 2010), but more tolerant to salt stress together under a water deficit situation (Saleh et al., 2008; Allario et al., 2011). Autotetraploid *A. thaliana* lines of several genetic backgrounds have higher leaf potassium content and better salinity tolerance than otherwise genetically identical diploids (Chao et al., 2013). All these physiological effects could be linked to the altered transcription, since many genes linked to stress senescence and ions were altered in the autotetraploids analyzed in the presented study (see Fig. 12).

A recent work found that autotetraploid *Arabidopsis* plants were more sensitive to glucose treatment than diploid with decreased number of rosette leaves and suppressed root elongation (Li et al., 2012). The authors found a subtle but stable increase in the expression of *ICK* genes in tetraploids seedlings. Among those genes, *ICK4* and *ICK2* also appeared in this analyses presented here. These *ICK* genes are likely to be involved in the regulation of growth-rhythm in tetraploid and diploid *Arabidopsis*, i.e. the vegetative growth period which produces the rosette leaf numbers (Li et al., 2012).

4.8 The usage of the induced tetraploid RILs

The intention of the RI tetraploid population establishment and trait screening was to test whether tetraploid-affected traits could be mapped by this strategy. There are over thousand markers such as single nucleotide polymorphisms (SNPs), simple sequence repeats (SSRs,or microsatellites), restriction fragment length polymorphisms (RFLPs), and transposable element positions, along the five

chromosomes for the 30 RI diploid lines (when L*er*-0 and Col-0 are compared). A map has been generated according to these markers so that the distribution of the L*er*-0 and Col-0 segments along the chromosome in the RILs is known (Reiter et al., 1992). The data is available at http://arabidopsis.info/new_ri_map.html. The strategy would be as follows: if there is a trait found common among several RI tetraploid lines and different from the corresponding diploid lines and other RI tetraploids and if there is a region where the marker shows the common origin (all L*er*-0 like or all Col-0 like) of these lines and the reverse origin in the other lines, then this region is linked to this tetraploid-affected trait (Fig. 31).

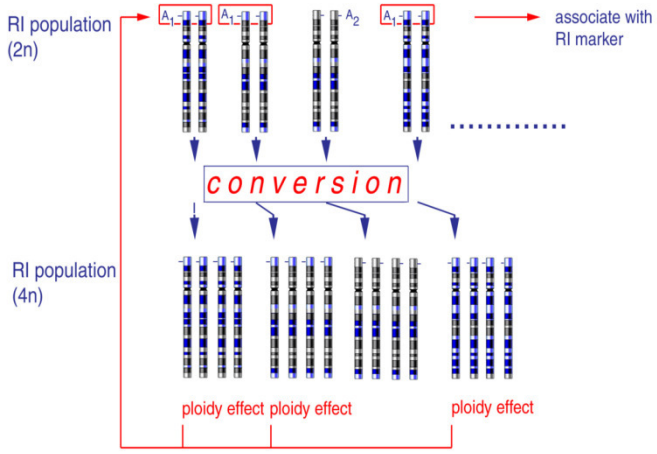


Fig. 31 Use the converted RILs to find the polyploidy effect
The diploid RILs (NASC) were converted into tetraploids. If in some tetraploid RILs (for example here the three lines marked with "ploidy effect"), there is a common trait which is distinct from the diploid lines and the other RI tetraploid lines (e.g. abnormal vs. normal flowers) linked to a known marker in a certain region among these tetraploid RILs (for example, A1 different from A2), then A1 is associated with this trait as a polyploidy effect. (Figure provided by Prof. Dr. R .A. Torres Ruiz)

The region between the marker "LRRPK" and the marker "CDs11" on chromosome 4 may link to the abnormal flower trait in tetraploid RILs, since all the lines having abnormal flowers showed Col-0 origin in this region while other lines displayed L*er*-0 origin. It is anticipated that a flower phenotype might not be controlled by a single factor, and there should be many factors involved indeed. The variation of the frequency of the occurrence of the abnormal flowers among these tetraploid RILs might point to the combined function of the different factors in different genetic backgrounds. However, this result requires further thests given that the number of used tetraploid lines was limited. The same holds true for the seed weight trait. More RILs need to be converted into tetraploids and be confirmed as tetraploids by flow cytometry to do a meaningful mapping.

Individual seeds of tetraploid plants accumulated more proteins, fatty acids, soluble sugars and starch compared to diploid plants (Li et al., 2012). This explains the heavier seeds in tetraploids. For L*er*-0, it seems the seeds doubled the weight when

converted into tetraploids. While for Col-0 tetraploids seed, the weight is less than the double of the weight of the diploid seed. Since the gene expression alteration pattern was already ecotype dependent, it is not surprising that the affected metabolic difference was also ecotype dependent. The seed weight alteration difference can further prove this.

4.9 Implications for evolution and plant breeding

A former study indicated that newly formed polyploids are unstable and their genomes undergo rapid repatterning (Wendel, 2000). However, such genomic repatterning has been mostly observed in allopolyploids and there are many reasons to expect that hybridization may be causally responsible. Transposable elements that are repressed within each parent lineage but activated in hybrids can facilitate the movement of genes and promote unequal crossing over. For example, Josefsson et al. (2006) found that maternally derived siRNAs are not sufficient to repress retrotransposons in the paternal genome of *Arabidopsis thaliana* × *A. arenosa* hybrids. Divergence of centromeres and centromeric histones can lead to segregation distortion and nondisjunction in hybrids (Malik and Bayes, 2006). In addition, nonhomologous recombination and nonreciprocal exchanges are particularly likely among homeologous chromosomes that bear structural rearrangements. In autopolyploids, reduction in genome size through chromosome loss has been observed in both *Candida albicans* (Bennett et al., 2005) and *S. cerevisiae* (Gerstein et al., 2006), largely restoring the diploid complement. Therefore, genomic repatterning in polyploids is not entirely driven by hybridization. Through this work, the induced autotetraploid of *A. thaliana* is relatively stable for the first three generations with respect to gene expression. The genomic repatterning can occur over the long term and will not immediately be affected after polyploidisation. Genomic repatterning can increase the genetic variability available to newly formed polyploid populations and fuel the evolution of a polyploid population if individuals can survive the onslaught of genomic mutations (Otto, 2007).

Allopolyploids and their homoploid progenitors could resort to numerous alterations in gene expression, allowing for rapid adaptations to extreme habitats (Wang et al., 2006 a and b; Ni et al., 2009). On the other hand, they might be prone to developmental accidents due to the interference of ploidy, heterosis, and effects that result from the reunion of divergent genomes (Otto and Whitton, 2000; Comai, 2005; Mallet, 2007; Osborn et al., 2003; Soltis, 2009; Rieseberg and Willis, 2007; Leitch and Leitch, 2008). Neo-autopolyploids could resort to a lower and stably heritable

number of ploidy-induced alterations allowing selective adaptations. In the long term, these processes might entail mutations that would act to fix such alterations (Osborn et al., 2003), which would otherwise be lost. If so, this mechanism could appreciably impact the evolution of autopolyploids, together with other known mechanisms such as point mutations or genetic drift.

Additional aspects for these two forms of ploidy need to be considered. First, autopolyploidy can occur recurrently (Soltis et al., 2003; De et al., 2005; Leitch and Leitch, 2008). Second, autopolyploids could "feed" allopolyploid evolution. For instance, the generation of synthetic *A. suecica* allopolyploids was only possible through crosses of a synthetic autotetraploid of *A. thaliana* with *A. arenosa* because of the lethality of homoploid hybrids (Comail et al., 2000).

Several studies showed that autopolyploids are more adaptable to stressful conditions. An induced autotetraploid *Arabidopsis* line is more tolerant to boron deficiency (Kasajima et al., 2010). Autotetraploid turnip (*Brassica rapa*) and autotetraploid *A. thaliana* lines exhibit a better adaptation to salinity stress (Meng et al., 2011; Chao et al., 2013). Furthermore, field research showed that autohexaploids have a fivefold fitness advantage over autotetraploids in dune habitats in *Achilea borealis*, a widespread autotetraploid plant with localized hexaploid populations (Ramsey, 2011). This indicates that autopolyploidisation can already give the better adaption to the environment and allow the invasion of a habitat.

Allopolyploids are taxonomically predominated, but a reliable estimate for the frequency of autopolyploid species is yet to be found. In fact, autopolyploids might be much more prevalent in nature than presently known (Otto and Whitton, 2000; Ramsey and Schemske, 2002; Soltis et al., 2003; Mallet, 2007; Darlington CD, 1963; Soltis, 2009; Rieseberg and Willis, 2007), because they are sometimes difficult to recognize based on morphology. The presented results support this notion and indicate that the success of autotetraploids might critically depend on the magnitude of a species' natural genetic variability. The plant surviving ratio after polyploidy induction and the gene expression alteration pattern strongly depends on the genetic background (different among different ecotypes). This observation could impact plant breeding because autopolyploidy might be much better exploited if the natural variability of a species is considered.

5. References

Adams, K. L., Cronn, R., Percifield, R., & Wendel, J. F. (2003). Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. *Proceedings of the National Academy of Sciences of the United States of America*, 100(8), 4649–54.

Adams, K. L., Percifield, R., Wendel, J. F. (2004). Organ-specific silencing of duplicated genes in a newly synthesized cotton allotetraploid. *Genetics*,168(4),2217-26.

Adams, K. L., & Wendel, J. F. (2005). Polyploidy and genome evolution in plants. *Current Opinion in Plant Biology*, *8*(2), 135–41.

Albertin, W., Balliau, T., Brabant, P., Chèvre, A.-M., Eber, F., Malosse, C., & Thiellement, H. (2006). Numerous and rapid nonstochastic modifications of gene products in newly synthesized Brassica napus allotetraploids. *Genetics*, *173*(2), 1101–13.

Albertin, W., Brabant, P., Catrice, O., Eber, F., Jenczewski, E., Chèvre, A.-M.and Thiellement, H. (2005). Autopolyploidy in cabbage (*Brassica oleracea L.*) does not alter significantly the proteomes of green tissues. *Proteomics*, 5: 2131–2139.

Allario, T., Brumos, J., Colmenero-Flores, J. M., Tadeo, F., Froelicher, Y., Talon, M., Navarro, L., et al. (2011). Large changes in anatomy and physiology between diploid Rangpur lime (*Citrus limonia*) and its autotetraploid are not associated with large changes in leaf gene expression. *Journal of Experimental Botany*, *62*(8), 2507–19.

Alonso-Blanco, C., Peeters, A. J. M., Koornneef, M., Lister, C., Dean, C., van den Bosch, N. , Pot, J. and Kuiper, M. T. R. (1998). Development of an AFLP based linkage map of L*er*, Col and Cvi *Arabidopsis thaliana* ecotypes and construction of a L*er*/Cvi recombinant inbred line population. *The Plant Journal*, 14: 259–271. doi: 10.1046/j.1365-313X.1998.00115.

The *Arabidopsis* Genome Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, *408*(6814), 796–815. doi:10.1038/35048692

Armengaud, P., Breitling, R., & Amtmann, A. (2004). The Potassium-Dependent Transcriptome of *Arabidopsis* Reveals a Prominent Role of Jasmonic Acid in Nutrient Signaling. *Plant Physiology*, *136*(September), 2556–2576. doi:10.1104/pp.104.046482.2556

Aufsatz, W., Mette, M., Matzke, A., Matzke, M. (2004). The role of *MET1* in RNA-directed de novo and maintenance methylation of CG dinucleotides. *Plant Molecular Biology*,54(6):793-804.

Becker, C., Hagmann, J., Muller, J., Koenig, D., Stegle, O., Borgwardt, K., & Weigel, D. (2011). Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. *Nature*, *480*(7376), 245–249.

Bedford, T., & Hartl, D. L. (2009). Optimization of gene expression by natural selection. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(4), 1133–8. doi:10.1073/pnas.0812009106

Benkova, E., Michniewicz, M., Sauer, M., Teichmann, T., Seifertová, D., Jürgens, G., & Friml, J. (2003). Local, efflux-dependent auxin gradients as a common module for plant organ formation. *Cell*, *115*(5), 591–602.

Bennett, M., Leitch, I. (2005). Nuclear DNA amounts in angiosperms: progress, problems and prospects. *Annals of Botany*.2005 Jan;95(1):45-90.

Bhattacharyya, B., Panda, D., Gupta, S. and Banerjee, M. (2008). Anti-mitotic activity of colchicine and the structural basis for its interaction with tubulin. *Medicinal Research Reviews,* 28: 155–183. doi: 10.1002/med.20097

Bhutani, N., Burns, D. M., & Blau, H. M. (2011). DNA demethylation dynamics. *Cell*, *146*(6), 866–72. doi:10.1016/j.cell.2011.08.042

Birchler, J., Bhadra, U., Bhadra, M., Auger, D. (2001). Dosage-dependent gene regulation in multicellular eukaryotes: Implications for dosage compensation, aneuploid syndromes, and quantitative traits. *Developmental Biology* 234: 275–288. doi: 10.1006/dbio.2001.0262.

Birchler, J., Riddle, N. C., Auger, D. L., & Veitia, R. (2005). Dosage balance in gene regulation: biological implications. *Trends in Genetics : TIG*, *21*(4), 219–26. doi:10.1016/j.tig.2005.02.010

Birnbaum, K., Shasha, D. E., Wang, J. Y., Jung, J. W., Lambert, G. M., Galbraith, D. W., & Benfey, P. N. (2003). A gene expression map of the *Arabidopsis* root. *Science (New York, N.Y.)*, *302*(5652), 1956–60. doi:10.1126/science.1090022

Blanc, G., & Wolfe, K. H. (2004). Functional divergence of duplicated genes formed by polyploidy during *Arabidopsis* evolution, *Plant Cell*, *16*(July), 1679–1691. doi:10.1105/tpc.021410.tion

Borron, S., Scherrmann, J., Baud, F. (1996). Markedly altered colchicine kinetics in a fatal intoxication: examination of contributing factors. *Human & Experimental Toxicology*, Nov, 15(11):885-90

Borsani, O., Zhu, J., Verslues, P. E., Sunkar, R., & Zhu, J.-K. (2005). Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in *Arabidopsis*. *Cell*, *123*(7), 1279–91. doi:10.1016/j.cell.2005.11.035

Bustin, S. a, Benes, V., Nolan, T., & Pfaffl, M. W. (2005). Quantitative real-time RT-PCR--a perspective. *Journal of Molecular Endocrinology*, *34*(3), 597–601. doi:10.1677/jme.1.01755

Cai, G., Yang, Q., Yang, Q., Zhao, Z., Chen, H., Wu, J., Fan, C., et al. (2012). Identification of candidate genes of QTLs for seed weight in Brassica napus through comparative mapping among *Arabidopsis* and *Brassica* species. *BMC Genetics*, *13*(1), 105.doi:10.1186/1471-2156-13-105

Camacho, J. P. M., Perfectti, F., Teruel, M., López-León, M. D., & Cabrero, J. (2004). The odd-even effect in mitotically unstable B chromosomes in grasshoppers. *Cytogenetic and Genome Research*, *106*(2-4), 325–331.

Campbell, P. and Braam, J. (1999). In vitro activities of four xyloglucan endotransglycosylases from *Arabidopsis*. *The Plant Journal*, 18: 371–382. doi: 10.1046/j.1365-313X.1999.00459.x

Chandler, V., Alleman, M. (2008). Paramutation. Epigenetic instructions passed across generations. *Genetics.*178 (4), 1839-44.

Chao, D., Dilkes, B., Luo, H. (2013). Polyploids exhibit higher potassium uptake and salinity tolerance in *Arabidopsis. Science*. 341:658-659

Chen, Z. J., Comai, L., & Pikaard, C. S. (1998).Gene dosage and stochastic effects determine the severity and direction of uniparental ribosomal RNA gene silencing (nucleolar dominance) in *Arabidopsis* allopolyploids. *Proceedings of the National Academy of Sciences of the United States of America*, 95(25), 14891–6.

Chinnusamy, V., & Zhu, J.-K. (2009). Epigenetic regulation of stress responses in plants. *Current Opinion in Plant Biology*, *12*(2), 133–139. doi:http://dx.doi.org/10.1016/j.pbi.2008.12.006

Comai, L, Tyagi, a P., Winter, K., Holmes-Davis, R., Reynolds, S. H., Stevens, Y., & Byers, B. (2000). Phenotypic instability and rapid gene silencing in newly formed *Arabidopsis* allotetraploids. *The Plant Cell*, *12*(9), 1551–68.

Comai, Luca. (2005). The advantages and disadvantages of being polyploid. *Genetics*, *6*(11), 836–46.doi:10.1038/nrg1711

Conesa, A., & Götz, S. (2008). Blast2GO: A comprehensive suite for functional analysis in plant genomics. *International Journal of Plant Genomics*, *2008*, 619832. doi:10.1155/2008/619832

Darlington, CD. (1963). Psychology, genetics and the process of history. *British Journal of Psychology,* 54:293-8.

Dawe, R., Freeling, M. (1991). Cell lineage and its consequences in higher plants. *The Plant Journal,*1:3–8

De Bodt, S., Maere, S., & Van de Peer, Y. (2005). Genome duplication and the origin of angiosperms. *Trends in Ecology & Evolution*, *20*(11), 591–7. doi:10.1016/j.tree.2005.07.008

Del Pozo, J. C., Ramirez-Parra, E. (2014). Deciphering the molecular bases for drought tolerance in *Arabidopsis* autotetraploids. *Plant Cell Environment*, Apr 9. Doi 10.1111/pce 12344

De Smet, R., & Van de Peer, Y. (2012). Redundancy and rewiring of genetic networks following genome-wide duplication events. *Current Opinion in Plant Biology*, *15*(2), 168–76. doi:10.1016/j.pbi.2012.01.003

De Storme, N., & Geelen, D. (2013). Sexual polyploidization in plants - cytological mechanisms and molecular regulation. *The New Phytologist*, doi:10.1111/nph.12184

De Veylder, L., Beeckman, T., & Inzé, D. (2007). The ins and outs of the plant cell cycle. *Nature Reviews, Molecular Cell Biology*, *8*(8), 655–65.doi:10.1038/nrm2227

Dilkes, B. P., Spielman, M., Weizbauer, R., Watson, B., Burkart-Waco, D., Scott, R. J., & Comai, L. (2008). The maternally expressed WRKY transcription factor TTG2 controls lethality in interploidy crosses of *Arabidopsis*. *PLoS Biology*, *6*(12), 2707–20. doi:10.1371/journal.pbio.0060308

Dong, S., & Adams, K. L. (2011). Differential contributions to the transcriptome of duplicated genes in response to abiotic stresses in natural and synthetic polyploids. *The New Phytologist*, *190*(4), 1045–57. doi:10.1111/j.1469-8137.2011.03650.x

Dubcovsky, J., & Dvorak, J. (2007). Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science (New York, N.Y.)*, *316*(5833), 1862–6. doi:10.1126/science.1143986

Etheridge, N., Chen, Y.-F., & Schaller, G. E. (2005). Dissecting the ethylene pathway of *Arabidopsis*. *Briefings in Functional Genomics & Proteomics*, *3*(4), 372–81.

Fodde, R., Smits, R. (2002).Cancer biology. A matter of dosage. *Science,* 298: 761–763. doi: 10.1126/science.1077707

Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y. L., & Postlethwait, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. *Genetics*, 151(4), 1531–45.

Fox, D. T., & Duronio, R. J. (2013). Endoreplication and polyploidy: insights into development and disease. *Development (Cambridge, England)*, *140*(1), 3–12. doi:10.1242/dev.080531

Furutani, M., Kajiwara, T., Kato, T., Treml, B. S., Stockum, C., Torres-Ruiz, R. a, & Tasaka, M. (2007). The gene *MACCHI-BOU 4/ENHANCER OF PINOID* encodes a NPH3-like protein and reveals similarities between organogenesis and phototropism at the molecular level. *Development (Cambridge, England)*, *134*(21), 3849–59. doi:10.1242/dev.009654

Galitski, T., Saldanha, A., Styles, C., Lander, E., Fink, G., (1999). Ploidy regulation of gene expression. *Science*, Vol. 285 no. 5425 pp. 251-254 doi:10.1126/science.285.5425.251

Garber, E. (1955). Cytogenetics of Sorghum. I. The orientation of interchange complexes and quadrivalents at metaphase I in S. purpureo-sericeum. *Botanical Gazette*. 116:369–72

Gerstein, A., Chun, H., Grant, A., Otto, S. (2006). Genomic convergence toward diploidy in Saccharomyces cerevisiae. *PLoS Genet* 2(9):e145.doi:10.1371/journal.pgen.0020145

Giaever, G., Shoemaker, D., Jones, T., Liang, H., Winzeler, E., Astromoff, A., Davis, R. (1999). Genomic profiling of drug sensitivities via induced haploinsufficiency. *Nature Genetics* 21: 278–283.

Gil, P., Liu, Y., Orbović, V., Verkamp, E., Poff, K. L., & Green, P. J. (1994). Characterization of the auxin-inducible *SAUR-AC1* gene for use as a molecular genetic tool in *Arabidopsis*. *Plant Physiology*, 104(2), 777–84.

Glawischnig, E., Tomas, A., Eisenreich, W., Spiteller, P., Bacher, A., Gierl, A. (2000). Auxin Biosynthesis in Maize Kernels. *Plant Physiology,* Vol. 123, pp. 1109–1119

Goto, D. B., & Naito, S. (2002). At*MRD1* and At*MRU1*, two novel genes with altered mRNA levels in the methionine over-accumulating *mto1-1* mutant of *Arabidopsis thaliana*. *Plant & Cell Physiology*, *43*(8), 923–31.

Grant, V. (1981). Plant speciation. *Columbia University Press*, New York, USA.

Gülersönmez, M. C. (2011) Master of science Thesis, Technische Universität München

Guo, M., Davis, D., & Birchler, J. a. (1996). Dosage effects on gene expression in a maize ploidy series. *Genetics*, *142*(4), 1349–55.

Haage, K. (2005) Diploma Thesis, Technische Universität München

Henry, I M, Dilkes, B. P., Tyagi, a P., Lin, H.-Y., & Comai, L. (2009). Dosage and parent-of-origin effects shaping aneuploid swarms in *A. thaliana*. *Heredity*, *103*(6), 458–68. doi:10.1038/hdy.2009.81

Henry, Isabelle M, Dilkes, B. P., & Comai, L. (2007). Genetic basis for dosage sensitivity in *Arabidopsis thaliana*. *PLoS Genetics*, *3*(4), e70. doi:10.1371/journal.pgen.0030070

Henry, Isabelle M, Dilkes, B. P., Young, K., Watson, B., Wu, H., & Comai, L. (2005). Aneuploidy and genetic variation in the *Arabidopsis thaliana* triploid response. *Genetics*, *170*(4), 1979–88. doi:10.1534/genetics.104.037788

Henry, I., Dilkes, B., Comai, L. (2007) Genetic basis for dosage sensitivity in *Arabidopsis thaliana*. *PLoS Genet* 3(4): e70. doi:10.1371/journal.pgen.0030070

Herridge, R., Day, R. C., Baldwin, S., & Macknight, R. C. (2011). Rapid analysis of seed size in *Arabidopsis* for mutant and QTL discovery. *Plant Methods*, *7*(1), 3. doi:10.1186/1746-4811-7-3

Huettel, B., Kreil, D. P., Matzke, M., & Matzke, A. J. M. (2008). Effects of aneuploidy on genome structure, expression, and interphase organization in *Arabidopsis thaliana. PLoS Genetics*, *4*(10), e1000226. doi:10.1371/journal.pgen.1000226

Ilic, K., SanMiguel, P. J., & Bennetzen, J. L. (2003). A complex history of rearrangement in an orthologous region of the maize, sorghum, and rice genomes. *Proceedings of the National Academy of Sciences of the United States of America*, 100(21), 12265–70. doi:10.1073/pnas.1434476100

Inácio, A., Pinho, J., Pereira, P. M., Comai, L., & Coelho, M. M. (2012). Global analysis of the small RNA transcriptome in different ploidies and genomic combinations of a vertebrate complex--the Squalius alburnoides. *PloS One*, *7*(7), e41158. doi:10.1371/journal.pone.0041158

Jones, R.N., Rees, H. (1982). B chromosomes. *Academic Press, London*

Josefsson, C., Dilkes, B., & Comai, L. (2006, July 11). Parent-Dependent Loss of Gene Silencing during Interspecies Hybridization. *Current Biology*: *CB. Cell Press*, 11;16 (13):1322-8.

Jürgens, G., Mayer, U., Busch, M., Lukowitz, W., & Laux, T. (1995). Pattern Formation in the *Arabidopsis* Embryo: A Genetic Perspective. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* ,*350* (1331), 19–25. doi:10.1098/rstb.1995.0132

Kasajima, I., Ide, Y., Yokota Hirai, M. and Fujiwara, T. (2010), WRKY6 is involved in the response to boron deficiency in *Arabidopsis thaliana*. *Physiologia Plantarum*, 139: 80–92. doi: 10.1111/j.1399-3054.2010.01349.x

Kashkush, K., Feldman, M., & Levy, A. a. (2002). Gene loss, silencing and activation in a newly synthesized wheat allotetraploid. *Genetics*, *160*(4), 1651–9.

Kauschmann, A.,Jessop, A.,Koncz, C.,Szekeres, M.,Willmitzer, L.,Altmann, T.(1996). Genetic evidence for an essential role of brassinosteroids in plant development. *The Plant Journal*, 9(5), 701

Kawade, K., Ishizaki, T., & Masuda, K. (2008).Differential expression of ribosome-inactivating protein genes during somatic embryogenesis in spinach (Spinacia oleracea). *Physiologia Plantarum*, *134*(2), 270–81. doi:10.1111/j.1399-3054.2008.01129.x

Kingsbury, M., Friedman, B., McConnell, M., Rehen, S., Yang, A., Kaushal, D., Chun, J. (2005). Aneuploid neurons are functionally active and integrated into brain circuitry. *Proceedings of the National Academy of Sciences of the United States of America* 102: 6143–6147. doi: 10.1073/pnas.0408171102

Bomblies, K., Madlung, A. (2014). Polyploidy in the *Arabidopsis* genus. *Chromosome Res*, doi: 10.1007/s10577-014-9416-x

Kobe, B., & Kajava, A. V. (2001). The leucine-rich repeat as a protein recognition motif. *Current Opinion in Structural Biology*, *11*(6), 725–732. doi:http://dx.doi.org/10.1016/S0959-440X(01)00266-4

Kung, A. L., Sherwood, S. W., & Schimke, R. T. (1990). Cell line-specific differences in the control of cell cycle progression in the absence of mitosis. *Proceedings of the National Academy of Sciences of the United States of America*, *87* (24), 9553–9557.

Lai, J., Ma, J., Swigonová, Z., Ramakrishna, W., Linton, E., Llaca, V., Tanyolac, B., et al. (2004). Gene loss and movement in the maize genome. *Genome Research*, 14(10A), 1924–31. doi:10.1101/gr.2701104

Larochelle, S. (2013). Directing DNA demethylation. *Nature Structural & Molecular Biology*, *20*(3), 308–308. doi:10.1038/nsmb.2540

Law, J. A., & Jacobsen, S. E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Reviewes Genetics*, *11*(3), 204–220. doi:10.1038/nrg2719.

Lee, B., Henderson, D. A., & Zhu, J. (2005). The *Arabidopsis* cold-responsive transcriptome and its regulation by *ICE1*. *The Plant Cell*,*17*(November), 3155–3175. doi:10.1105/tpc.105.035568.1

Leitch, AR., Leitch, IJ.(2008). Genomic plasticity and the diversity of polyploidy plants. *Science*, Apr 25;320(5875):481-3. doi: 10.1126/science.1153585.

Leonhardt, N., Kwak, J. M., Robert, N., Waner, D., Leonhardt, G., & Schroeder, J. I. (2004). Microarray Expression Analyses of *Arabidopsis* Guard Cells and Isolation of a Recessive Abscisic Acid Hypersensitive Protein Phosphatase 2C Mutant. *The Plant Cell*,*16*(March), 596–615. doi:10.1105/tpc.019000.2

Levin, D. (2011). Polyploidy and ecological trans figuration in Achillea. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(17), 6697–8. doi:10.1073/pnas.1103568108

Levy, A. A., Feldman, M. (2002).The impact of polyploidy on grass genome evolution. *Plant Physiology.*130(4):1587-93.

Li, J., Yang, Z., Yu, B., Liu, J., & Chen, X. (2005).Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in *Arabidopsis*. *Current Biology : CB*, *15*(16), 1501–7. doi:10.1016/j.cub.2005.07.029

Ling, H.-Q., Zhao, S., Liu, D., Wang, J., Sun, H., Zhang, C., Fan, H., et al. (2013). Draft genome of the wheat A-genome progenitor *Triticum urartu*. *Nature*, 3–6. doi:10.1038/nature11997

Lister, R., O'Malley, R. C., Tonti-Filippini, J., Gregory, B. D., Berry, C. C., Millar, a H., & Ecker, J. R. (2008). Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. *Cell*, *133*(3), 523–36. doi:10.1016/j.cell.2008.03.029

Li, X., Yu, E., Fan, C., Zhang, C., Fu, T., & Zhou, Y. (2012). Developmental, cytological and transcriptional analysis of autotetraploid *Arabidopsis*. *Planta*, *236*(2), 579–596. doi:10.1007/s00425-012-1629-7

Lu, C., Tej, S. S., Luo, S., Haudenschild, C. D., Meyers, B. C., & Green, P. J. (2005). Elucidation of the small RNA component of the transcriptome. *Science (New York, N.Y.)*, *309*(5740), 1567–9. doi:10.1126/science.1114112

Lukens, L. N., Pires, J. C., Leon, E., Vogelzang, R., Oslach, L., & Osborn, T. (2006). Patterns of Sequence Loss and Cytosine Methylation within a Population of Newly Resynthesized *Brassica napus* allopolyploids. *Plant Physiology*, *140* (1), 336–348. doi:10.1104/pp.105.066308

Madlung, A., Masuelli, R. W., Watson, B., Reynolds, S. H., Davison, J., Comai, L., & Washington, A. M. (2002). Remodeling of DNA methylation and phenotypic and

transcriptional changes in synthetic *Arabidopsis*. *Plant Physiology*, *129*(June), 733–746. doi:10.1104/pp.003095.widespread

Madlung, A., Tyagi, A. P., Watson, B., Jiang, H., Kagochi, T., Doerge, R. W., Martienssen, R. and Comai, L. (2005), Genomic changes in synthetic *Arabidopsis* polyploids. *The Plant Journal*, 41: 221–230. doi: 10.1111/j.1365-313X.2004.02297.x

Malik, H., Bayes, J. (2006). Genetic conflicts during meiosis and the evolutionary origins of centromere complexity. *Biochemical Society Transactions*, 34(Pt 4):569-73.

Mallet, J. (2007). Hybrid speciation. *Nature*, *446*(7133), 279–83.doi:10.1038/nature05706

Matzke, M., Kanno, T., Huettel, B., Daxinger, L., & Matzke, A. J. M. (2007). Targets of RNA-directed DNA methylation. *Current Opinion in Plant Biology*, *10*(5), 512–519. doi:http://dx.doi.org/10.1016/j.pbi.2007.06.007

Mayfield, D., Chen, Z. J., & Pires, J. C. (2011). Epigenetic regulation of flowering time in polyploids. *Current opinion in plant biology*, *14*(2), 174–8. doi:10.1016/j.pbi.2011.03.008

Mayfield-Jones, D., Washburn, J. D., Arias, T., Edger, P. P., Pires, J. C., & Conant, G. C. (2013). Watching the grin fade: Tracing the effects of polyploidy on different evolutionary time scales. *Seminars in Cell & Developmental Biology*, 1–12. doi:10.1016/j.semcdb.2013.02.002

McCollum, G.(1958). Comparative studies of chromosome pairing in natural and induced tetraploid *Dactylis*. *Chromosoma,* 9:571– 605

Meng, H., Jiang, S., Hua, S., Lin, X., Li, Y., Guo, W., Jiang, L. (2011). Comparison between a tetraploid turnip and its diploid progenitor (*Brassica rapa* L.): the adaptation to salinity stress. *Agr Sci China,* 10:363–375

Messing, J., Bharti, A. K., Karlowski, W. M., Gundlach, H., Kim, H. R., Yu, Y., Wei, F., et al. (2004). Sequence composition and genome organization of maize. *Proceedings of the National Academy of Sciences of the United States of America*, 101(40), 14349–14354.

Miller, D. J., Ball, E. E., & Technau, U. (2005). Cnidarians and ancestral genetic complexity in the animal kingdom. *Trends in Genetics*, *21*(10), 536–9. doi:10.1016/j.tig.2005.08.002

Minton, A.P. (2000). Implications of macromolecular crowding for protein assembly. *Current Opinion in Structural Biology,* 10, 34–39.

Miséra, S., Müller, A., Weiland-Heidecker, U., Jürgens, G.(1994). The *FUSCA* genes of *Arabidopsis*: Negative regulators of light responses. *Mol Gen Genet, 244:242–252.*

Misteli, T. (2007). Beyond the sequence: cellular organization of genome function. *Cell*, *128*(4), 787–800. doi:10.1016/j.cell.2007.01.028

Mittelsten Scheid, O, Jakovleva, L., Afsar, K., Maluszynska, J., & Paszkowski, J. (1996). A change of ploidy can modify epigenetic silencing. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(14), 7114–9.

Mittelsten Scheid, Ortrun, Afsar, K., & Paszkowski, J. (2003). Formation of stable epialleles and their paramutation-like interaction in tetraploid *Arabidopsis thaliana*. *Nature Genetics*, *34*(4), 450–4.doi:10.1038/ng1210

Montgomery, S. B., & Dermitzakis, E. T. (2011). From expression QTLs to personalized transcriptomics. *Nature Reviews. Genetics*, *12*(4), 277–82.doi:10.1038/nrg2969

Montgomery, T. a, Yoo, S. J., Fahlgren, N., Gilbert, S. D., Howell, M. D., Sullivan, C. M., Alexander, A., et al. (2008). *AGO1-miR173* complex initiates phased siRNA formation in plants. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(51), 20055–62. doi:10.1073/pnas.0810241105

Mosher, R. A., Schwach, F., Studholme, D., & Baulcombe, D. C. (2008). PolIVb influences RNA-directed DNA methylation independently of its role in siRNA biogenesis. Proceedings of the National Academy of Sciences , of the United States of America, 105 (8), 3145–3150. doi:10.1073/pnas.0709632105

Ni, Z., Kim, E.-D., Ha, M., Lackey, E., Liu, J., Zhang, Y., Sun, Q., et al. (2009). Altered circadian rhythms regulate growth vigour in hybrids and allopolyploids. *Nature*, *457*(7227), 327–31. doi:10.1038/nature07523

Niel, E., Scherrmann, J. (2006). Colchicine today. *Joint Bone Spine*, Dec;73(6):672-8.

Niittylä, T., Fuglsang, A. T., Palmgren, M. G., Frommer, W. B., & Schulze, W. X. (2007). Temporal analysis of sucrose-induced phosphorylation changes in plasma membrane proteins of *Arabidopsis*. *Molecular & Cellular Proteomics* : MCP, 6(10), 1711–26. doi:10.1074/mcp.M700164-MCP200

Noggle, S., Fung, H.-L., Gore, A., Martinez, H., Satriani, K. C., Prosser, R., Oum, K., et al. (2011). Human oocytes reprogram somatic cells to a pluripotent state. *Nature*, *478*(7367), 70–5.doi:10.1038/nature10397

Nossal, N. G., Franklin, J. L., Kutter, E., & Drake, J. W. (2004). Paramutation: Epigenetic Instructions Passed Across Generations. *Genetics*, *168*(3), 1097–104.

Ohno S., (1970). Evolution by Gene Duplication. *New York: Springer-Verlag*

Osborn, T. C., Chris Pires, J., Birchler, J. a., Auger, D. L., Jeffery Chen, Z., Lee, H.-S., Comai, L., et al. (2003). Understanding mechanisms of novel gene expression in polyploids. *Trends in Genetics*, *19*(3), 141–147. doi:10.1016/S0168-9525(03)00015-5

Otto, S. P. (2007). The evolutionary consequences of polyploidy. *Cell*, *131*(3), 452–62. doi:10.1016/j.cell.2007.10.022

Otto, S. P., & Whitton, J. (2000).Polyploid incidence and evolution. *Annual Reviews Genet* , 34:401–37

Papp, B., Pal, C., Hurst, L. (2003). Dosage sensitivity and the evolution of gene families in yeast. *Nature,* 424: 194–197. doi: 10.1038/nature01771

Paun, O., Bateman, R. M., Fay, M. F., Luna, J. a, Moat, J., Hedrén, M., & Chase, M. W. (2011). Altered gene expression and ecological divergence in sibling allopolyploids of *Dactylorhiza* (Orchidaceae). *BMC Evolutionary Biology*, *11*, 113.doi:10.1186/1471-2148-11-113

Penterman, J., Zilberman, D., Huh, J. H., Ballinger, T., Henikoff, S., & Fischer, R. L. (2007).DNA demethylation in the *Arabidopsis* genome.*Proceedings of the National Academy of Sciences of the United States of America*, *104*(16), 6752–7. doi:10.1073/pnas.0701861104

Perazza, D., Herzog, M., Hülskamp, M., Brown, S., Dorne, A., Bonneville, J. (1999). Trichome cell growth *in Arabidopsis thaliana* can be derepressed by mutations in at least five genes. *Genetics,* May 1, 1999 vol. 152 no. 1 461-476

Pignatta, D., Dilkes, B. P., Yoo, S.-Y., Henry, I. M., Madlung, A., Doerge, R. W., Jeffrey Chen, Z., et al. (2010). Differential sensitivity of the *Arabidopsis thaliana* transcriptome and enhancers to the effects of genome doubling. *The New Phytologist*, *186*(1), 194–206. doi:10.1111/j.1469-8137.2010.03198.x

Pignatta, D., Dilkes, B., Wroblewski, T., Michelmore, R. W., & Comai, L. (2008). Transgene-induced gene silencing is not affected by a change in ploidy level. *PloS One*, *3*(8), e3061. doi:10.1371/journal.pone.0003061

Pikaard, C. S., Haag, J. R., Ream, T., & Wierzbicki, A. T. (2008, July 1). Roles of RNA polymerase IV in gene silencing.*Trends in Plant Science*.13(7),390-397.

Qutob, D., Kemmerling, B., Brunner, F., Küfner, I., Engelhardt, S., Gust, A. a, Luberacki, B., et al. (2006). Phytotoxicity and innate immune responses induced by Nep1-like proteins. *The Plant Cell*, *18*(12), 3721–44. doi:10.1105/tpc.106.044180

Ramsey, J. (2011). Polyploidy and ecological adaptation in wild yarrow. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(17), 7096–101. doi:10.1073/pnas.1016631108

Ramsey, J., & Schemske, D. W. (2002). Neopolyploidy in Flowering Plants. *Annual Review of Ecology and Systematics*, *33*(1), 589–639. doi:10.1146/annurev.ecolsys.33.010802.150437

Ramsey, J. (2007). Unreduced gametes and neopolyploids in natural populations of *Achillea borealis* (Asteraceae). *Heredity*, 98(3), 143–50. doi:10.1038/sj.hdy.6800912

Ravelli, R. B. G., Gigant, B., Curmi, P. A., Jourdain, I., Lachkar, S., Sobel, A., & Knossow, M. (2004). Insight into tubulin regulation from a complex with colchicine and a stathmin-like domain. *Nature*, *428*(6979), 198–202.

Reiter, R., Williams, J., Feldmann, K., Rafalski, J., Tingey, S., Scolnik, P. (1992). Global and local genome mapping in *Arabidopsis thaliana* by using recombinant inbred lines and random amplified polymorphic DNAs. *Proceedings of the National Academy of Sciences of the United States of America,* 15; 89(4): 1477–1481.

Riddle, N. C., Jiang, H., An, L., Doerge, R. W., & Birchler, J. a. (2010). Gene expression analysis at the intersection of ploidy and hybridity in maize.*Theoretical and Applied Genetics*, 120(2), 341–53.doi:10.1007/s00122-009-1113-3

Rieseberg, L. H., & Willis, J. H. (2007). Plant speciation. *Science, (New York, N.Y.)*, *317*(5840), 910–4. doi:10.1126/science.1137729

Sánchez, J.-P., Duque, P. and Chua, N.-H. (2004), ABA activates ADPR cyclase and cADPR induces a subset of ABA-responsive genes in *Arabidopsis*. *The Plant Journal*, 38: 381–395. doi: 10.1111/j.1365-313X.2004.02055.x

Saleh, B., Allario, T., Dambier, D., Ollitrault, P., Morillon, R. (2008). Tetraploid citrus rootstocks are more tolerant to salt stress than diploid. *CR Biologies*, 331(9), 703–10

Sairanen, I., Novák, O., Pencík, A., Ikeda, Y., Jones, B., Sandberg, G., & Ljung, K. (2012). Soluble Carbohydrates Regulate Auxin Biosynthesis via PIF Proteins in *Arabidopsis*. *The Plant Cell*, *24*(12), 4907–16. doi:10.1105/tpc.112.104794

Sarilar, V., Palacios, P. M., Rousselet, A., Ridel, C., Falque, M., Eber, F., Chèvre, A.-M., et al. (2013). Allopolyploidy has a moderate impact on restructuring at three contrasting transposable element insertion sites in resynthesized *Brassica napus* allotetraploids. *The New Phytologist*, *198*(2), 593–604. doi:10.1111/nph.12156

Sauna, Z., Kim, I.-W., & Ambudkar, S. (2007). Genomics and the mechanism of P-glycoprotein (ABCB1). *Journal of Bioenergetics and Biomembranes*, *39*(5-6), 481–487. doi:10.1007/s10863-007-9115-9

Saze, H., Mittelsten Scheid, O., Paszkowski, J. (2003). Maintenance of CpG methylation is essential for epigenetic inheritance during plant gameto genesis. *Nature Genetics*, 34:65–69

Schmid, M., Uhlenhaut, N. H., Godard, F., Demar, M., Bressan, R., Weigel, D., & Lohmann, J. U. (2003). Dissection of floral induction pathways using global expression analysis. *Development (Cambridge, England)*, *130*(24), 6001–12. doi:10.1242/dev.00842

Shibuya, K., Fukushima, S., & Takatsuji, H. (2009). RNA-directed DNA methylation induces transcriptional activation in plants. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(5), 1660–5. doi:10.1073/pnas.0809294106

Singh, R., Tsuchiya, T. (1993). Cytogenetics of an unstable trisomie in barley (Hordeum vulgare). *Genome*, 1993 Apr;36(2):350-5.

Slotkin, R. K., Vaughn, M., Borges, F., Tanurdžić, M., Becker, J. D., Feijó, J. A., & Martienssen, R. A. (2009). Epigenetic Reprogramming and Small RNA Silencing of Transposable Elements in Pollen. *Cell*, 136 (3), 461-472.

Soltis, D. E., Soltis, P. S.,Bennett, M. D. Leitch, I. J. (2003). Evolution of genome size in the angiosperms1. *American Journal of Botany*, 90(11), 1596–1603.

Soltis, D. E., Soltis, P. S., Schemske, D. W., Hancock, J. F., Thompson, J. N., Husband, B. C., & Judd, W. S. (2007). Autopolyploidy in angiosperms : have we grossly underestimated the number of species ? *Taxon*, *56*(February), 13–30.

Soltis, P. S., & Soltis, D. E. (2009). The role of hybridization in plant speciation. *Annual review of the plant biology,* 60, 561-588.

Soppa, J. (2013). Evolutionary advantages of polyploidy in halophilic archaea. *Biochemical Society Transactions*, *41*(1), 339–43.doi:10.1042/BST20120315

Storchova, Z., Breneman, A., Cande, J., Dunn, J., Burbank, K., O'Toole, E., & Pellman, D. (2006). Genome-wide genetic analysis of polyploidy in yeast. *Nature*, *443*(7111), 541–547.

Stupar, R. M., Bhaskar, P. B., Yandell, B. S., Rensink, W. a, Hart, A. L., Ouyang, S., Veilleux, R. E., et al. (2007). Phenotypic and transcriptomic changes associated with potato autopolyploidization. *Genetics*, *176*(4), 2055–67. doi:10.1534/genetics.107.074286

Takada, S., & Jürgens, G. (2007). Transcriptional regulation of epidermal cell fate in the *Arabidopsis* embryo. *Development (Cambridge, England)*, *134*(6), 1141–50. doi:10.1242/dev.02803

Teixeira, F. K., Heredia, F., Sarazin, A., Roudier, F., Boccara, M., Ciaudo, C., Cruaud, C., et al. (2009). A role for RNAi in the selective correction of DNA methylation defects. *Science*, *323*(5921), 1600–1604

Terkeltaub, R. A. (2008) Colchicine Update: 2008. *Seminarsin Arthritis and Rheumatism - June 2009* (Vol. 38, Issue 6, Pages 411-419, DOI: 10.1016/j.semarthrit.2008.08.006)

Tilney-Bassett, R. (1986).Plant chimeras. *Edward Arnold Ltd, Baltimore*

Treml, B. S., Winderl, S., Radykewicz, R., Herz, M., Schweizer, G., Hutzler, P., Glawischnig, E., Torres Ruiz, R. A. (2005). The gene *ENHANCER OF PINOID* controls cotyledon development in the *Arabidopsis* embryo. *Development (Cambridge, England)*, *132*(18), 4063–74. doi:10.1242/dev.01969

Tsukaya, H. (2008). Controlling size in multicellular organs: focus on the leaf. *PLoS Biology*, *6*(7), e174. doi:10.1371/journal.pbio.0060174

Uilfoyle, T. O. M. J. G. (1999). Activation and repression of transcription by auxin response factors. *Proceedings of the National Academy of Sciences of the United States of America.96*(May), 5844–5849.

Vaucheret, H., Vazquez, F., Crété, P., & Bartel, D. P. (2004). The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes & Development*, *18*(10), 1187–97. doi:10.1101/gad.1201404

Vaughn, M. W., Tanurdzić, M., Lippman, Z., Jiang, H., Carrasquillo, R., Rabinowicz, P. D., Dedhia, N., et al. (2007). Epigenetic natural variation in *Arabidopsis thaliana. PLoS Biology*, *5*(7), e174. doi:10.1371/journal.pbio.0050174

Vazquez, F., Cre, P., & Bartel, D. P. (2004). The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes & Development*, 1187–1197. doi:10.1101/gad.1201404.called

Veitia, R. (2005). Gene dosage balance: Deletions, duplications and dominance. *Trends Genet* 21(1) 33–35.

Walker, J. D., Oppenheimer, D. G., Concienne, J., & Larkin, J. C. (2000). SIAMESE, a gene controlling the endoreduplication cell cycle in *Arabidopsis thaliana* trichomes. *Development (Cambridge, England)*, *127*(18), 3931–40.

Wang, J., Tian, L., Lee, H.-S., & Chen, Z. J. (2006a). Nonadditive regulation of FRI and FLC loci mediates flowering-time variation in *Arabidopsis* allopolyploids. *Genetics*, *173*(2), 965–74. doi:10.1534/genetics.106.056580

Wang, J., Tian, L., Lee, H.-S., Wei, N. E., Jiang, H., Watson, B., Madlung, A., et al. (2006b). Genomewide nonadditive gene regulation in *Arabidopsis* allotetraploids. *Genetics*, *172*(1), 507–17. doi:10.1534/genetics.105.047894

Wang, J., Tian, L., Madlung, A., Lee, H.-S., Chen, M., Lee, J. J., Watson, B., et al. (2004). Stochastic and epigenetic changes of gene expression in *Arabidopsis* polyploids. *Genetics*, *167*(4), 1961–73. doi:10.1534/genetics.104.027896

Wassenegger, M., Heimes, S., Riedel, L., & Sänger, H. L. (1994, February 11). RNA-directed de novo methylation of genomic sequences in plants. *Cell*, 76(3), 567-576.

Wendel, J. (2000). Genome evolution in polyploids. *Plant Molecular Biology*, 42(1):225-49.

Weier, J., Ferlatte, C., Baumgartner, A., Jung, C., Nguyen, H., Chu, L., Pedersen, R., Fisher, S., Weier, H. (2006). Molecular cytogenetic studies towards the full karyotype analysis of human blastocysts and cytotrophoblasts. *Cytogenetic and Genome Research,* 114: 302–311.doi: 10.1159/000094218.

Xiao, J., Song, C., Liu, S., Tao, M., Hu, J., Wang, J., Liu, W.(2013). DNA Methylation Analysis of Allotetraploid Hybrids of Red Crucian Carp (*Carassius auratus* red var.) and Common Carp (*Cyprinus carpio* L.). *PloS One*, *8*(2), e56409. doi:10.1371/journal.pone.0056409

Xiong, Z., Gaeta, R. T., & Pires, J. C. (2011). Homoeologous shuffling and chromosome compensation maintain genome balance in resynthesized allopolyploid *Brassica napus. Proceedings of the National Academy of Sciences of the United States of America,* 10:108(19):7908-13

Yu, Z., Haberer, G., Matthes, M., Rattei, T., Mayer,K. F., Gierl, A., Torres-Ruiz, R. A. (2010). Impact of natural genetic variation on the transcriptome of autotetraploid *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America,* 107:17809–17814.

Yu, Z., Haage, K., Streit, V. E., Gierl, A., Torres-Ruiz, R. A. (2009). A large number of tetraploid *Arabidopsis thaliana* lines, generated by a rapid strategy, reveal high stability of neo-tetraploids during consecutive generations. *Theoretical and Applied Genetics,* 118:1107–1119.

Zhang, H., & Zhu, J.-K. (2012). Seeing the forest for the trees: a wide perspective on RNA-directed DNA methylation. *Genes & Development*, *26*(16), 1769–73. doi:10.1101/gad.200410.112

Zhang, X., Deng, M., Fan, G. (2014). Differential transcriptome analysis between *paulownia fortune* and its synthesized autopolyploid. *International Journal of Molecular Science,* 15, 5079-5093.

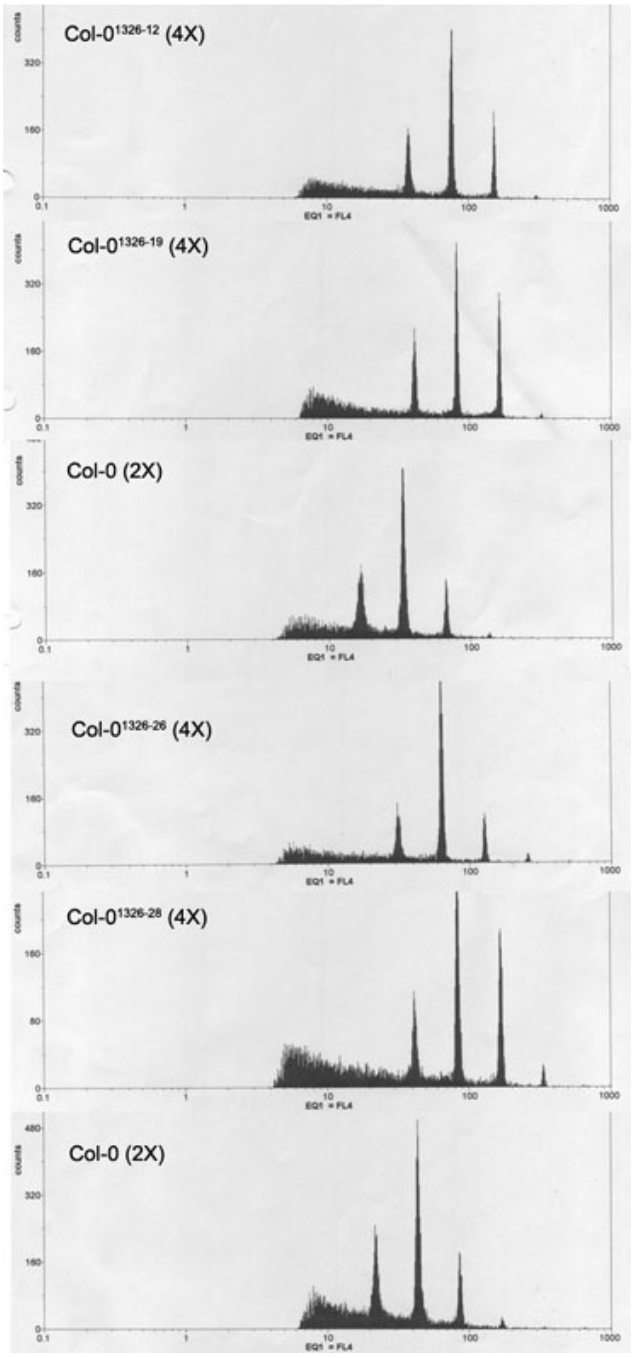
Zhang, X., Hu, C., Yao, J. (2010). Tetraploidization of diploid *Dioscorea* results in activation of the antioxidant defense system and increased heat tolerance. *Journal of Plant Physiology,* 167(2), 88-94.

Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, S. W.-L., Chen, H., Henderson, I. R., et al. (2006). Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. *Cell*, *126*(6), 1189–201. doi:10.1016/j.cell.2006.08.003

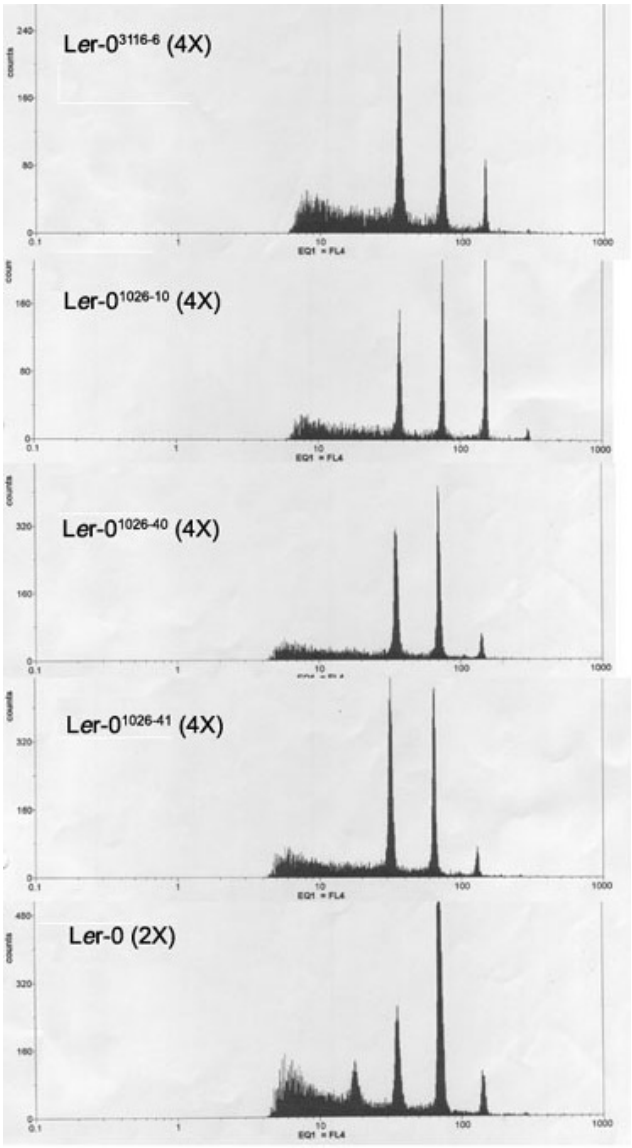
Zhao, Y., Christensen, S.K., Fankhauser, C., Cashman, J.R., Cohen, J.D., Weigel, D., Chory, J. (2001). A role for flavin monooxygenase-like enzymes in auxin biosynthesis. *Science, 291:306–309.*

Zheng, X., Pontes, O., Zhu, J., Miki, D., Zhang, F., Li, W.-X., Iida, K., et al. (2008). ROS3 is an RNA-binding protein required for DNA demethylation in *Arabidopsis*. *Nature*, *455*(7217), 1259–62. doi:10.1038/nature07305

6. Appendix Figures

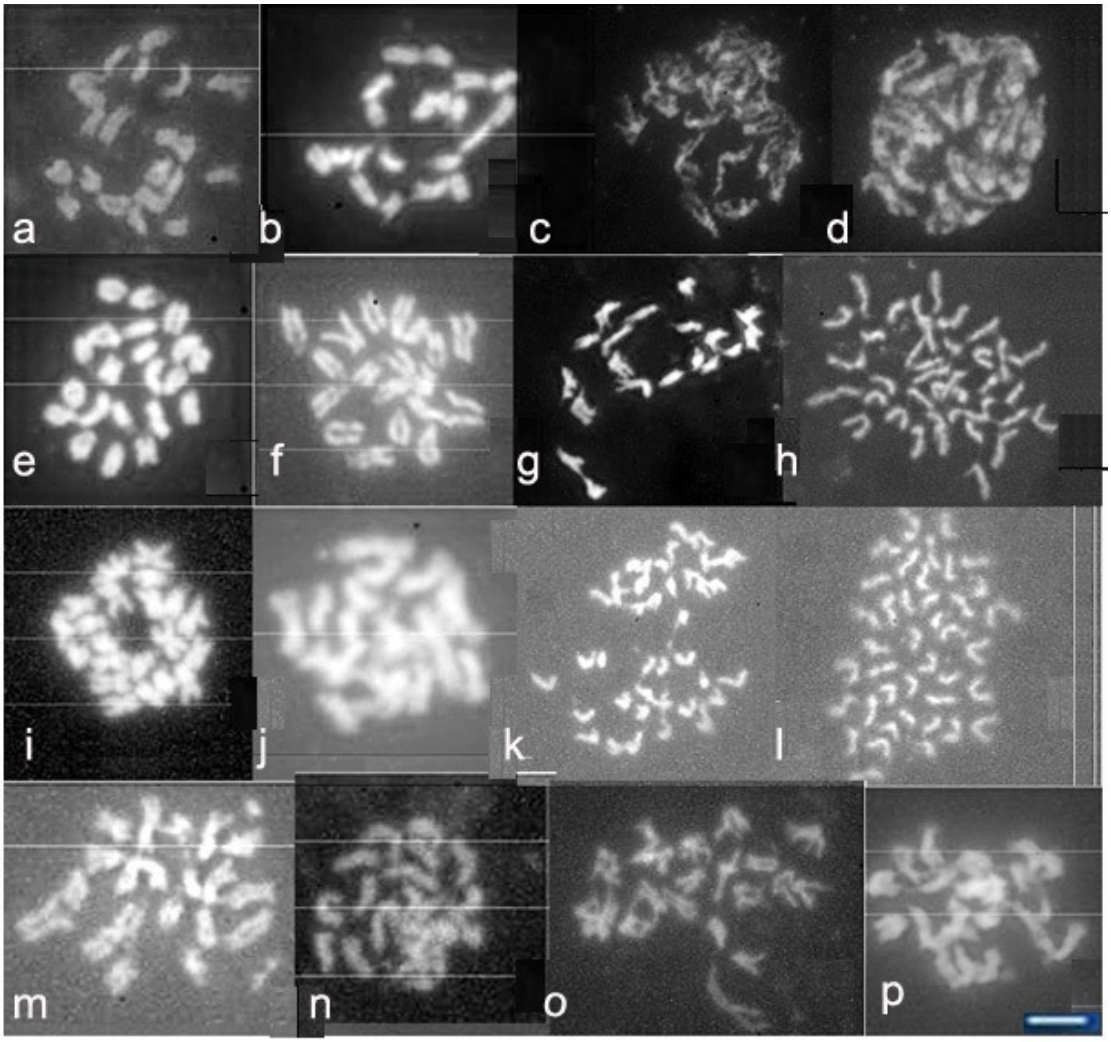


App. Fig. 1 Flow cytometry analysis of the tetraploid *Arabidopsis thaliana* ecotype lines with Col-0 background.
Note the first 2C peak in the diploid line and its absence in the tetraploid lines.



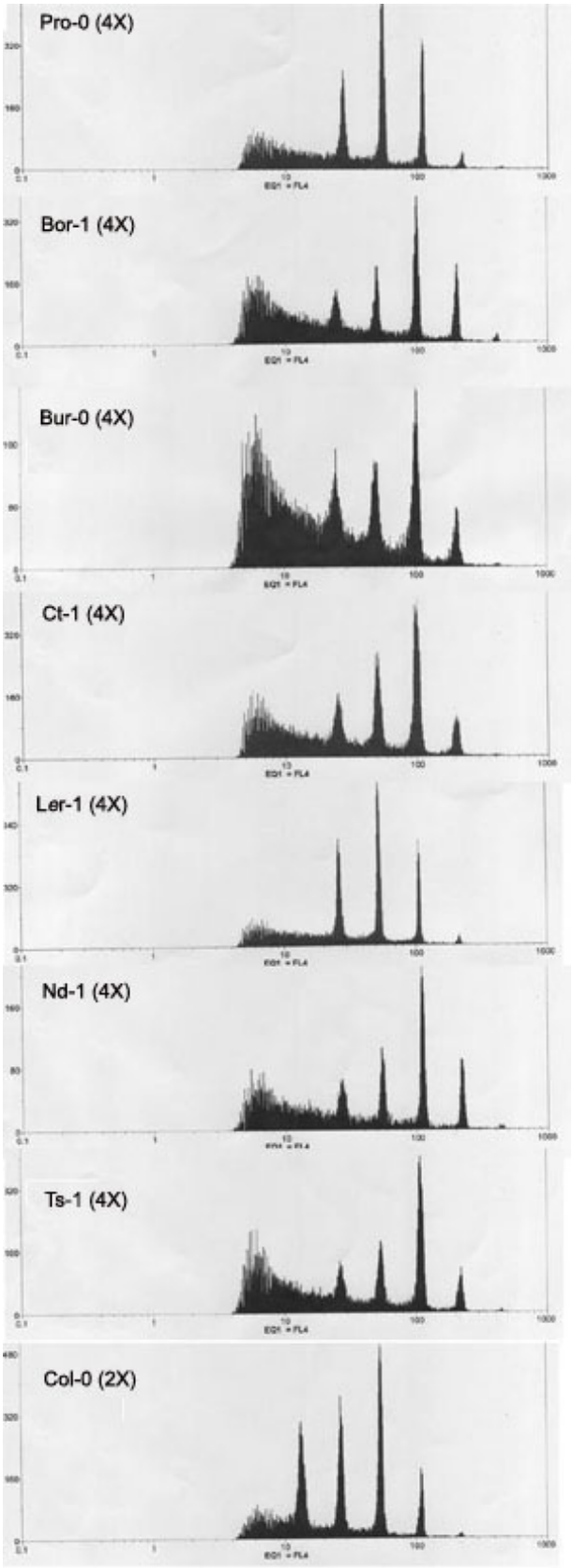
App. Fig. 2 Flow cytometry analysis of the tetraploid *Arabidopsis thaliana* ecotype lines with L*er*-0 background

Note the first 2C peak in the diploid line and its absence in the tetraploid lines.

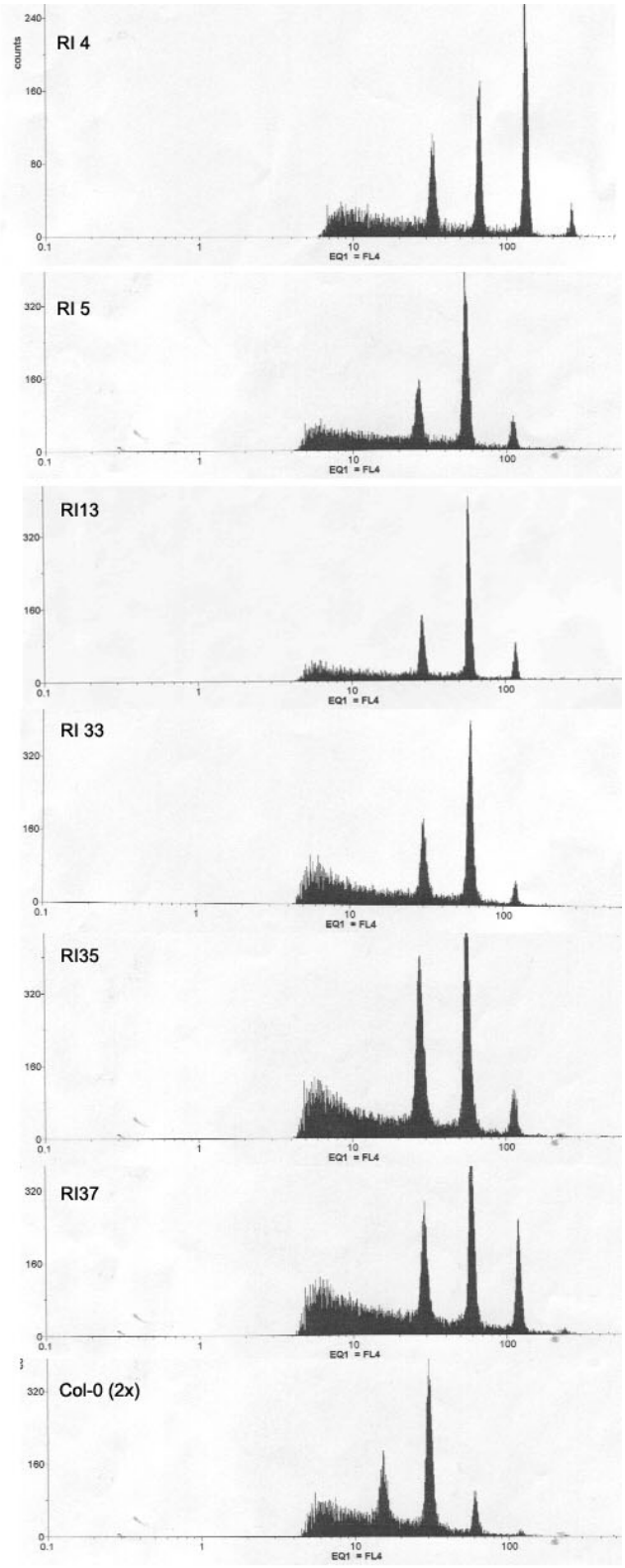


App. Fig. 3 Mitotic chromosome figures of Col-0 and L*er*-0 tetraploid lines used in microarray analysis.

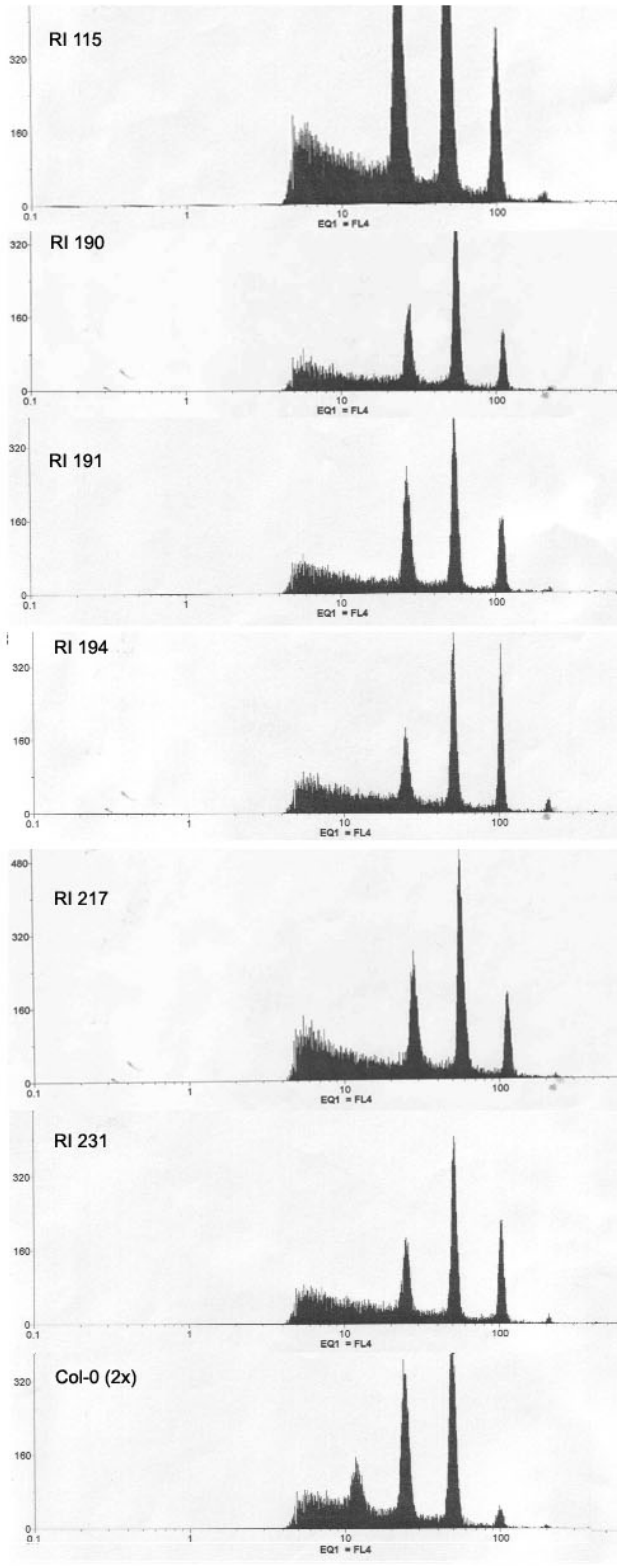
a) and b): Col-$^{0\ 1326\text{-}12}$ (4x); c) and d): Col-0 $^{1326\text{-}19}$ (4x); e) and f) : Col-0 $^{1326\text{-}26}$ (4x);
g) and h): Col-0 $^{1326\text{-}28}$ (4x); i) and j): L*er*-0 $^{3116\text{-}6}$ (4x); k) and l): L*er*-0 $^{1026\text{-}10}$ (4x);
m) and n): L*er*-0 $^{1026\text{-}40}$ (4x); o) and p): L*er*-0 $^{1026\text{-}41}$ (4x). Scale bar in p) is the same as for a-p): 10 µm.



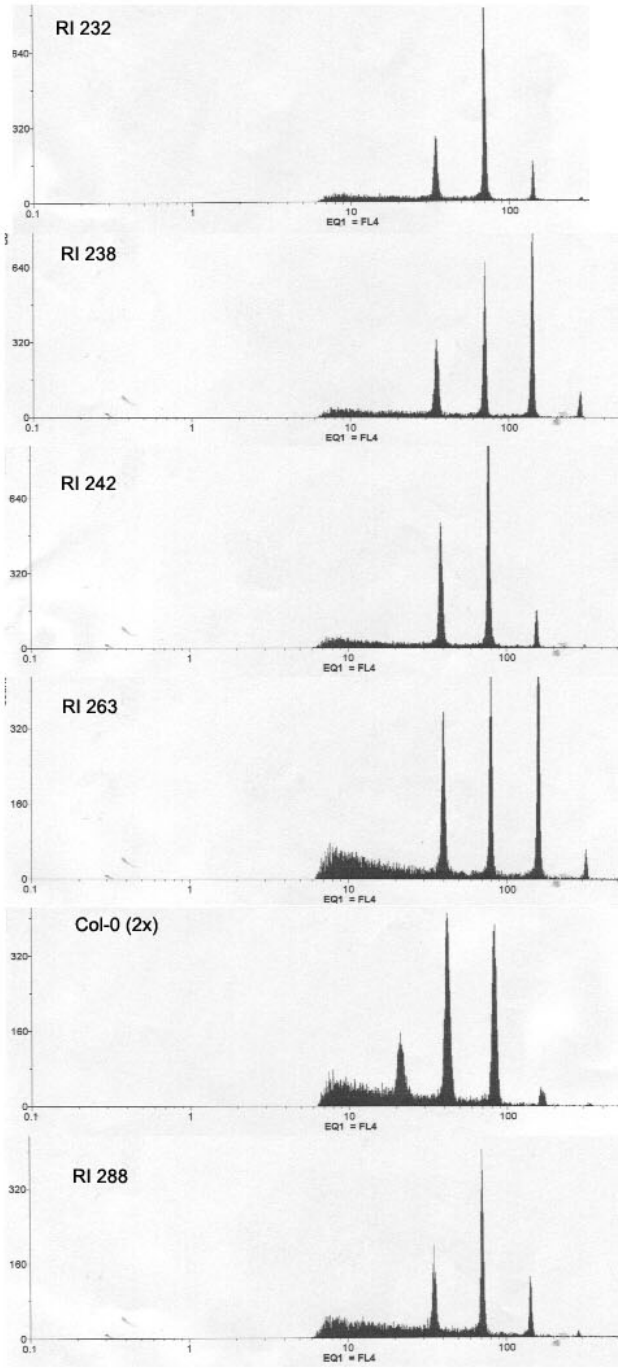
App. Fig. 4 Flow cytometry of the tetraploid *Arabidopsis thaliana* lines with seven different ecotype backgrounds.
Note the first 2C peak in the diploid line and its absence in the tetraploid lines.



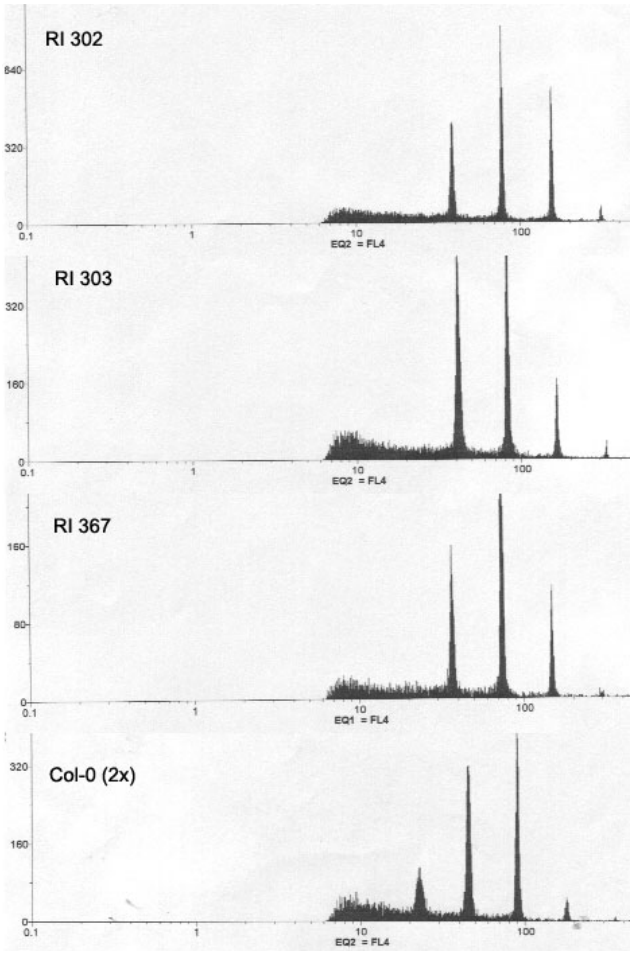
App. Fig. 5 a) Flow cytometry analysis of RI tetraploid lines used for seed weight measurement. RI 4 was newly induced , was not a pure tetraploid line. The other RI lines here are tetraploids.
Note the first 2C peak in the diploid line and its absence in the tetraploid lines.



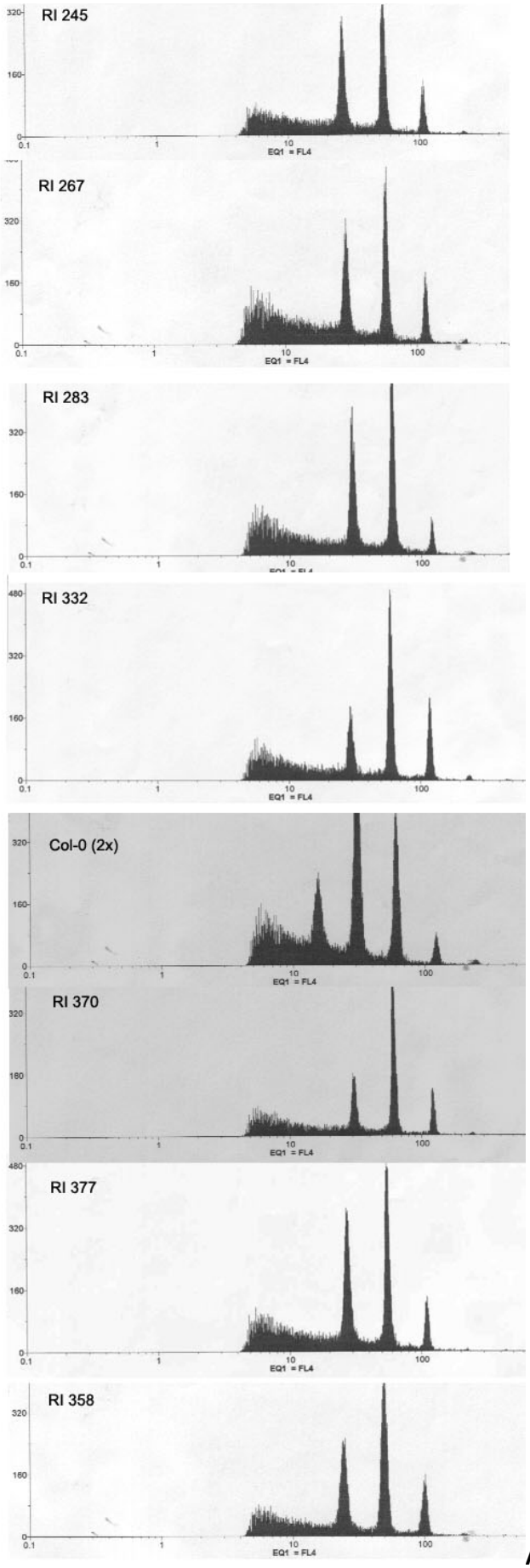
App. Fig. 5 b) Flow cytometry analysis of RI tetraploid lines used for seed weight measurement.
Note the first 2C peak in the diploid line and its absence in the tetraploid lines.



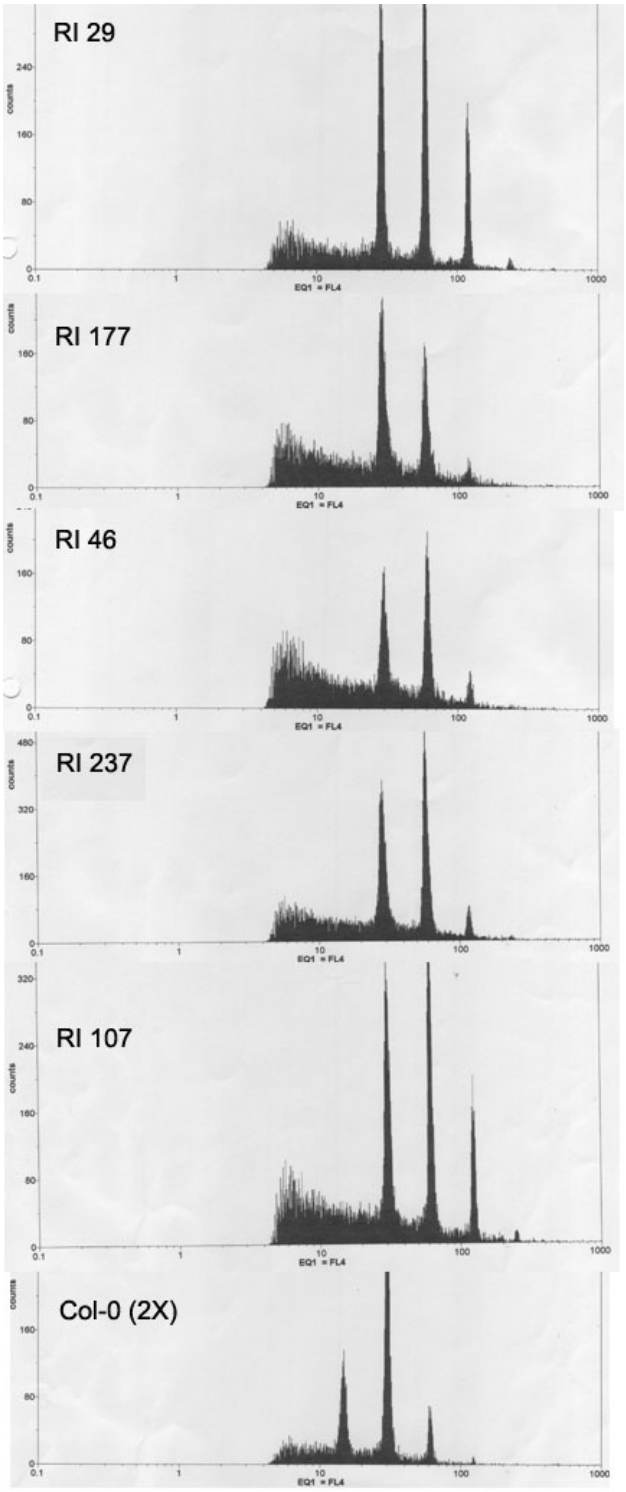
App. Fig. 5 c) Flow cytometry analysis of RI tetraploid lines used for seed weight measurement.
Note the first 2C peak in the diploid line and its absence in the tetraploid lines.



App. Fig. 5 d) Flow cytometry analysis of RI tetraploid lines used for seed weight measurement.
Note the first 2C peak in the diploid line and its absence in the tetraploid lines.



App. Fig. 5 e) RI tetraploid lines



App. Fig. 6 Flow cytometry analysis of newly induced RI tetraploid lines in table 12. Note the first 2C peak in the diploid line and its absence in the tetraploid lines.

7. Appendix Tables

App. Table 1. Chromosome counts

Lines	First Generation		Second Generation		Third Generation	
Col-0[13-2-6-12]	20>>10	15,18,20,11,20,20,18,20,10,18,20,20	.20.	20,18,20,20	.20.	20,20,20
Col-0[13-2-6-15]	20>>10	15,18,20,20,13,11,10,17	22/20			
Col-0[13-2-6-18]	20>>10	18,20,17,18,18,20,20,10,10				
Col-0[13-2-6-19]	.20.	20,20,19,20,18,19,18	.20.	20,20,20,20	.20.	20,20,20,18
Col-0[13-2-6-26]	.20.	20,20,2020,2018,19,18	.20.	20,20,20,19	.20.	20,18,20,19
Col-0[13-2-6-28]	.20.	20,20,20,18,19,20,20,19	.20.	20,18,20,20	.20.	20,19,20,20
Col-0[4n]	.20.	20,20,20,18,19,20,20,18,19	.20.	20,17,20,20		
Pcol31-1-5-1	40>>20>10	20,25,32,40	.20.	19,20,18	.20.	20,19
Pcol31-1-5-2	10>>15	15,10,10,10,10,10	.10.	10,10,10,10,10		
Pcol31-1-5-3	.40.	40,40,40,40,40				
ColP9A			.20.	18,17,20,19		
Ler-0[10-2-6-5]	20>>10	20,20,19,17,13,11,19,10	24/.20.			
Ler-0[10-2-6-10]	20>>10	19,13,10,20,20,20,20,20,20,17	.20.	20,20,20	.20.	20,20,20
Ler-0[10-2-6-19]	10>>20	18,10,9,10				
Ler-0[10-2-6-40]	20>>10	16,20,20,19,20,20,29,17,20,20,18,20	.20.	20,19,20	.20.	20,18,20,20
Ler-0[10-2-6-41]	20>>10	20,18,18,8,15,20,19,13	.20.	20,18,20,20	.20.	20,20,20,18
Ler-0[10-2-6-27]	10>>20	10,10,10,18				

L*er*-0$^{31\text{-}1\text{-}6\text{-}1}$	30/20/10					
L*er*-0$^{31\text{-}1\text{-}6\text{-}2}$	20/10					
L*er*-0$^{31\text{-}1\text{-}6\text{-}6}$	.20.	20,16,20	.20.	20,20,20,18	.20.	20,20,20,19

Chromosome numbers got from the counts of chromosome staining pictures

App. Table 2. Differentially expressed genes in seedlings - diploid Col-0 vs. diploid L*er*-0 (four biological replicates)

Gene ID	Description	log_2 FC	P-value
AT1G52990.1	thioredoxin family protein	4,3274	4,29E-05
AT5G16110.1	unknown protein	-4,1546	4,37E-05
AT4G00585.1	unknown protein	-4,3454	4,67E-05
AT5G28030	cysteine synthase, putative / O-acetylserine (thiol)-lyase, putative / O-acetylserine sulfhydrylase, putative	3,6438	8,71E-05
AT3G52780	ATPAP20/PAP20; acid phosphatase/ protein serine/threonine phosphatase	-4,2367	1,00E-04
AT1G30835	unknown protein	-4,6232	1,17E-04
AT3G27470	unknown protein	-3,8405	1,21E-04
AT1G08780.1	prefoldin, putative	-3,9570	1,24E-04
AT2G43620.1	chitinase, putative	4,5025	1,27E-04
AT1G72800.1	nuM1-related	-3,8977	1,50E-04
AT4G14930.1	acid phosphatase survival protein SurE, putative	-4,3567	1,51E-04
AT5G24850.1	CRY3 (CRYPTOCHROME 3); DNA binding / DNA photolyase/ FMN binding	-4,2676	1,73E-04
AT2G30230.1	unknown protein	3,5473	2,14E-04
AT3G47200	unknown protein	-3,7663	2,24E-04
AT2G16990	tetracycline transporter	-3,2455	2,28E-04
AT5G17880.1	CSA1 (CONSTITUTIVE SHADE-AVOIDANCE1); ATP binding / protein binding / transmembrane receptor	-3,7347	2,39E-04
AT2G26030	F-box family protein	-3,5569	2,40E-04
AT1G21690	EMB1968 (EMBRYO DEFECTIVE 1968); ATPase	-4,7300	2,44E-04
AT4G00970.1	protein kinase family protein	3,6952	2,49E-04
AT5G46490.2	disease resistance protein (TIR-NBS class), putative	-4,7087	2,53E-04
AT5G03460.1	unknown protein	-3,7264	2,70E-04
AT3G63330.1	protein kinase family protein	-3,0299	2,77E-04
AT1G44920.1	unknown protein	-3,4607	2,90E-04
AT5G19100.1	extracellular dermal glycoprotein-related / EDGP-related	3,3157	3,04E-04
AT1G27385	unknown protein	-3,8718	3,09E-04
AT1G12010.1	1-aminocyclopropane-1-carboxylate oxidase, putative / ACC oxidase, putative	-5,0377	3,33E-04
AT5G05750.1	DNAJ heat shock N-terminal domain-containing protein	-2,6895	3,59E-04
AT5G23395.1	unknown protein	-2,9154	3,85E-04
AT3G43520.1	unknown protein	-3,1102	3,90E-04
AT3G18485.1	ILR2 (*IAA*-LEUCINE RESISTANT 2)	2,8085	4,06E-04
AT3G02690.1	integral membrane family protein	-3,9061	4,12E-04
AT5G54830.1	DOMON domain-containing protein / dopamine beta-monooxygenase N-terminal domain-containing protein	-3,7436	4,13E-04
AT5G65440.1	unknown protein	-3,9684	4,14E-04
AT1G71960.1	ABC transporter family protein	-3,0838	4,17E-04
AT5G24570.1	unknown protein	-2,9907	4,53E-04
AT3G50770.1	calmodulin-related protein, putative	-3,7996	4,57E-04
AT1G30835	unknown protein	-3,5399	4,58E-04
AT3G02700.1	NC domain-containing protein	-3,5534	4,63E-04
AT5G45750.1	AtRABA1c (*Arabidopsis* Rab GTPase homolog A1c); GTP binding	-3,0296	4,78E-04

AT5G38350.1	disease resistance protein (NBS-LRR class), putative	-3,9418	4,81E-04
AT5G42850.2	electron carrier/ protein disulfide oxidoreductase	-3,0607	4,89E-04
AT2G42270.1	U5 small nuclear ribonucleoprotein helicase, putative	-4,0151	4,96E-04
AT2G16970.1	MEE15 (maternal effect embryo arrest 15); tetracycline transporter	-3,1493	4,99E-04
AT1G34200.1	oxidoreductase family protein	-4,3794	5,15E-04
AT4G26320.1	AGP13 (ARABINOGALACTAN PROTEIN 13)	-3,7647	5,36E-04
AT1G65985.1	unknown protein	-4,5137	5,54E-04
AT2G02250.1	ATPP2-B2 (Phloem protein 2-B2)	3,5565	5,68E-04
AT4G01380.1	plastocyanin-like domain-containing protein	2,9168	6,00E-04
AT3G49360.1	glucosamine/galactosamine-6-phosphate isomerase family protein	-3,0395	6,16E-04
AT5G47810.1	phosphofructokinase family protein	-2,5323	6,79E-04
AT1G13420.1	sulfotransferase family protein	-2,5974	7,12E-04
AT3G61180.1	zinc finger (C3HC4-type RING finger) family protein	-2,6879	7,20E-04
AT5G17090.1	unknown protein	4,0007	7,31E-04
AT4G12060.1	Clp amino terminal domain-containing protein	-2,8079	7,33E-04
AT4G23600	CORI3 (CORONATINE INDUCED 1, JASMONIC ACID RESPONSIVE 2); transaminase	3,3147	7,61E-04
AT3G09380.1	unknown protein	3,0070	7,80E-04
AT3G03950	ECT1; protein binding	-4,5234	8,05E-04
AT5G48335.1	unknown protein	-4,4153	8,23E-04
AT4G01590	unknown protein	-2,9036	8,23E-04
AT4G32850	nPAP (NUCLEAR POLY(A) POLYMERASE); nucleotidyltransferase	-2,2502	8,47E-04
AT4G00895.1	ATP synthase delta chain-related	-3,4883	8,47E-04
AT5G02460.1	Dof-type zinc finger domain-containing protein	-3,3742	8,53E-04
AT5G47990.1	CYP705A5 (cytochrome P450, family 705, subfamily A, polypeptide 5); oxygen binding	-2,6813	8,64E-04
AT5G43150.1	unknown protein	-2,5824	8,87E-04
AT1G21350	electron carrier	-3,2276	8,94E-04
AT5G35450.1	disease resistance protein (CC-NBS-LRR class), putative	-2,4957	9,00E-04
AT4G05380.1	AAA-type ATPase family protein	3,8809	9,18E-04
AT4G01950.1	ATGPAT3/GPAT3 (GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE 3); acyltransferase	-3,6199	9,19E-04
AT1G70080.1	terpene synthase/cyclase family protein	3,5914	9,42E-04
AT1G59900.1	AT-E1 ALPHA (pyruvate dehydrogenase complex E1 alpha subunit); pyruvate dehydrogenase (acetyl-transferring)	-4,0820	9,49E-04
AT2G25670	unknown protein	-3,3603	9,54E-04
AT3G20970	NFU4 (NFU domain protein 4)	-3,2696	9,62E-04
AT1G18730	unknown protein	-2,1801	9,80E-04
AT5G40500	unknown protein	-4,9834	9,97E-04
AT4G19840.1	ATPP2-A1 (*Arabidopsis thaliana* phloem protein 2-A1)	-4,5704	1,02E-03
AT2G44200.1	unknown protein	-2,6489	1,05E-03
AT5G25010.1	unknown protein	-3,6277	1,07E-03
AT3G18380	DNA binding / sequence-specific DNA binding / transcription factor	-2,6242	1,08E-03
AT2G03710	SEP4 (SEPALLATA4); DNA binding / transcription factor	-2,8747	1,08E-03
AT2G20720.1	pentatricopeptide (PPR) repeat-containing protein	2,2956	1,12E-03
AT1G52920.1	catalytic	2,2470	1,13E-03
AT2G13800.1	ATSERK5 (SOMATIC EMBRYOGENESIS RECEPTOR LIKE KINASE 5); ATP binding / protein kinase/ transmembrane receptor protein serine/threonine kinase	-4,0948	1,14E-03
AT4G02450.1	glycine-rich protein	-4,6687	1,14E-03

AT1G31885.1	major intrinsic family protein / MIP family protein	2,4423	1,16E-03
AT4G16920.1	disease resistance protein (TIR-NBS-LRR class), putative	2,4927	1,19E-03
AT4G23496.1	SP1L5 (SPIRAL1-LIKE5)	2,3753	1,20E-03
AT5G19950	unknown protein	-2,2907	1,20E-03
AT5G45060.1	disease resistance protein (TIR-NBS-LRR class), putative	-2,7893	1,23E-03
AT5G16080.1	unknown protein	2,5046	1,25E-03
AT3G11220.1	ELO1 (ELONGATA 1)	-2,8340	1,33E-03
AT1G13220.1	nuclear matrix constituent protein-related	-2,0876	1,34E-03
AT1G27180.1	disease resistance protein (TIR-NBS-LRR class), putative	-3,3644	1,37E-03
AT2G29460.1	ATGSTU4 (GLUTATHIONE S-TRANSFERASE 22); glutathione transferase	-2,4620	1,38E-03
AT3G46490.1	oxidoreductase, 2OG-Fe(II) oxygenase family protein	-2,4484	1,45E-03
AT1G19394	unknown protein	-3,2355	1,46E-03
AT1G66540.1	cytochrome P450, putative	-4,5525	1,48E-03
AT5G05340.1	peroxidase, putative	2,3820	1,49E-03
AT2G17560	HMGB4 (HIGH MOBILITY GROUP B 4); transcription factor	-5,6827	1,50E-03
AT5G55940.1	EMB2731 (EMBRYO DEFECTIVE 2731)	-5,8991	1,50E-03
AT5G67290.1	FAD-dependent oxidoreductase family protein	-2,9252	1,51E-03
AT1G60730	aldo/keto reductase family protein	-2,4315	1,52E-03
AT4G00970.1	protein kinase family protein	3,1907	1,52E-03
AT3G32040.1	geranylgeranyl pyrophosphate synthase, putative / GGPP synthetase, putative / farnesyltranstransferase, putative	-4,5480	1,57E-03
AT5G23990.1	ATFRO5/FRO5 (FERRIC REDUCTION OXIDASE 5); ferric-chelate reductase	-2,3424	1,58E-03
AT1G59950.1	aldo/keto reductase, putative	2,3733	1,60E-03
AT1G55940.1	CYP708A1 (cytochrome P450, family 708, subfamily A, polypeptide 1); oxygen binding	-2,1523	1,62E-03
AT3G48080.1	lipase class 3 family protein / disease resistance protein-related	-2,8219	1,63E-03
AT1G80960	F-box protein-related	-4,3700	1,64E-03
AT5G56370	F-box family protein	-2,9236	1,66E-03
AT3G19370.1	unknown protein	-2,4980	1,67E-03
AT1G19720.1	pentatricopeptide (PPR) repeat-containing protein	-3,1421	1,67E-03
AT1G65350.1	UBQ13 (ubiquitin 13)	3,7791	1,68E-03
AT5G24880.1	unknown protein	-2,5868	1,71E-03
AT2G28820.1	structural constituent of ribosome	2,4505	1,74E-03
AT4G11210.1	disease resistance-responsive family protein / dirigent family protein	-4,2122	1,76E-03
AT5G66055	AKRP/EMB2036 (EMBRYO DEFECTIVE 2036); protein binding	-2,6458	1,76E-03
AT3G55020.1	RabGAP/TBC domain-containing protein	-4,1684	1,78E-03
AT4G09680.1	unknown protein	-2,9051	1,79E-03
AT2G21830.1	DC1 domain-containing protein	-3,0386	1,80E-03
AT5G44440.1	FAD-binding domain-containing protein	3,4788	1,80E-03
AT4G36430.1	peroxidase, putative	2,8264	1,86E-03
AT5G13200.1	GRAM domain-containing protein / ABA-responsive protein-related	2,1946	1,86E-03
AT1G23460.1	polygalacturonase	-3,1977	1,88E-03
AT5G15980.1	pentatricopeptide (PPR) repeat-containing protein	-3,2812	1,91E-03
AT5G48090.1	ELP1 (EDM2-LIKE PROTEIN1); zinc ion binding	2,2584	1,92E-03
AT1G69720.1	HO3 (HEME OXYGENASE 3); heme oxygenase (decyclizing)	4,8904	1,95E-03
AT4G20480.1	unknown protein	-2,8038	2,00E-03
AT5G42825.1	unknown protein	-5,2154	2,03E-03
AT2G20790	unknown protein	-3,7654	2,03E-03
AT2G15220.1	secretory protein, putative	-2,0783	2,05E-03

AT5G40910.1	disease resistance protein (TIR-NBS-LRR class), putative	-3,8898	2,06E-03
AT4G23860	PHD finger protein-related	-1,9722	2,08E-03
AT4G13520.1	SMAP1 (SMALL ACIDIC PROTEIN 1)	-2,2228	2,09E-03
AT3G01220.1	ATHB20 (*Arabidopsis Thaliana* Homeobox Protein 20); DNA binding / transcription factor	-2,7190	2,09E-03
AT2G21860.1	violaxanthin de-epoxidase-related	-5,1839	2,11E-03
AT1G60740.1	peroxiredoxin type 2, putative	-3,2898	2,12E-03
AT3G07525	autophagocytosis-associated family protein	-1,9000	2,14E-03
AT1G15660.1	unknown protein	-2,8339	2,16E-03
AT4G05320	UBQ10 (POLYUBIQUITIN 10); protein binding	-3,5510	2,17E-03
AT5G02760.1	protein phosphatase 2C family protein / PP2C family protein	2,1826	2,24E-03
AT4G39180.1	SEC14 (secretion 14)	-2,2710	2,25E-03
AT5G35730.1	EXS family protein / ERD1/XPR1/SYG1 family protein	-3,5251	2,28E-03
AT5G63180.1	pectate lyase family protein	2,6041	2,38E-03
AT4G24420.1	RNA recognition motif (RRM)-containing protein	5,5492	2,41E-03
AT4G15400.1	transferase family protein	-2,7166	2,42E-03
AT5G43270	SPL2 (SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 2); DNA binding / transcription factor	-2,0419	2,43E-03
AT4G37800.1	xyloglucan:xyloglucosyl transferase, putative / xyloglucan endotransglycosylase, putative / endo-xyloglucan transferase, putative	-5,9032	2,49E-03
AT5G24240.1	phosphatidylinositol 3- and 4-kinase family protein / ubiquitin family protein	5,3444	2,50E-03
AT4G37530	peroxidase, putative	-3,4092	2,51E-03
AT4G19240.1	unknown protein	3,6949	2,52E-03
AT4G19770.1	glycosyl hydrolase family 18 protein	2,9163	2,53E-03
AT5G01350.1	unknown protein	-3,5174	2,53E-03
AT2G18570.1	UDP-glucoronosyl/UDP-glucosyl transferase family protein	-3,4036	2,55E-03
AT1G71200.1	basic helix-loop-helix (bHLH) family protein	-1,8442	2,58E-03
AT3G27850.1	RPL12-C (RIBOSOMAL PROTEIN L12-C); structural constituent of ribosome	-2,2770	2,60E-03
AT1G68910	unknown protein	-4,5452	2,64E-03
AT1G63290.1	ribulose-phosphate 3-epimerase, cytosolic, putative / pentose-5-phosphate 3-epimerase, putative	-1,9597	2,71E-03
AT5G17190.1	unknown protein	-5,2908	2,74E-03
AT1G45474.1	LHCA5 (Photosystem I light harvesting complex gene 5)	-2,0288	2,75E-03
AT3G28200.1	peroxidase, putative	-3,4313	2,78E-03
AT2G01100	unknown protein	-2,7647	2,79E-03
AT2G46950.1	CYP709B2 (cytochrome P450, family 709, subfamily B, polypeptide 2); oxygen binding	3,8934	2,79E-03
AT5G10140.1	FLC (FLOWERING LOCUS C); transcription factor	-3,1202	2,81E-03
AT5G37175.1	unknown protein	2,2719	2,84E-03
AT4G19080.1	unknown protein	2,2986	2,87E-03
AT4G10950.1	GDSL-motif lipase/hydrolase family protein	2,5682	2,88E-03
AT5G17790.1	VAR3 (VARIEGATED 3); binding	-4,5687	2,89E-03
AT5G23980.1	ATFRO4/FRO4 (FERRIC REDUCTION OXIDASE 4); ferric-chelate reductase	-2,3233	2,90E-03
AT2G02960	zinc finger (C3HC4-type RING finger) family protein	-6,1560	2,93E-03
AT1G56120.1	leucine-rich repeat family protein / protein kinase family protein	2,3395	2,94E-03
AT2G02350.1	SKIP3 (SKP1 INTERACTING PARTNER 3)	-1,7991	2,97E-03
AT2G33340	transducin family protein / WD-40 repeat family protein	-3,0269	2,98E-03
AT1G53885.1	senescence-associated protein-related	-1,9654	2,98E-03
AT3G09160.1	RNA recognition motif (RRM)-containing protein	-2,9054	3,01E-03
AT1G09380.1	integral membrane family protein / nodulin MtN21-related	-3,4128	3,01E-03

AT4G12330.1	CYP706A7 (cytochrome P450, family 706, subfamily A, polypeptide 7); oxygen binding	-5,2842	3,03E-03
AT3G28130	nodulin MtN21 family protein	-3,8753	3,05E-03
AT4G18590.1	unknown protein	-5,4320	3,05E-03
AT5G65070.1	MAF4 (MADS AFFECTING FLOWERING 4); transcription factor	-1,9652	3,08E-03
AT2G32290.1	BMY5 (beta-amylase 5); beta-amylase	2,3837	3,11E-03
AT2G21280.1	GC1 (GIANT CHLOROPLAST 1); catalytic/ coenzyme binding	-4,4609	3,12E-03
AT5G63820.1	unknown protein	2,0468	3,13E-03
AT1G53340.1	DC1 domain-containing protein	-3,4371	3,15E-03
AT3G30350.1	unknown protein	-2,6549	3,16E-03
AT2G18670.1	zinc finger (C3HC4-type RING finger) family protein	-3,5307	3,20E-03
AT3G16370.1	GDSL-motif lipase/hydrolase family protein	2,0833	3,20E-03
AT1G67700	unknown protein	-5,7780	3,21E-03
AT4G10910.1	unknown protein	2,2029	3,30E-03
AT1G47760.1	MADS-box protein (AGL102)	-2,0735	3,32E-03
AT4G39955.1	hydrolase, alpha/beta fold family protein	-2,1802	3,32E-03
AT1G63540.1	hydroxyproline-rich glycoprotein family protein	3,7706	3,35E-03
AT1G29720.1	protein kinase family protein	-3,4161	3,36E-03
AT3G16370.1	GDSL-motif lipase/hydrolase family protein	2,1717	3,37E-03
AT1G12330.1	unknown protein	1,8859	3,40E-03
AT1G33490.1	unknown protein	-1,9021	3,40E-03
AT3G52180	SEX4 (STARCH-EXCESS 4); oxidoreductase/ polysaccharide binding / protein tyrosine/serine/threonine phosphatase	1,7829	3,41E-03
AT5G18140.1	DNAJ heat shock N-terminal domain-containing protein	-2,5759	3,45E-03
AT3G46730.1	disease resistance protein (CC-NBS class), putative	-2,0875	3,45E-03
AT1G29610.1	unknown protein	-2,4112	3,47E-03
AT1G34630.1	unknown protein	-4,1543	3,48E-03
AT3G47810	MAG1 (MAIGO 1); protein serine/threonine phosphatase	-5,2953	3,52E-03
AT4G04925.1	unknown protein	-2,2087	3,55E-03
AT3G01720.1	unknown protein	-5,9749	3,56E-03
AT5G64630	FAS2 (FASCIATA 2); nucleotide binding	2,6495	3,60E-03
AT1G64220.1	preprotein translocase-related	-2,0658	3,60E-03
AT4G19985.1	GCN5-related N-acetyltransferase (GNAT) family protein	-2,3206	3,68E-03
AT3G61390.2	U-box domain-containing protein	-3,3475	3,68E-03
AT4G13110.1	BSD domain-containing protein	-3,3276	3,74E-03
AT1G19485.1	AT hook motif-containing protein	-1,8042	3,78E-03
AT1G12650	unknown protein	-5,9890	3,90E-03
AT1G17890	GER2; catalytic	-2,0066	3,92E-03
AT3G44430.1	unknown protein	-4,3216	3,98E-03
AT4G11790.1	Ran-binding protein 1 domain-containing protein / RanBP1 domain-containing protein	-3,0733	3,98E-03
AT1G78360.1	ATGSTU21 (*Arabidopsis thaliana* Glutathione S-transferase (class tau) 21); glutathione transferase	-2,0508	4,04E-03
AT5G60230	SEN2 (SPLICING ENDONUCLEASE 2); tRNA-intron endonuclease	-3,9262	4,07E-03
AT3G24890.1	ATVAMP728 (*Arabidopsis thaliana* vesicle-associated membrane protein 728)	-2,4805	4,08E-03
AT1G63530.1	unknown protein	-2,3097	4,14E-03
AT2G39850.1	subtilase	2,8982	4,17E-03
AT5G26270.1	unknown protein	-4,5796	4,18E-03
AT2G29480.1	ATGSTU2 (GLUTATHIONE S-TRANSFERASE 20); glutathione transferase	-2,3439	4,19E-03
AT5G18020.1	auxin-responsive protein, putative	2,1067	4,20E-03
AT4G21326.1	subtilase family protein	2,4604	4,22E-03

AT2G23130.1	AGP17 (ARABINOGALACTAN PROTEIN 17)	-4,5879	4,22E-03
AT2G18660.1	expansin family protein (EXPR3)	-3,3632	4,24E-03
AT1G59700.1	ATGSTU16 (*Arabidopsis thaliana* Glutathione S-transferase (class tau) 16); glutathione transferase	-1,8487	4,25E-03
AT3G29430.1	geranylgeranyl pyrophosphate synthase, putative / GGPP synthetase, putative / farnesyltranstransferase, putative	3,4341	4,26E-03
AT5G38900.1	DSBA oxidoreductase family protein	2,1931	4,27E-03
AT2G06025.1	GCN5-related N-acetyltransferase (GNAT) family protein	-3,0458	4,28E-03
AT5G37580.1	binding / protein binding	-2,7632	4,30E-03
AT3G27990.1	unknown protein	-1,8467	4,30E-03
AT1G16540.1	ABA3/ATABA3/LOS5/SIR3 (ABA DEFICIENT 3); Mo-molybdopterin cofactor sulfurase/ selenocysteine lyase	-2,8627	4,31E-03
AT5G39080.1	transferase family protein	-6,0403	4,33E-03
AT4G05040	ankyrin repeat family protein	-2,3994	4,36E-03
AT4G11830	phospholipase D gamma 2 / PLD gamma 2 (PLDGAMMA2)	-2,2698	4,41E-03
AT1G43160.1	RAP2.6 (related to AP2 6); DNA binding / transcription factor	-2,3000	4,42E-03
AT3G03790	ankyrin repeat family protein / regulator of chromosome condensation (RCC1) family protein	-2,0901	4,45E-03
AT5G23840	MD-2-related lipid recognition domain-containing protein / ML domain-containing protein	-2,8264	4,46E-03
AT1G58602.1	disease resistance protein (CC-NBS-LRR class), putative	2,1463	4,47E-03
AT1G65430.1	zinc finger protein-related	-4,4331	4,47E-03
AT2G03220.1	FT1 (FUCOSYLTRANSFERASE 1); fucosyltransferase/ transferase, transferring glycosyl groups	-1,6514	4,48E-03
AT1G58520.1	RXW8	1,8250	4,49E-03
AT4G05060.1	vesicle-associated membrane family protein / VAMP family protein	-2,3313	4,49E-03
AT5G23570.1	SGS3 (SUPPRESSOR OF GENE SILENCING 3)	2,2026	4,53E-03
AT1G51420.1	sucrose-phosphatase, putative	-2,4292	4,54E-03
AT1G22150.1	SULTR1;3 (sulfate transporter); sulfate transporter	2,6517	4,61E-03
AT3G29670.1	transferase family protein	-2,8593	4,65E-03
AT5G05890.1	UDP-glucoronosyl/UDP-glucosyl transferase family protein	-5,5822	4,67E-03
AT1G80130.1	binding	2,5365	4,69E-03
AT3G01170.1	structural constituent of ribosome	-2,0228	4,73E-03
AT1G78460.1	SOUL heme-binding family protein	-1,9994	4,78E-03
AT4G15280.1	UDP-glucoronosyl/UDP-glucosyl transferase family protein	2,0068	4,84E-03
AT5G45080.1	ATPP2-A6 (Phloem protein 2-A6); transmembrane receptor	-3,2379	4,91E-03
AT3G50560.1	short-chain dehydrogenase/reductase (SDR) family protein	-2,2457	4,92E-03
AT1G57630.1	disease resistance protein (TIR class), putative	-2,8796	4,96E-03
AT3G26290.1	CYP71B26 (cytochrome P450, family 71, subfamily B, polypeptide 26); oxygen binding	-1,9781	4,97E-03
AT5G49820.1	EMB1879 (EMBRYO DEFECTIVE 1879)	-4,0729	4,99E-03
AT4G01860	transducin family protein / WD-40 repeat family protein	-1,6106	5,00E-03
AT4G08480.1	MAPKKK9 (Mitogen-activated protein kinase kinase kinase 9); kinase	-3,8072	5,01E-03
AT1G74790.1	unknown protein	-5,5952	5,05E-03
AT3G26180	CYP71B20 (cytochrome P450, family 71, subfamily B, polypeptide 20); oxygen binding	-3,5613	5,10E-03
AT1G61180.1	disease resistance protein (CC-NBS-LRR class), putative	-5,2324	5,11E-03
AT2G40610.1	ATEXPA8 (*ARABIDOPSIS THALIANA* EXPANSIN A8)	1,8405	5,12E-03
AT1G15890.1	disease resistance protein (CC-NBS-LRR class), putative	-4,0821	5,14E-03
AT1G33790.1	jacalin lectin family protein	-1,7323	5,19E-03
AT1G53980.1	polyubiquitin-related	-3,4329	5,20E-03
AT2G14060.1	S-adenosyl-L-methionine:carboxyl methyltransferase family protein	-2,5999	5,24E-03

AT1G71400.1	disease resistance family protein / LRR family protein	-5,5137	5,25E-03
AT1G66980.1	protein kinase family protein / glycerophosphoryl diester phosphodiesterase family protein	2,6214	5,28E-03
AT1G13460	serine/threonine protein phosphatase 2A (PP2A) regulatory subunit B', putative	-5,6168	5,30E-03
AT1G70830	Bet v I allergen family protein	2,4372	5,35E-03
AT3G03280.1	unknown protein	-2,9346	5,42E-03
AT4G20690.1	unknown protein	2,0659	5,43E-03
AT5G40950.1	50S ribosomal protein L27, chloroplast, putative (RPL27)	-4,4143	5,48E-03
AT5G48530.1	unknown protein	-2,9215	5,49E-03
AT4G15390.1	transferase family protein	-2,6873	5,51E-03
AT5G17910.1	unknown protein	-4,0921	5,55E-03
AT5G10230.1	ANN7 (ANN7, ANNEXIN *ARABIDOPSIS* 7); calcium ion binding / calcium-dependent phospholipid binding	-2,0627	5,58E-03
AT4G19090.1	unknown protein	-1,9483	5,67E-03
AT4G02920	unknown protein	-4,6657	5,71E-03
AT5G43470	RPP8 (RECOGNITION OF PERONOSPORA PARASITICA 8)	-4,9667	5,72E-03
AT5G03200.1	zinc finger (C3HC4-type RING finger) family protein	-5,7551	5,73E-03
AT5G55570.1	unknown protein	2,1101	5,75E-03
AT3G44300.1	NIT2 (NITRILASE 2)	-3,1381	5,76E-03
AT5G26660.1	ATMYB4 (myb domain protein 4); transcriptional repressor	-1,8823	5,93E-03
AT4G12030	bile acid:sodium symporter family protein	-2,1065	5,95E-03
AT4G08780.1	peroxidase, putative	1,8941	5,98E-03
AT5G24750.1	unknown protein	-2,1627	5,99E-03
AT1G29800	zinc ion binding	-5,7821	6,03E-03
AT5G20730	NPH4 (NON-PHOTOTROPHIC HYPOCOTYL); transcription factor	-2,1444	6,08E-03
AT5G23020.1	MAM-L (METHYLTHIOALKYMALATE SYNTHASE-LIKE); 2-isopropylmalate synthase	-1,5965	6,20E-03
AT2G44290.1	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein (YLS3)	-2,0432	6,21E-03
AT1G05600.1	pentatricopeptide (PPR) repeat-containing protein	-2,2422	6,25E-03
AT2G45730.1	eukaryotic initiation factor 3 gamma subunit family protein	-1,6994	6,27E-03
AT2G46330	AGP16 (ARABINOGALACTAN PROTEIN 16)	-3,7972	6,29E-03
AT5G39640.1	unknown protein	2,1108	6,29E-03
AT3G58940.1	F-box family protein	2,0345	6,42E-03
AT5G36220.1	CYP81D1 (CYTOCHROME P450 91A1); oxygen binding	-2,2535	6,44E-03
AT5G13890	unknown protein	-2,5683	6,47E-03
AT4G18250.1	receptor serine/threonine kinase, putative	-2,0158	6,51E-03
AT2G44650.1	CHL-CPN10 (chloroplast chaperonin 10)	-1,9047	6,63E-03
AT1G30900.1	vacuolar sorting receptor, putative	-1,7100	6,66E-03
AT2G34840.1	coatomer protein epsilon subunit family protein / COPE family protein	-4,6328	6,68E-03
AT5G25250.1	unknown protein	-6,4048	6,69E-03
AT1G17240.1	leucine-rich repeat family protein	-2,7022	6,74E-03
AT5G27640	TIF3B1 (EUKARYOTIC TRANSLATION INITIATION FACTOR 3B); nucleic acid binding / translation initiation factor	-1,6795	6,78E-03
AT2G30670.1	tropinone reductase, putative / tropine dehydrogenase, putative	2,9864	6,78E-03
AT5G43740	disease resistance protein (CC-NBS-LRR class), putative	-4,7144	6,83E-03
AT1G14960.1	major latex protein-related / MLP-related	-2,0480	6,96E-03
AT5G44920	Toll-Interleukin-Resistance (TIR) domain-containing protein	-1,8132	7,16E-03
AT3G49810.1	U-box domain-containing protein	-1,9383	7,21E-03
AT3G03520.1	phosphoesterase family protein	-1,5286	7,27E-03
AT4G16880.1	disease resistance protein-related	-3,5161	7,27E-03

AT5G23390.1	unknown protein	-3,6394	7,34E-03
AT3G50570.1	hydroxyproline-rich glycoprotein family protein	-2,1692	7,35E-03
AT1G68340.1	unknown protein	-1,5294	7,37E-03
AT1G80940	unknown protein	-2,9537	7,38E-03
AT2G43520.1	ATTI2 (*ARABIDOPSIS THALIANA* TRYPSIN INHIBITOR PROTEIN 2); trypsin inhibitor	2,1260	7,42E-03
AT3G29410.1	terpene synthase/cyclase family protein	5,1104	7,44E-03
AT3G03480.1	CHAT (ACETYL COA:(Z)-3-HEXEN-1-OL ACETYLTRANSFERASE); acetyl CoA:(Z)-3-hexen-1-ol acetyltransferase	2,2567	7,49E-03
AT5G19990.1	ATSUG1; ATPase	-4,7439	7,50E-03
AT1G30500	CCAAT-binding transcription factor (CBF-B/NF-YA) family protein	-2,8076	7,50E-03
AT2G02240.1	MEE66 (maternal effect embryo arrest 66)	2,2997	7,62E-03
AT5G65690.1	phosphoenolpyruvate carboxykinase (ATP), putative / PEP carboxykinase, putative / PEPCK, putative	-1,7286	7,65E-03
AT4G13030	unknown protein	-2,1066	7,80E-03
AT1G05150.1	calcium-binding EF hand family protein	-1,7134	7,80E-03
AT4G15340.1	ATPEN1 (*Arabidopsis thaliana* pentacyclic triterpene synthase 1); catalytic/ lyase	-2,6247	7,84E-03
AT4G26150.1	zinc finger (GATA type) family protein	-1,4917	7,85E-03
AT4G03960.1	tyrosine specific protein phosphatase family protein	-2,6571	7,86E-03
AT2G42170.1	actin, putative	-3,5789	8,05E-03
AT1G27540.1	F-box family protein	2,9244	8,08E-03
AT4G22460.1	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein	-2,3286	8,09E-03
AT4G14370.1	disease resistance protein (TIR-NBS-LRR class), putative	-2,5562	8,12E-03
AT5G37290.1	armadillo/beta-catenin repeat family protein	-1,6817	8,18E-03
AT1G65280.1	heat shock protein binding / unfolded protein binding	-6,3987	8,20E-03
AT4G11000.1	ankyrin repeat family protein	3,4533	8,24E-03
AT4G22280	F-box family protein	1,7332	8,31E-03
AT4G22235	unknown protein	2,0963	8,50E-03
AT2G40920	F-box family protein	-3,5205	8,53E-03
AT5G20790.1	unknown protein	-2,9374	8,53E-03
AT4G11830	phospholipase D gamma 2 / PLD gamma 2 (PLDGAMMA2)	-2,0705	8,56E-03
AT5G18030.1	auxin-responsive protein, putative	2,3014	8,59E-03
AT3G21810.1	zinc finger (CCCH-type) family protein	-2,6776	8,65E-03
AT1G66130.1	oxidoreductase N-terminal domain-containing protein	-5,3580	8,67E-03
AT5G08370	ATAGAL2 (*ARABIDOPSIS THALIANA* ALPHA-GALACTOSIDASE 2); alpha-galactosidase	-3,0034	8,71E-03
AT4G23130	CRK5 (CYSTEINE-RICH RLK5); kinase	2,0928	8,76E-03
AT3G47360.1	short-chain dehydrogenase/reductase (SDR) family protein	-1,5726	8,78E-03
AT4G25900.1	aldose 1-epimerase family protein	-2,4504	8,84E-03
AT4G31570.1	unknown protein	2,2756	8,85E-03
AT3G44290.1	ANAC060 (*Arabidopsis* NAC domain containing protein 60); transcription factor	-2,5865	8,88E-03
AT3G48510.1	unknown protein	2,0465	8,95E-03
AT1G09590.1	60S ribosomal protein L21 (RPL21A)	-1,9346	9,07E-03
AT4G01350.1	DC1 domain-containing protein	-4,8380	9,09E-03
AT1G63130.1	pentatricopeptide (PPR) repeat-containing protein	-2,8823	9,20E-03
AT2G02850.1	ARPN (PLANTACYANIN); copper ion binding	1,6048	9,25E-03
AT1G16420.1	latex-abundant protein, putative (AMC8) / caspase family protein	-1,5331	9,35E-03
AT3G48720.1	transferase family protein	1,5794	9,44E-03

AT3G44740.1	tRNA synthetase class II (G, H, P and S) family protein	-3,5289	9,44E-03
AT2G22350.1	RNase H domain-containing protein	1,7375	9,53E-03
AT4G00390.1	transcription regulator	-2,7190	9,55E-03
AT4G18910.1	NIP1;2/NLM2 (NOD26-like intrinsic protein 1;2); water channel	-2,8308	9,57E-03
AT2G14510.1	leucine-rich repeat protein kinase, putative	-1,9599	9,60E-03
AT5G26300.1	meprin and TRAF homology domain-containing protein / MATH domain-containing protein	1,4389	9,61E-03
AT1G28610	GDSL-motif lipase, putative	1,6504	9,63E-03
AT3G45410.1	lectin protein kinase family protein	-1,4643	9,67E-03
AT4G29760.1	unknown protein	4,9467	9,73E-03
AT1G13250.1	GATL3 (Galacturonosyltransferase-like 3); polygalacturonate 4-alpha-galacturonosyltransferase/ transferase, transferring glycosyl groups / transferase, transferring hexosyl groups	1,7652	9,77E-03
AT4G18970.1	GDSL-motif lipase/hydrolase family protein	1,6080	9,84E-03
AT1G65295.1	unknown protein	-4,4410	9,85E-03
AT1G28340.1	leucine-rich repeat family protein	-1,8108	1,01E-02
AT5G45420.1	myb family transcription factor	-1,5935	1,02E-02
AT5G55550	RNA recognition motif (RRM)-containing protein	-2,3504	1,02E-02
AT5G09930.1	ATGCN2 (*Arabidopsis thaliana* general control non-repressible 2)	-1,7314	1,02E-02
AT1G49660.1	unknown protein	-5,6872	1,02E-02
AT3G27420.1	unknown protein	-5,4778	1,03E-02
AT1G22650.1	beta-fructofuranosidase, putative / invertase, putative / saccharase, putative / beta-fructosidase, putative	-1,6803	1,03E-02
AT4G20030.1	RNA recognition motif (RRM)-containing protein	-1,9326	1,03E-02
AT5G37350	RIO1 family protein	-1,5702	1,04E-02
AT1G70810.1	C2 domain-containing protein	-1,7073	1,04E-02
AT4G17060.1	unknown protein	-1,7309	1,05E-02
AT2G17430.1	ATMLO7/MLO7 (MILDEW RESISTANCE LOCUS O 7); calmodulin binding	-3,4478	1,05E-02
AT4G25400.1	basix helix-loop-helix (bHLH) family protein	-1,8900	1,06E-02
AT2G17420.1	NTRA (NADPH-dependent thioredoxin reductase 2)	-1,4930	1,08E-02
AT5G43440	2-oxoglutarate-dependent dioxygenase, putative	-2,5393	1,09E-02
AT4G01910.1	DC1 domain-containing protein	-5,3319	1,09E-02
AT5G64190.1	unknown protein	-2,3535	1,10E-02
AT5G65870.1	ATPSK5 (PHYTOSULFOKINE 5 PRECURSOR); growth factor	-2,3433	1,10E-02
AT3G46340.1	leucine-rich repeat protein kinase, putative	-1,9065	1,11E-02
AT5G24680.1	unknown protein	-3,8591	1,11E-02
AT3G07800.1	thymidine kinase, putative	-5,1280	1,12E-02
AT5G42030.1	ABIL4 (ABL INTERACTOR-LIKE PROTEIN 4)	-4,3293	1,12E-02
AT2G30750.1	CYP71A12 (cytochrome P450, family 71, subfamily A, polypeptide 12); oxygen binding	2,9997	1,12E-02
AT3G26080.1	plastid-lipid associated protein PAP / fibrillin family protein	-1,8333	1,12E-02
AT1G64780.1	ATAMT1;2 (AMMONIUM TRANSPORTER 1;2); ammonium transporter	-2,1929	1,12E-02
AT2G28950.1	ATEXPA6 (*ARABIDOPSIS THALIANA* EXPANSIN A6)	1,6384	1,12E-02
AT5G27100.1	ATGLR2.1 (Arabidopsis thaliana glutamate receptor 2.1)	-4,4036	1,13E-02
AT5G59530.1	2-oxoglutarate-dependent dioxygenase, putative	-2,2291	1,13E-02
AT1G27540.1	F-box family protein	-4,9205	1,14E-02
AT1G29790.1	unknown protein	-5,7244	1,14E-02
AT3G48850.1	mitochondrial phosphate transporter, putative	2,0933	1,15E-02
AT1G70260.1	nodulin MtN21 family protein	-3,5122	1,17E-02
AT1G67360	rubber elongation factor (REF) family protein	-1,5066	1,17E-02
AT3G45443.1	unknown protein	-1,7668	1,17E-02

AT5G24210.1	lipase class 3 family protein	-3,9359	1,17E-02
AT4G02210.1	unknown protein	-1,9021	1,17E-02
AT4G00650.1	FRI (FRIGIDA)	-2,6742	1,17E-02
AT1G49030.1	unknown protein	-1,7368	1,18E-02
AT4G01390.1	meprin and TRAF homology domain-containing protein / MATH domain-containing protein	2,4626	1,18E-02
AT4G14305.1	unknown protein	-2,3218	1,19E-02
AT2G30920.1	ATCOQ3 (EMBRYO DEFECTIVE 3002)	-1,7089	1,19E-02
AT1G20696	HMGB3 (HIGH MOBILITY GROUP B 3); transcription factor	-2,2017	1,19E-02
AT1G12160.1	flavin-containing monooxygenase family protein / FMO family protein	-1,5608	1,20E-02
AT2G34050.1	unknown protein	-3,2526	1,20E-02
AT4G39480.1	CYP96A9 (cytochrome P450, family 96, subfamily A, polypeptide 9); oxygen binding	2,1493	1,21E-02
AT2G02360.1	ATPP2-B10 (Phloem protein 2-B10)	-6,0408	1,22E-02
AT5G38590	F-box family protein	1,7919	1,24E-02
AT5G43660.1	oxidoreductase, acting on paired donors, with incorporation or reduction of molecular oxygen, 2-oxoglutarate as one donor, and incorporation of one atom each of oxygen into both donors	2,7915	1,25E-02
AT1G16390.1	organic cation transporter-related	-1,9565	1,25E-02
AT5G48620.1	disease resistance protein (CC-NBS-LRR class), putative	-5,6349	1,25E-02
AT5G49830.1	unknown protein	-5,8291	1,26E-02
AT3G22860.1	TIF3C2 (eukaryotic translation initiation factor 3 subunit C2); translation initiation factor	2,1051	1,26E-02
AT2G06005	unknown protein	-4,5752	1,27E-02
AT5G28020	ATCYSD2 (*Arabidopsis thaliana* cysteine synthase D2); cysteine synthase	-1,7316	1,27E-02
AT5G48530.1	unknown protein	-2,2012	1,27E-02
AT2G39880.1	MYB25 (myb domain protein 25); DNA binding / transcription factor	1,8896	1,28E-02
AT1G66040.1	zinc finger (C3HC4-type RING finger) family protein	-2,5592	1,29E-02
AT4G39330.1	mannitol dehydrogenase, putative	1,6155	1,30E-02
AT5G04885.1	glycosyl hydrolase family 3 protein	-1,4542	1,31E-02
AT3G19515.2	binding	-2,3254	1,33E-02
AT5G07770.1	formin homology 2 domain-containing protein / FH2 domain-containing protein	-1,6133	1,34E-02
AT5G18050.1	auxin-responsive protein, putative	2,7317	1,35E-02
AT3G27440.1	uracil phosphoribosyltransferase, putative / UMP pyrophosphorylase, putative / UPRTase, putative	-1,4290	1,36E-02
AT2G14880.1	SWIB complex BAF60b domain-containing protein	-2,2515	1,37E-02
AT2G44260	unknown protein	-5,4006	1,38E-02
AT1G66440.1	DC1 domain-containing protein	-1,4884	1,38E-02
AT4G01935.1	unknown protein	2,5248	1,38E-02
AT4G00880.1	auxin-responsive family protein	-3,0423	1,38E-02
AT3G53370.1	DNA-binding S1FA family protein	-1,7019	1,39E-02
AT5G42590.1	CYP71A16 (cytochrome P450, family 71, subfamily A, polypeptide 16); oxygen binding	-4,4431	1,40E-02
AT1G21250.1	WAK1 (CELL WALL-ASSOCIATED KINASE); kinase	2,7017	1,40E-02
AT3G59480.1	pfkB-type carbohydrate kinase family protein	-1,4272	1,42E-02
AT5G65350.1	histone H3	-2,6246	1,43E-02
AT4G01880.1	unknown protein	-3,4757	1,44E-02
AT1G33811.1	GDSL-motif lipase/hydrolase family protein	1,6787	1,45E-02
AT5G05730.1	ASA1 (ANTHRANILATE SYNTHASE ALPHA SUBUNIT 1); anthranilate synthase	1,5423	1,45E-02
AT5G08460.1	GDSL-motif lipase/hydrolase family protein	2,6269	1,46E-02

AT3G45300.1	IVD (ISOVALERYL-COA-DEHYDROGENASE)	-3,2439	1,46E-02
AT4G15160.1	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein	1,8104	1,47E-02
AT3G44630	disease resistance protein RPP1-WsB-like (TIR-NBS-LRR class), putative	-5,4895	1,47E-02
AT2G25490.1	EBF1 (EIN3-BINDING F BOX PROTEIN 1); ubiquitin-protein ligase	-3,3443	1,47E-02
AT2G29300.1	tropinone reductase, putative / tropine dehydrogenase, putative	1,5459	1,47E-02
AT3G45440.1	lectin protein kinase family protein	1,4545	1,47E-02
AT5G23770.1	agenet domain-containing protein	2,0525	1,48E-02
AT4G37320.1	CYP81D5 (cytochrome P450, family 81, subfamily D, polypeptide 5); oxygen binding	-1,5651	1,49E-02
AT1G48690.1	auxin-responsive GH3 family protein	-3,7337	1,51E-02
AT3G05150.1	sugar transporter family protein	-1,3747	1,51E-02
AT1G33470	RNA recognition motif (RRM)-containing protein	-1,6290	1,51E-02
AT1G71340.1	glycerophosphoryl diester phosphodiesterase family protein	-1,5118	1,52E-02
AT5G23010.1	MAM1 (2-isopropylmalate synthase 3); 2-isopropylmalate synthase	-4,8278	1,53E-02
AT2G44390.1	DC1 domain-containing protein	-2,3173	1,53E-02
AT2G14095.1	unknown protein	1,7743	1,53E-02
AT2G43590.1	chitinase, putative	-1,9865	1,54E-02
AT1G54040	ESP (EPITHIOSPECIFIER PROTEIN)	5,4663	1,55E-02
AT1G33520.1	MOS2 (MODIFIER OF SNC1, 2); RNA binding / nucleic acid binding / protein binding	-1,7534	1,55E-02
AT3G60480.1	unknown protein	-4,6529	1,56E-02
AT5G33210.1	SRS8 (SHI-RELATED SEQUENCE 8)	1,3972	1,57E-02
AT5G64170	dentin sialophosphoprotein-related	-1,5903	1,58E-02
AT1G51880.1	leucine-rich repeat protein kinase, putative	-2,0604	1,59E-02
AT2G15790.1	SQN (SQUINT)	2,8424	1,60E-02
AT2G02950.1	PKS1 (PHYTOCHROME KINASE SUBSTRATE 1)	-1,3956	1,61E-02
AT3G48330	PIMT1 (PROTEIN-L-ISOASPARTATE METHYLTRANSFERASE 1); protein-L-isoaspartate (D-aspartate) O-methyltransferase	-1,3658	1,61E-02
AT1G22400.1	UGT85A1 (UDP-glucosyl transferase 85A1); UDP-glycosyltransferase/ transferase, transferring glycosyl groups / transferase, transferring hexosyl groups	-1,6142	1,61E-02
AT2G36870.1	xyloglucan:xyloglucosyl transferase, putative / xyloglucan endotransglycosylase, putative / endo-xyloglucan transferase, putative	1,4756	1,61E-02
AT4G27070.1	TSB2 (TRYPTOPHAN SYNTHASE BETA-SUBUNIT); tryptophan synthase	1,3627	1,62E-02
AT4G01630.1	ATEXPA17 (*ARABIDOPSIS THALIANA* EXPANSIN A17)	-1,3860	1,62E-02
AT2G05830	eukaryotic translation initiation factor 2B family protein / eIF-2B family protein	-6,1940	1,63E-02
AT5G50400.1	ATPAP27/PAP27 (purple acid phosphatase 27); acid phosphatase/ protein serine/threonine phosphatase	-1,4545	1,64E-02
AT3G47050	glycosyl hydrolase family 3 protein	1,3219	1,65E-02
AT3G21600	senescence/dehydration-associated protein-related	-2,2932	1,65E-02
AT5G46500.1	unknown protein	-1,6805	1,66E-02
AT3G23550.1	MATE efflux family protein	-1,9653	1,66E-02
AT2G17590.1	DC1 domain-containing protein	-2,3657	1,67E-02
AT4G04810.1	methionine sulfoxide reductase domain-containing protein / SeIR domain-containing protein	-3,2992	1,68E-02
AT1G59780.1	disease resistance protein (CC-NBS-LRR class), putative	-2,9862	1,68E-02
AT1G70880.1	Bet v I allergen family protein	-1,6036	1,69E-02

AT1G42560.1	ATMLO9/MLO9 (MILDEW RESISTANCE LOCUS O 9); calmodulin binding	-1,8027	1,69E-02
AT3G28430.1	unknown protein	-1,6466	1,70E-02
AT2G01430.1	homeobox-leucine zipper protein 17 (HB-17) / HD-ZIP transcription factor 17	-1,4188	1,70E-02
AT1G58410.1	disease resistance protein (CC-NBS-LRR class), putative	-5,2384	1,70E-02
AT5G27220.1	protein transport protein-related	1,5570	1,70E-02
AT5G49780.1	leucine-rich repeat transmembrane protein kinase, putative	-1,3067	1,71E-02
AT2G25260.1	unknown protein	-5,6205	1,71E-02
AT4G08460	unknown protein	-1,9470	1,72E-02
AT4G13810.1	disease resistance family protein / LRR family protein	-3,0561	1,73E-02
AT5G15920.1	structural maintenance of chromosomes (SMC) family protein (MSS2)	-1,3512	1,73E-02
AT3G46400.1	leucine-rich repeat protein kinase, putative	-2,7284	1,73E-02
AT1G63880.1	disease resistance protein (TIR-NBS-LRR class), putative	-6,2092	1,74E-02
AT2G25450.1	2-oxoglutarate-dependent dioxygenase, putative	-6,0261	1,74E-02
AT5G39530.1	unknown protein	-1,6439	1,74E-02
AT1G80760.1	NIP6;1 (NOD26-like intrinsic protein 6;1); water channel	-4,7401	1,74E-02
AT4G01525.1	unknown protein	1,9721	1,76E-02
AT4G09760	choline kinase, putative	-1,4551	1,78E-02
AT3G26830.1	PAD3 (PHYTOALEXIN DEFICIENT 3); oxygen binding	-2,0968	1,78E-02
AT4G37410.1	CYP81F4 (cytochrome P450, family 81, subfamily F, polypeptide 4); oxygen binding	-3,3089	1,79E-02
AT1G17950.1	MYB52 (myb domain protein 52); DNA binding / transcription factor	-1,6978	1,79E-02
AT5G53990.1	glycosyltransferase family protein	1,4301	1,80E-02
AT5G41170.1	pentatricopeptide (PPR) repeat-containing protein	-1,9107	1,80E-02
AT1G59660.1	nucleoporin family protein	-1,6027	1,81E-02
AT5G14650.1	polygalacturonase, putative / pectinase, putative	-1,4640	1,82E-02
AT3G26730.1	zinc finger (C3HC4-type RING finger) family protein	-2,1303	1,82E-02
AT4G04830.1	methionine sulfoxide reductase domain-containing protein / SeIR domain-containing protein	-1,8417	1,83E-02
AT1G24490.1	ALB4 (ALBINA 4)	-3,6138	1,83E-02
AT2G23030.1	SNRK2-9/SnRK2.9 (SNF1-RELATED PROTEIN KINASE 2-9, SNF1-RELATED PROTEIN KINASE 2.9); kinase	-1,3986	1,84E-02
AT3G46480.1	oxidoreductase, 2OG-Fe(II) oxygenase family protein	-2,1370	1,85E-02
AT5G45260.1	SLH1 (sensitive to low humidity 1); transcription factor	-3,6654	1,85E-02
AT2G04380.1	unknown protein	-3,6739	1,85E-02
AT5G42960.1	unknown protein	-1,4492	1,85E-02
AT4G11810.1	SPX (SYG1/Pho81/XPR1) domain-containing protein	1,5422	1,85E-02
AT2G37120.1	DNA-binding S1FA family protein	-5,9968	1,86E-02
AT1G58150.1	unknown protein	-3,7404	1,88E-02
AT2G19200.1	unknown protein	-1,4080	1,88E-02
AT4G16850.1	unknown protein	-1,7883	1,89E-02
AT1G52000.1	jacalin lectin family protein	3,7063	1,89E-02
AT5G50130	short-chain dehydrogenase/reductase (SDR) family protein	-1,3711	1,89E-02
AT3G43210.1	TES (TETRASPORE); microtubule motor	-1,5841	1,90E-02
AT4G34120.1	LEJ1 (LOSS OF THE TIMING OF ET AND JA BIOSYNTHESIS 1)	-2,1321	1,90E-02
AT3G26100	regulator of chromosome condensation (RCC1) family protein	-3,5404	1,90E-02
AT5G38790.1	unknown protein	-2,6369	1,91E-02
AT2G24960.1	DNA binding	-1,5441	1,91E-02
AT5G25130.1	CYP71B12 (cytochrome P450, family 71, subfamily B, polypeptide 12); oxygen binding	-1,3762	1,91E-02

AT1G44180.1	aminoacylase, putative / N-acyl-L-amino-acid amidohydrolase, putative	-1,9462	1,92E-02
AT2G38620	CDKB1;2 (cyclin-dependent kinase B1;2); kinase	-2,8198	1,92E-02
AT5G27660.1	serine-type peptidase/ trypsin	1,7848	1,92E-02
AT5G39100.1	GLP6 (GERMIN-LIKE PROTEIN 6); manganese ion binding / metal ion binding / nutrient reservoir	2,0916	1,93E-02
AT5G44575.1	unknown protein	-5,4212	1,93E-02
AT5G37770.1	TCH2 (TOUCH 2); calcium ion binding	-3,7020	1,93E-02
AT2G02630.1	DC1 domain-containing protein	-1,8961	1,95E-02
AT3G60960.1	pentatricopeptide (PPR) repeat-containing protein	-1,8628	1,95E-02
AT5G24910.1	CYP714A1 (cytochrome P450, family 714, subfamily A, polypeptide 1); oxygen binding	-1,3009	1,97E-02
AT1G79080.1	pentatricopeptide (PPR) repeat-containing protein	-1,2981	1,97E-02
AT5G43860.1	ATCLH2 (Chlorophyll-chlorophyllido hydrolase 2)	-1,8028	1,98E-02
AT3G14650.1	CYP72A11 (cytochrome P450, family 72, subfamily A, polypeptide 11); oxygen binding	-5,3748	1,98E-02
AT5G40560.1	DEGP13 (DEGP PROTEASE 13); serine-type peptidase/ trypsin	-1,4379	1,99E-02
AT1G64710	alcohol dehydrogenase, putative	1,7491	2,00E-02
AT1G50520.1	CYP705A27 (cytochrome P450, family 705, subfamily A, polypeptide 27); oxygen binding	-5,9266	2,00E-02
AT1G34200.1	oxidoreductase family protein	-5,6347	2,00E-02
AT3G13480.1	unknown protein	-1,3585	2,00E-02
AT4G12170.1	thioredoxin family protein	5,5424	2,01E-02
AT5G24140.1	SQP2 (Squalene monooxygenase 2); oxidoreductase	-4,7428	2,02E-02
AT2G17580.1	polynucleotide adenylyltransferase family protein	-5,4642	2,02E-02
AT5G52070.1	agenet domain-containing protein	-6,3583	2,02E-02
AT3G30340.1	nodulin MtN21 family protein	1,8713	2,03E-02
AT1G23560.1	unknown protein	-3,7950	2,03E-02
AT2G24710.1	ATGLR2.3 (*Arabidopsis thaliana* glutamate receptor 2.3)	-2,8844	2,05E-02
AT2G25510.1	unknown protein	-5,1335	2,05E-02
AT2G41480.1	peroxidase	-1,7252	2,06E-02
AT2G25140.1	CLPB-M/CLPB4/HSP98.7 (HEAT SHOCK PROTEIN 98.7); ATP binding / ATPase	-4,5166	2,07E-02
AT2G11270.1	citrate synthase-related	-1,5970	2,08E-02
AT5G06530	ABC transporter family protein	-5,4222	2,09E-02
AT5G46780	VQ motif-containing protein	-5,2954	2,09E-02
AT1G62760.1	invertase/pectin methylesterase inhibitor family protein	-1,4321	2,10E-02
AT1G23440.1	pyrrolidone-carboxylate peptidase family protein	-1,8653	2,10E-02
AT2G21550.1	bifunctional dihydrofolate reductase-thymidylate synthase, putative / DHFR-TS, putative	-2,7440	2,10E-02
AT2G20750.1	ATEXPB1 (*ARABIDOPSIS THALIANA* EXPANSIN B1)	-1,7537	2,11E-02
AT5G65000	nucleotide-sugar transporter family protein	-1,5313	2,13E-02
AT1G52150	ATHB-15 (INCURVATA 4); DNA binding / transcription factor	-1,8298	2,14E-02
AT5G65080	MAF5 (MADS AFFECTING FLOWERING 5); transcription factor	-2,0160	2,14E-02
AT4G14690.1	ELIP2 (EARLY LIGHT-INDUCIBLE PROTEIN 2); chlorophyll binding	-2,6626	2,15E-02
AT4G37700.1	unknown protein	-2,5153	2,16E-02
AT5G47780.1	GAUT4 (Galacturonosyltransferase 4); polygalacturonate 4-alpha-galacturonosyltransferase/ transferase, transferring glycosyl groups	-1,8462	2,16E-02
AT5G67280.1	leucine-rich repeat transmembrane protein kinase, putative	1,7228	2,17E-02
AT2G18620.1	geranylgeranyl pyrophosphate synthase, putative / GGPP synthetase, putative / farnesyltranstransferase, putative	2,7374	2,17E-02
AT5G49500.1	signal recognition particle 54 kDa protein 2 / SRP54 (SRP-54B)	-1,7029	2,17E-02

AT1G10110.1	F-box family protein	1,8132	2,18E-02
AT3G21720.1	isocitrate lyase, putative	-1,7300	2,19E-02
AT5G03495.1	nucleotide binding	-2,4569	2,19E-02
AT3G45860.1	receptor-like protein kinase, putative	-1,4975	2,20E-02
AT1G09910.1	lyase	-1,7084	2,20E-02
AT3G54040.1	photoassimilate-responsive protein-related	-1,4262	2,20E-02
AT2G45403.1	unknown protein	-2,1556	2,20E-02
AT5G18080.1	auxin-responsive protein, putative	2,6390	2,22E-02
AT3G32930.1	unknown protein	-5,8099	2,23E-02
AT2G26370.1	MD-2-related lipid recognition domain-containing protein / ML domain-containing protein	-2,8282	2,24E-02
AT2G16650.1	unknown protein	-1,4278	2,25E-02
AT4G15010	mitochondrial substrate carrier family protein	-4,7210	2,25E-02
AT3G16740.1	F-box family protein	-5,6847	2,25E-02
AT4G03070.1	AOP1 (2-oxoglutarate?dependent dioxygenase 1.1); oxidoreductase, acting on paired donors, with incorporation or reduction of molecular oxygen, 2-oxoglutarate as one donor, and incorporation of one atom each of oxygen into both donors	1,4516	2,26E-02
AT2G30700.1	unknown protein	-2,2599	2,26E-02
AT5G54300.1	unknown protein	-2,0840	2,27E-02
AT5G09970.1	CYP78A7 (cytochrome P450, family 78, subfamily A, polypeptide 7); oxygen binding	1,4741	2,27E-02
AT4G13840.1	transferase family protein	1,6414	2,28E-02
AT5G16170.1	unknown protein	-3,3279	2,28E-02
AT5G10850.1	unknown protein	-2,0894	2,30E-02
AT1G78950.1	beta-amyrin synthase, putative	-1,5669	2,32E-02
AT5G19070.1	unknown protein	-4,9683	2,36E-02
AT3G48640.1	unknown protein	-2,7235	2,36E-02
AT1G43800.1	acyl-(acyl-carrier-protein) desaturase, putative / stearoyl-ACP desaturase, putative	-1,3476	2,37E-02
AT1G11300.1	carbohydrate binding / kinase	1,3810	2,38E-02
AT3G03070.1	NADH-ubiquinone oxidoreductase-related	-1,7720	2,40E-02
AT5G24510.1	60s acidic ribosomal protein P1, putative	-1,6475	2,41E-02
AT1G51820.1	leucine-rich repeat protein kinase, putative	-6,2491	2,41E-02
AT1G31580.1	ECS1	-3,6386	2,41E-02
AT1G02580.1	MEA (MEDEA); transcription factor	-2,1528	2,42E-02
AT2G37720.1	unknown protein	1,3174	2,42E-02
AT5G60430	unknown protein	1,6336	2,46E-02
AT2G21140.1	ATPRP2 (PROLINE-RICH PROTEIN 2)	1,7296	2,47E-02
AT5G62950	unknown protein	-1,3300	2,47E-02
AT1G02770.1	unknown protein	1,9997	2,48E-02
AT3G61010.1	glycosyl hydrolase family protein 85	-3,7875	2,48E-02
AT4G38770.1	PRP4 (PROLINE-RICH PROTEIN 4)	1,3548	2,49E-02
AT3G16520	UDP-glucoronosyl/UDP-glucosyl transferase family protein	-2,5094	2,49E-02
AT2G33855.1	unknown protein	-2,7633	2,51E-02
AT4G24920.1	protein transport protein SEC61 gamma subunit, putative	-1,5812	2,52E-02
AT1G62150.1	mitochondrial transcription termination factor-related / mTERF-related	-1,2192	2,53E-02
AT4G01920.1	DC1 domain-containing protein	-4,7335	2,54E-02
AT1G07550.1	leucine-rich repeat protein kinase, putative	-2,1909	2,55E-02
AT1G30530.1	UDP-glucoronosyl/UDP-glucosyl transferase family protein	-4,5045	2,55E-02
AT1G12860.1	basic helix-loop-helix (bHLH) family protein / F-box family protein	1,7173	2,56E-02
AT5G42700.1	transcriptional factor B3 family protein	1,7899	2,58E-02

AT1G75030.1	ATLP-3 (*Arabidopsis* thaumatin-like protein 3)	1,3122	2,58E-02
AT5G25910.1	disease resistance family protein	-1,3414	2,60E-02
AT5G17890.1	LIM domain-containing protein / disease resistance protein-related	-5,8554	2,60E-02
AT5G45030	unknown protein	-1,2857	2,61E-02
AT4G25710.1	kelch repeat-containing F-box family protein	-1,6748	2,62E-02
AT1G15870.1	mitochondrial glycoprotein family protein / MAM33 family protein	1,4505	2,64E-02
AT3G27550.1	group II intron splicing factor CRS1-related	-2,6689	2,64E-02
AT4G03500.1	ankyrin repeat family protein	-2,8200	2,64E-02
AT1G77380.1	AAP3 (amino acid permease 3); amino acid permease	1,3627	2,64E-02
AT1G64950.1	CYP89A5 (cytochrome P450, family 87, subfamily A, polypeptide 5); oxygen binding	-6,2952	2,66E-02
AT3G52870.1	calmodulin-binding family protein	-1,3061	2,66E-02
AT5G27730.1	unknown protein	1,4088	2,66E-02
AT5G43140.1	peroxisomal membrane 22 kDa family protein	-4,2359	2,67E-02
AT3G46130	MYB111 (myb domain protein 111); DNA binding / transcription factor	-1,4275	2,70E-02
AT1G76620.1	unknown protein	1,3570	2,71E-02
AT2G04500.1	DC1 domain-containing protein	-2,1093	2,71E-02
AT3G57090.1	binding	-2,4271	2,71E-02
AT5G40340.1	PWWP domain-containing protein	-3,5616	2,72E-02
AT2G37130.1	peroxidase 21 (PER21) (P21) (PRXR5)	-1,8395	2,72E-02
AT5G42850.1	electron carrier/ protein disulfide oxidoreductase	-5,5718	2,72E-02
AT5G28440.1	unknown protein	-1,2660	2,73E-02
AT4G07410	transducin family protein / WD-40 repeat family protein	-1,2753	2,73E-02
AT3G14210.1	ESM1 (EPITHIOSPECIFIER MODIFIER 1); carboxylic ester hydrolase	3,3521	2,75E-02
AT4G27050	F-box family protein	1,6607	2,77E-02
AT5G43520.1	DC1 domain-containing protein	-1,2159	2,77E-02
AT1G68570.1	proton-dependent oligopeptide transport (POT) family protein	1,3247	2,77E-02
AT5G45490.1	disease resistance protein-related	-5,7473	2,77E-02
AT2G35810.1	unknown protein	1,4844	2,78E-02
AT1G51480.1	disease resistance protein (CC-NBS-LRR class), putative	2,4242	2,79E-02
AT3G44020.1	thylakoid lumenal P17.1 protein	1,4607	2,80E-02
AT5G37450.1	leucine-rich repeat transmembrane protein kinase, putative	-1,7691	2,81E-02
AT1G24370.1	unknown protein	6,7029	2,81E-02
AT1G05650.1	polygalacturonase, putative / pectinase, putative	-1,8786	2,82E-02
AT2G31940.1	oxidoreductase/ transition metal ion binding	-1,3561	2,82E-02
AT4G11100	unknown protein	-1,2748	2,82E-02
AT5G24820.1	aspartyl protease family protein	-3,1831	2,82E-02
AT1G48280.1	hydroxyproline-rich glycoprotein family protein	-1,6799	2,82E-02
AT3G46410.1	protein kinase family protein	1,9507	2,83E-02
AT3G25250.1	AGC2-1 (OXIDATIVE SIGNAL-INDUCIBLE1); kinase	-2,4739	2,83E-02
AT1G64405.1	unknown protein	1,5815	2,84E-02
AT3G44400.1	disease resistance protein (TIR-NBS-LRR class), putative	-3,3707	2,84E-02
AT1G65820.1	microsomal glutathione s-transferase, putative	-1,5524	2,84E-02
AT5G50330.1	ATP binding / protein kinase	-1,3734	2,88E-02
AT5G23830.1	MD-2-related lipid recognition domain-containing protein / ML domain-containing protein	-5,4801	2,89E-02
AT1G80660.1	AHA9 (*Arabidopsis* H(+)-ATPase 9); hydrogen-exporting ATPase, phosphorylative mechanism	-1,8945	2,89E-02
AT2G38720.1	microtubule associated protein (MAP65/ASE1) family protein	1,8391	2,89E-02
AT5G44930.1	ARAD2 (ARABINAN DEFICIENT 2); catalytic	-1,2575	2,90E-02

AT4G16845	VRN2 (REDUCED VERNALIZATION RESPONSE 2); transcription factor	-2,9202	2,90E-02
AT4G22120	early-responsive to dehydration protein-related / ERD protein-related	-5,4863	2,94E-02
AT1G24145.1	unknown protein	-3,1620	2,94E-02
AT4G29240.1	leucine-rich repeat family protein / extensin family protein	1,3655	2,96E-02
AT4G21400.1	protein kinase family protein	1,1786	2,96E-02
AT2G26820.1	ATPP2-A3 (Phloem protein 2-A3); GTP binding	-6,5662	2,96E-02
AT5G24660.1	unknown protein	-4,3021	2,97E-02
AT3G21770.1	peroxidase 30 (PER30) (P30) (PRXR9)	-1,2543	2,97E-02
AT2G22970	SCPL11; serine carboxypeptidase	-1,3170	2,97E-02
AT2G01100.1	unknown protein	-2,1487	2,98E-02
AT3G51080.1	zinc finger (GATA type) family protein	1,4277	3,00E-02
AT2G16860.1	GCIP-interacting family protein	-1,5960	3,02E-02
AT1G75040.1	PR5 (PATHOGENESIS-RELATED GENE 5)	1,3567	3,02E-02
AT5G45670.1	GDSL-motif lipase/hydrolase family protein	1,3228	3,02E-02
AT5G12370	SEC10 (EXOCYST COMPLEX COMPONENT SEC10)	-1,3308	3,02E-02
AT5G13170.1	nodulin MtN3 family protein	1,8425	3,03E-02
AT5G65140.1	trehalose-6-phosphate phosphatase, putative	1,6888	3,03E-02
AT1G61870.1	pentatricopeptide (PPR) repeat-containing protein	-5,4428	3,03E-02
AT4G28490.1	HAESA (RECEPTOR-LIKE PROTEIN KINASE 5); ATP binding / kinase/ protein serine/threonine kinase	1,5887	3,03E-02
AT4G19810.1	glycosyl hydrolase family 18 protein	-1,6003	3,04E-02
AT1G51630.1	unknown protein	-6,7008	3,04E-02
AT4G16960.1	disease resistance protein (TIR-NBS-LRR class), putative	-4,2214	3,10E-02
AT2G26440.1	pectinesterase family protein	1,2122	3,11E-02
AT3G15570.1	phototropic-responsive NPH3 family protein	1,4348	3,14E-02
AT4G00440	unknown protein	-1,4081	3,15E-02
AT4G13460	SUVH9 (SU(VAR)3-9 HOMOLOG 9); histone-lysine N-methyltransferase/ zinc ion binding	-1,5996	3,15E-02
AT5G38260.1	serine/threonine protein kinase, putative	3,7214	3,16E-02
AT3G19620.1	glycosyl hydrolase family 3 protein	2,7965	3,16E-02
AT2G19150.1	pectinesterase family protein	-1,6243	3,17E-02
AT3G11010.1	disease resistance family protein / LRR family protein	-2,1393	3,17E-02
AT2G17700.1	protein kinase family protein	-2,3858	3,17E-02
AT3G44440.1	unknown protein	-3,3808	3,18E-02
AT2G03550.1	unknown protein	1,4436	3,20E-02
AT5G45940.1	ATNUDT11 (*Arabidopsis thaliana* Nudix hydrolase homolog 11); hydrolase	-1,5320	3,20E-02
AT2G04570.1	GDSL-motif lipase/hydrolase family protein	1,3165	3,20E-02
AT2G14255.1	zinc finger (DHHC type) family protein	-1,3341	3,21E-02
AT5G41140.1	unknown protein	-3,7003	3,21E-02
AT3G44770.1	unknown protein	1,4654	3,22E-02
AT3G11370.1	DC1 domain-containing protein	-3,7639	3,25E-02
AT4G14650.1	unknown protein	-2,9509	3,26E-02
AT5G38700.1	unknown protein	-1,4078	3,27E-02
AT5G03220.1	transcriptional co-activator-related	-1,3095	3,27E-02
AT1G77240.1	AMP-binding protein, putative	1,2129	3,28E-02
AT1G31814.1	FRL2 (FRIGIDA LIKE 2)	-4,0158	3,30E-02
AT3G61350.1	SKIP4 (SKP1 INTERACTING PARTNER 4)	-1,3538	3,30E-02
AT5G05640.1	nucleoprotein-related	-1,2924	3,30E-02
AT5G27200.1	ACP5 (ACYL CARRIER PROTEIN 5); acyl carrier	2,9358	3,30E-02
AT5G23460.1	unknown protein	-1,2987	3,31E-02

AT3G16030.1	CES101 (CALLUS EXPRESSION OF RBCS 101); carbohydrate binding / kinase	-2,1845	3,31E-02
AT1G22500.1	zinc finger (C3HC4-type RING finger) family protein	-1,1944	3,31E-02
AT1G49750.1	leucine-rich repeat family protein	1,3627	3,31E-02
AT2G43530.1	trypsin inhibitor, putative	1,4058	3,31E-02
AT4G20210.1	terpene synthase/cyclase family protein	-3,3248	3,33E-02
AT5G01330.1	PDC3 (PYRUVATE DECARBOXYLASE-3); pyruvate decarboxylase	-3,6934	3,34E-02
AT1G73165.1	CLE1 (CLAVATA3/ESR-RELATED 1); receptor binding	1,3817	3,35E-02
AT5G51730.1	nucleotide binding	-5,7241	3,36E-02
AT5G26180	NOL1/NOP2/sun family protein	-1,4102	3,43E-02
AT3G26730.1	zinc finger (C3HC4-type RING finger) family protein	-2,0573	3,43E-02
AT5G13880.1	unknown protein	1,4432	3,45E-02
AT1G17880.1	nascent polypeptide-associated complex (NAC) domain-containing protein / BTF3b-like transcription factor, putative	-1,2487	3,45E-02
AT1G08460.1	HDA08 (histone deacetylase 8); histone deacetylase	-1,6678	3,45E-02
AT4G16890.1	SNC1 (SUPPRESSOR OF NPR1-1, CONSTITUTIVE 1)	-2,0843	3,46E-02
AT1G70170.1	MMP (MATRIX METALLOPROTEINASE); metalloendopeptidase	-1,4368	3,46E-02
AT3G16750.1	unknown protein	-3,8470	3,47E-02
AT3G59130.1	unknown protein	1,4181	3,47E-02
AT5G24410.1	glucosamine/galactosamine-6-phosphate isomerase-related	-1,1656	3,48E-02
AT4G16750.1	DRE-binding transcription factor, putative	-1,2003	3,49E-02
AT5G04190.1	PKS4 (PHYTOCHROME KINASE SUBSTRATE 4)	1,5944	3,49E-02
AT1G78070	WD-40 repeat family protein	-2,0150	3,49E-02
AT5G56910.1	cysteine protease inhibitor	-1,8688	3,50E-02
AT5G59480	haloacid dehalogenase-like hydrolase family protein	-2,4542	3,50E-02
AT2G02690.1	protein binding / zinc ion binding	-2,2380	3,50E-02
AT4G03230.1	S-locus lectin protein kinase family protein	-1,5752	3,51E-02
AT5G04210.1	RNA recognition motif (RRM)-containing protein	-2,0468	3,53E-02
AT5G60260.1	unknown protein	-3,0017	3,54E-02
AT1G13790.1	XH/XS domain-containing protein / XS zinc finger domain-containing protein	1,2542	3,54E-02
AT5G64730.1	transducin family protein / WD-40 repeat family protein	-5,7716	3,56E-02
AT5G03960.1	IQD12 (IQ-domain 12); calmodulin binding	-1,6055	3,56E-02
AT3G46720.1	UDP-glucoronosyl/UDP-glucosyl transferase family protein	-1,1407	3,57E-02
AT3G06150.1	unknown protein	-1,7710	3,57E-02
AT2G35980.1	YLS9 (YELLOW-LEAF-SPECIFIC GENE 9)	-1,6877	3,59E-02
AT5G48410.1	ATGLR1.3 (*Arabidopsis thaliana* glutamate receptor 1.3)	-2,0257	3,60E-02
AT1G23330.1	unknown protein	1,2552	3,60E-02
AT1G48000.1	MYB112 (myb domain protein 112); DNA binding / transcription factor	-1,3966	3,60E-02
AT1G12750	rhomboid family protein	-1,1450	3,61E-02
AT4G22390.1	F-box family protein-related	1,2852	3,62E-02
AT5G49420.1	MADS-box protein (AGL84)	2,1750	3,63E-02
AT3G04720.1	PR4 (PATHOGENESIS-RELATED 4)	-1,3082	3,63E-02
AT1G15125.1	S-adenosyl-L-methionine:carboxyl methyltransferase family protein	-2,4016	3,64E-02
AT4G08540.1	unknown protein	-1,6525	3,64E-02
AT3G55130.1	ATWBC19 (WHITE-BROWN COMPLEX HOMOLOG 19); ATPase, coupled to transmembrane movement of substances	1,5769	3,64E-02
AT3G22250.1	UDP-glucoronosyl/UDP-glucosyl transferase family protein	-1,6607	3,66E-02
AT1G05220.1	unknown protein	-1,4655	3,67E-02
AT4G04990.1	unknown protein	1,5664	3,68E-02

AT2G37440	endonuclease/exonuclease/phosphatase family protein	-1,3864	3,69E-02
AT3G11240.1	arginine-tRNA-protein transferase, putative / arginyltransferase, putative / arginyl-tRNA-protein transferase, putative	-1,2767	3,70E-02
AT2G23910.1	cinnamoyl-CoA reductase-related	-3,4053	3,72E-02
AT2G31550.1	GDSL-motif lipase/hydrolase family protein	3,1875	3,72E-02
AT2G21130.1	peptidyl-prolyl cis-trans isomerase / cyclophilin (CYP2) / rotamase	1,3156	3,76E-02
AT1G34040.1	alliinase family protein	-2,6767	3,77E-02
AT4G34550.1	unknown protein	-2,3925	3,77E-02
AT5G66052.1	unknown protein	-5,5270	3,78E-02
AT5G62230.1	ERL1 (ERECTA-LIKE 1); kinase	1,5788	3,79E-02
AT2G05510	glycine-rich protein	-3,4157	3,79E-02
AT1G56290.1	CwfJ-like family protein	-2,1278	3,80E-02
AT1G78720.1	protein transport protein sec61, putative	-1,4648	3,82E-02
AT2G17650.1	AMP-dependent synthetase and ligase family protein	-1,1449	3,84E-02
AT3G46040.1	RPS15AD (ribosomal protein S15A D); structural constituent of ribosome	-7,0719	3,87E-02
AT5G18840.1	sugar transporter, putative	1,3983	3,88E-02
AT4G16800.1	enoyl-CoA hydratase, putative	1,4462	3,88E-02
AT1G31310.1	hydroxyproline-rich glycoprotein family protein	-1,7860	3,88E-02
AT5G02320.1	MYB3R-5 (myb domain protein 3R-5); DNA binding / transcription factor	1,1155	3,89E-02
AT5G08600.1	U3 ribonucleoprotein (Utp) family protein	-1,5822	3,89E-02
AT3G62410.1	CP12-2	-4,8284	3,89E-02
AT4G26170.1	unknown protein	1,4752	3,91E-02
AT1G75550.1	glycine-rich protein	1,5633	3,92E-02
AT2G13950.1	DC1 domain-containing protein	1,2302	3,92E-02
AT5G56380.1	F-box family protein	-5,8597	3,93E-02
AT5G48350.1	unknown protein	3,9126	3,93E-02
AT3G47965.1	unknown protein	-4,4277	3,95E-02
AT5G39860.1	PRE1 (PACLOBUTRAZOL RESISTANCE1); DNA binding / transcription factor	2,2718	3,96E-02
AT4G27820.1	glycosyl hydrolase family 1 protein	1,1784	3,97E-02
AT4G25720	glutamine cyclotransferase family protein	-2,7772	3,97E-02
AT2G15290.1	ATTIC21/CIA5/TIC21 (CHLOROPLAST IMPORT APPARATUS 5); protein homodimerization	-1,2557	3,99E-02
AT4G33910.1	oxidoreductase, 2OG-Fe(II) oxygenase family protein	-6,2053	4,00E-02
AT4G25810.1	XTR6 (XYLOGLUCAN ENDOTRANSGLYCOSYLASE 6); hydrolase, acting on glycosyl bonds	-2,5077	4,00E-02
AT4G11630.1	ribosomal protein L19 family protein	-1,4914	4,03E-02
AT1G57560.1	AtMYB50 (myb domain protein 50); DNA binding / transcription factor	-1,2126	4,04E-02
AT2G02010.1	glutamate decarboxylase, putative	1,4457	4,04E-02
AT1G60590.1	polygalacturonase, putative / pectinase, putative	1,7385	4,08E-02
AT5G38010.1	UDP-glucoronosyl/UDP-glucosyl transferase family protein	-2,0162	4,08E-02
AT5G45070.1	ATPP2-A8 (Phloem protein 2-A8); transmembrane receptor	-6,0255	4,12E-02
AT3G10370.1	glycerol-3-phosphate dehydrogenase, putative	-1,2599	4,12E-02
AT3G05790.1	Lon protease, putative	1,1996	4,17E-02
AT1G10610.1	DNA binding / transcription factor	-1,2729	4,18E-02
AT3G26730.1	zinc finger (C3HC4-type RING finger) family protein	-2,1026	4,18E-02
AT5G26000	TGG1 (THIOGLUCOSIDE GLUCOHYDROLASE 1); hydrolase, hydrolyzing O-glycosyl compounds	1,2064	4,19E-02
AT4G28940.1	catalytic	-1,3353	4,20E-02
AT2G23100.1	DC1 domain-containing protein	-1,2569	4,21E-02

AT1G52140.1	unknown protein	-1,2370	4,24E-02
AT1G05340.1	unknown protein	-1,8931	4,25E-02
AT3G44710.1	unknown protein	-2,0077	4,27E-02
AT2G30480	unknown protein	-1,6665	4,31E-02
AT5G44400.1	FAD-binding domain-containing protein	1,4866	4,31E-02
AT1G56500.1	haloacid dehalogenase-like hydrolase family protein	-1,5656	4,31E-02
AT1G49630	ATPREP2; metalloendopeptidase	-1,4929	4,33E-02
AT1G27520.1	glycoside hydrolase family 47 protein	-1,1820	4,34E-02
AT4G16765.2	oxidoreductase, 2OG-Fe(II) oxygenase family protein	-3,1839	4,34E-02
AT5G20450	motor	-3,1828	4,35E-02
AT1G14070.1	FUT7 (Fucosyltransferase 7); fucosyltransferase/ transferase, transferring glycosyl groups	1,1245	4,36E-02
AT4G19540.1	unknown protein	-1,3081	4,37E-02
AT5G47750.1	protein kinase, putative	-1,2771	4,38E-02
AT2G25680.1	sulfate transporter	-1,3577	4,38E-02
AT4G02330.1	pectinesterase family protein	2,3413	4,42E-02
AT2G14260	PIP (proline iminopeptidase); prolyl aminopeptidase	-1,2546	4,43E-02
AT5G64510.1	unknown protein	1,4541	4,43E-02
AT1G13410.1	binding	-1,4270	4,45E-02
AT1G15080.1	ATPAP2 (PHOSPHATIDIC ACID PHOSPHATASE 2); phosphatidate phosphatase	-1,4352	4,47E-02
AT1G64490.1	unknown protein	-1,3833	4,47E-02
AT4G11850.1	PLDGAMMA1 (maternal effect embryo arrest 54); phospholipase D	-1,3204	4,49E-02
AT3G46530.1	RPP13 (RECOGNITION OF PERONOSPORA PARASITICA 13); ATP binding	-6,2217	4,49E-02
AT1G10920.1	disease resistance protein (CC-NBS-LRR class), putative	1,5536	4,49E-02
AT5G62420.1	aldo/keto reductase family protein	1,2566	4,50E-02
AT4G29905.1	unknown protein	-3,1740	4,52E-02
AT3G11070.1	outer membrane OMP85 family protein	-4,8606	4,52E-02
AT4G21323.1	subtilase family protein	-1,4149	4,52E-02
AT3G20860.1	protein kinase family protein	-1,2557	4,53E-02
AT5G24170.1	unknown protein	-2,7804	4,55E-02
AT3G26730.1	zinc finger (C3HC4-type RING finger) family protein	-2,0660	4,56E-02
AT5G47760.1	ATPK5 (*Arabidopsis thaliana* serine/threonine protein kinase 5); phosphoglycolate phosphatase	-5,3670	4,59E-02
AT3G26740.1	CCL (CCR-LIKE)	1,3320	4,59E-02
AT4G29020.1	glycine-rich protein	1,6040	4,61E-02
AT2G06850.1	EXGT-A1 (ENDO-XYLOGLUCAN TRANSFERASE); hydrolase, acting on glycosyl bonds	1,3725	4,62E-02
AT5G25040.1	transporter	-2,3128	4,63E-02
AT1G18360.1	hydrolase, alpha/beta fold family protein	1,2093	4,68E-02
AT5G48340.1	unknown protein	-1,5308	4,68E-02
AT1G73270.1	SCPL6 (serine carboxypeptidase-like 6); serine carboxypeptidase	-1,4765	4,69E-02
AT1G51965.1	pentatricopeptide (PPR) repeat-containing protein	-1,1548	4,72E-02
AT5G23405	high mobility group (HMG1/2) family protein	-1,9487	4,72E-02
AT1G27170.1	disease resistance protein (TIR-NBS-LRR class), putative	-3,8052	4,73E-02
AT3G47770.1	ATATH5 (ABC2 homolog 5); ATPase, coupled to transmembrane movement of substances	-2,0380	4,73E-02
AT3G26730.1	zinc finger (C3HC4-type RING finger) family protein	-2,0137	4,74E-02
AT1G16150.1	WAKL4 (WALL ASSOCIATED KINASE-LIKE 4); kinase	1,9462	4,75E-02
AT4G01270.1	zinc finger (C3HC4-type RING finger) family protein	1,2716	4,79E-02
AT4G29730.1	NFC5 (NUCLEOSOME/CHROMATIN ASSEMBLY FACTOR	-1,1909	4,80E-02

	GROUP C 5)		
AT1G06360.1	fatty acid desaturase family protein	2,4978	4,81E-02
AT1G61190.1	disease resistance protein (CC-NBS-LRR class), putative	-1,1431	4,81E-02
AT5G37130.1	unknown protein	1,3147	4,83E-02
AT1G21520.1	unknown protein	2,3540	4,83E-02
AT2G01880.1	ATPAP7/PAP7 (purple acid phosphatase 7); acid phosphatase/ protein serine/threonine phosphatase	-1,5313	4,91E-02
AT4G11410.1	short-chain dehydrogenase/reductase (SDR) family protein	-1,2144	4,95E-02
AT5G59070.1	glycosyl transferase family 1 protein	1,5791	4,97E-02
AT3G56290.1	unknown protein	-1,1613	4,98E-02
AT4G32280.1	*IAA*29 (indoleacetic acid-induced protein 29); transcription factor	1,4078	4,98E-02
AT1G11100.1	SNF2 domain-containing protein / helicase domain-containing protein / zinc finger protein-related	-1,4074	4,99E-02

App. Table 3. Differentially expressed genes in seedlings- tetraploid Col-0 vs. tetraploid Ler-0 (four biological replicates)

Gene ID	Description	log_2 FC	P-value
AT5G45490.1	disease resistance protein-related	-5,5044	3,21E-06
AT5G45070.1	ATPP2-A8 (Phloem protein 2-A8); transmembrane receptor	-5,4779	8,80E-06
AT3G29410.1	terpene synthase/cyclase family protein	5,3146	1,56E-05
AT4G02860.1	catalytic	-4,2385	1,82E-05
AT1G74450.1	unknown protein	-4,5429	2,07E-05
AT1G29790.1	unknown protein	-6,5407	2,13E-05
AT4G16950	RPP5 (RECOGNITION OF PERONOSPORA PARASITICA 5)	-4,7489	2,41E-05
AT2G26820.1	ATPP2-A3 (Phloem protein 2-A3); GTP binding	-5,1593	2,42E-05
AT1G63880.1	disease resistance protein (TIR-NBS-LRR class), putative	-5,4622	3,44E-05
AT1G52100.1	jacalin lectin family protein	-6,3618	3,50E-05
AT1G65280.1	heat shock protein binding / unfolded protein binding	-6,0678	3,91E-05
AT5G66052.1	unknown protein	-4,3028	3,98E-05
AT5G42825.1	unknown protein	-4,5180	4,52E-05
AT4G15010	mitochondrial substrate carrier family protein	-4,6827	4,65E-05
AT1G58410.1	disease resistance protein (CC-NBS-LRR class), putative	-4,9609	5,82E-05
AT1G34200.1	oxidoreductase family protein	-4,5596	6,13E-05
AT4G02540.1	DC1 domain-containing protein	-5,8634	6,15E-05
AT5G24240.1	phosphatidylinositol 3- and 4-kinase family protein / ubiquitin family protein	5,0003	6,66E-05
AT3G16740.1	F-box family protein	-4,7286	6,94E-05
AT1G50520.1	CYP705A27 (cytochrome P450, family 705, subfamily A, polypeptide 27); oxygen binding	-4,5212	7,22E-05
AT3G26330.1	CYP71B37 (cytochrome P450, family 71, subfamily B, polypeptide 37); oxygen binding	-5,5276	7,57E-05
AT5G17090.1	unknown protein	3,8969	8,82E-05
AT5G03200.1	zinc finger (C3HC4-type RING finger) family protein	-5,1863	8,89E-05
AT5G19070.1	unknown protein	-3,7780	9,10E-05
AT5G54040.1	DC1 domain-containing protein	-6,2357	9,29E-05
AT3G60480.1	unknown protein	-4,4497	9,46E-05
AT1G34200.1	oxidoreductase family protein	-3,2508	9,49E-05
AT3G15840	unknown protein	-5,3928	1,02E-04
AT1G29800	zinc ion binding	-4,7386	1,04E-04
AT5G27100.1	ATGLR2.1 (*Arabidopsis thaliana* glutamate receptor 2.1)	-4,4533	1,22E-04
AT5G52070.1	agenet domain-containing protein	-4,2130	1,30E-04
AT4G16950.1	RPP5 (RECOGNITION OF PERONOSPORA PARASITICA 5)	-5,8824	1,30E-04
AT5G24760	alcohol dehydrogenase, putative	-6,0167	1,30E-04

AT5G48620.1	disease resistance protein (CC-NBS-LRR class), putative	-3,6278	1,36E-04
AT5G47760.1	ATPK5 (*Arabidopsis thaliana* serine/threonine protein kinase 5); phosphoglycolate phosphatase	-4,4848	1,40E-04
AT5G17890.1	LIM domain-containing protein / disease resistance protein-related	-5,5359	1,51E-04
AT1G30835	unknown protein	-4,0698	1,54E-04
AT1G80760.1	NIP6;1 (NOD26-like intrinsic protein 6;1); water channel	-3,1003	1,59E-04
AT5G23830.1	MD-2-related lipid recognition domain-containing protein / ML domain-containing protein	-5,0357	1,63E-04
AT5G55940.1	EMB2731 (EMBRYO DEFECTIVE 2731)	-6,0339	1,65E-04
AT2G25670	unknown protein	-3,3184	1,66E-04
AT5G45430.1	protein kinase, putative	-5,5304	1,68E-04
AT5G26270.1	unknown protein	-4,1042	1,93E-04
AT4G24420.1	RNA recognition motif (RRM)-containing protein	4,7418	1,93E-04
AT2G06005	unknown protein	-3,8681	2,08E-04
AT2G18193.1	AAA-type ATPase family protein	-5,7515	2,27E-04
AT1G58150.1	unknown protein	-3,2262	2,29E-04
AT1G74280.1	hydrolase, alpha/beta fold family protein	-4,6054	2,42E-04
AT4G37530	peroxidase, putative	-3,2294	2,56E-04
AT1G51630.1	unknown protein	-6,6133	2,58E-04
AT1G24370.1	unknown protein	5,7921	2,97E-04
AT5G49830.1	unknown protein	-5,5400	3,35E-04
AT4G01910.1	DC1 domain-containing protein	-2,7220	3,56E-04
AT1G51430.1	unknown protein	-3,6194	3,62E-04
AT1G44935.1	unknown protein	4,5884	3,62E-04
AT1G54040	ESP (EPITHIOSPECIFIER PROTEIN)	4,6335	3,66E-04
AT2G33340	transducin family protein / WD-40 repeat family protein	-2,9575	3,89E-04
AT4G16860.1	RPP4 (RECOGNITION OF PERONOSPORA PARASITICA 4)	-5,0814	3,91E-04
AT5G54020.1	unknown protein	-3,9654	3,99E-04
AT2G37750.1	unknown protein	-4,2290	4,05E-04
AT3G32930.1	unknown protein	-2,8985	4,07E-04
AT5G55790.1	unknown protein	-2,8036	4,18E-04
AT3G01720.1	unknown protein	-4,2796	4,47E-04
AT1G27385	unknown protein	-3,0607	5,00E-04
AT4G29760.1	unknown protein	3,1059	5,00E-04
AT2G02360.1	ATPP2-B10 (Phloem protein 2-B10)	-5,1529	5,33E-04
AT1G59900.1	AT-E1 ALPHA (pyruvate dehydrogenase complex E1 alpha subunit); pyruvate dehydrogenase (acetyl-transferring)	-3,8208	5,38E-04
AT2G07777	unknown protein	-5,3063	5,66E-04
AT4G22120	early-responsive to dehydration protein-related / ERD protein-related	-4,3162	6,56E-04
AT3G63330.1	protein kinase family protein	-2,8089	6,60E-04
AT4G20480.1	unknown protein	-2,5633	6,66E-04
AT5G45760	transducin family protein / WD-40 repeat family protein	-2,4347	6,74E-04
AT5G51730.1	nucleotide binding	-4,4614	7,19E-04
AT4G19500.1	disease resistance protein (TIR-NBS-LRR class), putative	-3,8181	7,26E-04
AT1G70080.1	terpene synthase/cyclase family protein	3,1588	7,32E-04
AT4G04402.1	two-component phosphorelay mediator, putative	2,3548	8,08E-04
AT1G44920.1	unknown protein	-3,7675	8,13E-04
AT1G12650	unknown protein	-2,3678	8,25E-04
AT5G36930.1	disease resistance protein (TIR-NBS-LRR class), putative	-3,4132	8,29E-04
AT2G25260.1	unknown protein	-3,2900	8,40E-04
AT1G51820.1	leucine-rich repeat protein kinase, putative	-3,3139	8,47E-04

AT4G03050	AOP3 (2-oxoglutarate?dependent dioxygenase 3); oxidoreductase, acting on paired donors, with incorporation or reduction of molecular oxygen, 2-oxoglutarate as one donor, and incorporation of one atom each of oxygen into both donors	3,4557	8,57E-04
AT5G40500	unknown protein	-3,3857	8,59E-04
AT5G45500.1	unknown protein	-4,6166	8,61E-04
AT5G40910.1	disease resistance protein (TIR-NBS-LRR class), putative	-2,6510	8,72E-04
AT3G61010.1	glycosyl hydrolase family protein 85	-2,6466	8,78E-04
AT3G52550.1	unknown protein	2,4944	8,79E-04
AT2G34840.1	coatomer protein epsilon subunit family protein / COPE family protein	-3,8205	9,08E-04
AT5G41740.1	disease resistance protein (TIR-NBS-LRR class), putative	-2,6346	9,38E-04
AT4G12170.1	thioredoxin family protein	5,6870	9,41E-04
AT3G46040.1	RPS15AD (ribosomal protein S15A D); structural constituent of ribosome	-6,5893	9,52E-04
AT2G37120.1	DNA-binding S1FA family protein	-5,2095	9,63E-04
AT2G04380.1	unknown protein	-2,8071	9,81E-04
AT2G17580.1	polynucleotide adenylyltransferase family protein	-4,8894	1,03E-03
AT4G34930.1	1-phosphatidylinositol phosphodiesterase-related	-3,0350	1,05E-03
AT2G15790.1	SQN (SQUINT)	2,3049	1,05E-03
AT3G14650.1	CYP72A11 (cytochrome P450, family 72, subfamily A, polypeptide 11); oxygen binding	-3,0650	1,06E-03
AT5G42850.1	electron carrier/ protein disulfide oxidoreductase	-6,1591	1,07E-03
AT3G27470	unknown protein	-2,8487	1,09E-03
AT4G20210.1	terpene synthase/cyclase family protein	-2,3978	1,13E-03
AT1G62190.1	unknown protein	3,1428	1,14E-03
AT5G15360.1	unknown protein	3,2534	1,14E-03
AT2G05830	eukaryotic translation initiation factor 2B family protein / eIF-2B family protein	-5,1075	1,16E-03
AT4G12280.1	copper amine oxidase family protein	3,7874	1,18E-03
AT5G36220.1	CYP81D1 (CYTOCHROME P450 91A1); oxygen binding	-2,5745	1,19E-03
AT4G01915	unknown protein	-2,7236	1,26E-03
AT5G05890.1	UDP-glucoronosyl/UDP-glucosyl transferase family protein	-4,9652	1,30E-03
AT2G42270.1	U5 small nuclear ribonucleoprotein helicase, putative	-3,5265	1,31E-03
AT2G18210.1	unknown protein	-3,5482	1,31E-03
AT1G31820.1	amino acid permease family protein	-3,0088	1,35E-03
AT3G47965.1	unknown protein	-4,3249	1,37E-03
AT4G09300.1	unknown protein	-2,8290	1,45E-03
AT2G14285.1	unknown protein	-3,0706	1,48E-03
AT2G44200.1	unknown protein	-2,4034	1,50E-03
AT4G16960.1	disease resistance protein (TIR-NBS-LRR class), putative	-3,0637	1,53E-03
AT4G14650.1	unknown protein	-3,4159	1,55E-03
AT4G14385	unknown protein	-4,0566	1,56E-03
AT4G01920.1	DC1 domain-containing protein	-3,2583	1,58E-03
AT4G09420.1	disease resistance protein (TIR-NBS class), putative	-3,0433	1,61E-03
AT3G32040.1	geranylgeranyl pyrophosphate synthase, putative / GGPP synthetase, putative / farnesyltranstransferase, putative	-2,7513	1,62E-03
AT5G24680.1	unknown protein	-2,7513	1,64E-03
AT1G53980.1	polyubiquitin-related	-2,1834	1,69E-03
AT2G09795.2	unknown protein	-3,2581	1,70E-03
AT5G43740	disease resistance protein (CC-NBS-LRR class), putative	-2,7183	1,72E-03
AT1G35320.1	unknown protein	-4,1160	1,75E-03
AT2G02340.1	ATPP2-B8 (Phloem protein 2-B8)	-2,4665	1,76E-03

AT5G56380.1	F-box family protein	-4,9046	1,77E-03
AT3G18380	DNA binding / sequence-specific DNA binding / transcription factor	-2,4119	1,79E-03
AT1G52760.1	esterase/lipase/thioesterase family protein	-2,4797	1,86E-03
AT3G19515.2	binding	-3,3128	1,88E-03
AT5G43320.1	CKL8 (Casein Kinase I-like 8); casein kinase I/ kinase	-2,2448	1,92E-03
AT3G16750.1	unknown protein	-2,3463	1,94E-03
AT5G45220.1	Toll-Interleukin-Resistance (TIR) domain-containing protein	-3,2478	1,95E-03
AT3G44430.1	unknown protein	-3,1111	2,08E-03
AT1G58280	unknown protein	-2,9126	2,08E-03
AT1G65190.1	protein kinase family protein	-3,3526	2,16E-03
AT3G11070.1	outer membrane OMP85 family protein	-2,1384	2,20E-03
AT1G47570	zinc finger (C3HC4-type RING finger) family protein	-2,5179	2,26E-03
AT3G28130	nodulin MtN21 family protein	-3,0684	2,26E-03
AT4G21326.1	subtilase family protein	2,9971	2,30E-03
AT1G61310.1	disease resistance protein (CC-NBS-LRR class), putative	-2,8202	2,36E-03
AT1G59780.1	disease resistance protein (CC-NBS-LRR class), putative	-2,1570	2,56E-03
AT4G31570.1	unknown protein	2,2779	2,60E-03
AT5G24140.1	SQP2 (Squalene monooxygenase 2); oxidoreductase	-4,3978	2,90E-03
AT5G17880.1	CSA1 (CONSTITUTIVE SHADE-AVOIDANCE1); ATP binding / protein binding / transmembrane receptor	-2,9743	2,92E-03
AT2G44770.1	phagocytosis and cell motility protein ELMO1-related	-2,0329	2,93E-03
AT3G26830.1	PAD3 (PHYTOALEXIN DEFICIENT 3); oxygen binding	-2,8847	2,93E-03
AT3G46980	transporter-related	-2,3770	2,94E-03
AT4G00650.1	FRI (FRIGIDA)	-2,5394	2,96E-03
AT1G27540.1	F-box family protein	2,3691	2,97E-03
AT5G65850.1	F-box family protein	-2,6985	3,05E-03
AT1G19720.1	pentatricopeptide (PPR) repeat-containing protein	-2,4289	3,06E-03
AT3G03070.1	NADH-ubiquinone oxidoreductase-related	-1,9833	3,09E-03
AT4G09680.1	unknown protein	-2,5358	3,10E-03
AT5G42850.2	electron carrier/ protein disulfide oxidoreductase	-2,0610	3,14E-03
AT5G02630.1	unknown protein	-1,9094	3,14E-03
AT2G09795	unknown protein	-2,9077	3,30E-03
AT5G64730.1	transducin family protein / WD-40 repeat family protein	-2,3434	3,31E-03
AT1G64950.1	CYP89A5 (cytochrome P450, family 87, subfamily A, polypeptide 5); oxygen binding	-6,4726	3,32E-03
AT3G50480.1	HR4 (HOMOLOG OF RPW8 4)	-8,5781	3,33E-03
AT1G71300.1	Vps52/Sac2 family protein	2,5073	3,36E-03
AT3G23070.1	unknown protein	-2,1363	3,38E-03
AT1G55380.1	DC1 domain-containing protein	-1,8597	3,54E-03
AT2G25450.1	2-oxoglutarate-dependent dioxygenase, putative	-3,9450	3,59E-03
AT1G27540.1	F-box family protein	-5,0365	3,75E-03
AT4G12330.1	CYP706A7 (cytochrome P450, family 706, subfamily A, polypeptide 7); oxygen binding	-2,5323	3,97E-03
AT5G43540.1	zinc finger (C2H2 type) family protein	-1,8280	4,02E-03
AT5G60230	SEN2 (SPLICING ENDONUCLEASE 2); tRNA-intron endonuclease	-2,5427	4,23E-03
AT4G07825.1	unknown protein	-2,3482	4,31E-03
AT5G37970.1	S-adenosyl-L-methionine:carboxyl methyltransferase family protein	-3,4363	4,32E-03
AT5G27110.1	pentatricopeptide (PPR) repeat-containing protein	-2,4647	4,34E-03
AT5G42965.1	nucleic acid binding / ribonuclease H	-2,5394	4,35E-03
AT1G34630.1	unknown protein	-1,7541	4,55E-03

AT5G64630	FAS2 (FASCIATA 2); nucleotide binding	1,9211	4,55E-03
AT1G69720.1	HO3 (HEME OXYGENASE 3); heme oxygenase (decyclizing)	2,9948	4,57E-03
AT1G65295.1	unknown protein	-1,9727	4,64E-03
AT5G45080.1	ATPP2-A6 (Phloem protein 2-A6); transmembrane receptor	-3,6350	4,65E-03
AT5G19100.1	extracellular dermal glycoprotein-related / EDGP-related	2,8834	4,66E-03
AT5G43150.1	unknown protein	-1,8833	4,68E-03
AT2G16530	3-oxo-5-alpha-steroid 4-dehydrogenase family protein / steroid 5-alpha-reductase family protein	-2,7467	4,80E-03
AT2G01090.1	ubiquinol-cytochrome C reductase complex 7.8 kDa protein, putative / mitochondrial hinge protein, putative	-3,2985	4,80E-03
AT4G33910.1	oxidoreductase, 2OG-Fe(II) oxygenase family protein	-2,0063	4,85E-03
AT2G44260	unknown protein	-1,9666	4,91E-03
AT3G03950	ECT1; protein binding	-2,0116	5,00E-03
AT4G19420	pectinacetylesterase family protein	2,1767	5,43E-03
AT5G56900	CwfJ-like family protein / zinc finger (CCCH-type) family protein	-1,8254	5,58E-03
AT5G18350.1	disease resistance protein (TIR-NBS-LRR class), putative	-3,6658	5,61E-03
AT5G17190.1	unknown protein	-2,4062	5,78E-03
AT1G12010.1	1-aminocyclopropane-1-carboxylate oxidase, putative / ACC oxidase, putative	-3,4935	5,81E-03
AT5G10850.1	unknown protein	-2,0590	5,83E-03
AT4G39955.1	hydrolase, alpha/beta fold family protein	-1,7857	5,89E-03
AT5G38590	F-box family protein	1,8899	6,07E-03
AT3G49360.1	glucosamine/galactosamine-6-phosphate isomerase family protein	-2,1097	6,09E-03
AT4G19080.1	unknown protein	1,7023	6,49E-03
AT2G42170.1	actin, putative	-2,1330	6,53E-03
AT2G17430.1	ATMLO7/MLO7 (MILDEW RESISTANCE LOCUS O 7); calmodulin binding	-2,1742	6,61E-03
AT3G44190.1	pyridine nucleotide-disulphide oxidoreductase family protein	2,1102	6,73E-03
AT2G30230.1	unknown protein	3,1437	6,90E-03
AT3G44070.1	unknown protein	3,4280	6,96E-03
AT5G28920.1	unknown protein	-4,0226	7,10E-03
AT5G44575.1	unknown protein	-3,2669	7,47E-03
AT5G24850.1	CRY3 (CRYPTOCHROME 3); DNA binding / DNA photolyase/ FMN binding	-2,1380	7,65E-03
AT2G31550.1	GDSL-motif lipase/hydrolase family protein	1,9249	7,73E-03
AT2G28990.1	leucine-rich repeat protein kinase, putative	-3,3934	7,86E-03
AT5G47980.1	transferase family protein	2,0343	7,92E-03
AT4G39900.1	unknown protein	-1,7677	7,96E-03
AT4G19770.1	glycosyl hydrolase family 18 protein	2,6229	8,14E-03
AT5G42590.1	CYP71A16 (cytochrome P450, family 71, subfamily A, polypeptide 16); oxygen binding	-3,8049	8,21E-03
AT5G23010.1	MAM1 (2-isopropylmalate synthase 3); 2-isopropylmalate synthase	-3,7730	8,62E-03
AT2G05510	glycine-rich protein	-3,4130	8,69E-03
AT5G43470	RPP8 (RECOGNITION OF PERONOSPORA PARASITICA 8)	-2,0452	8,70E-03
AT5G43140.1	peroxisomal membrane 22 kDa family protein	-2,0538	8,96E-03
AT1G12220.1	RPS5 (RESISTANT TO P. SYRINGAE 5)	-2,6018	9,21E-03
AT5G22140	pyridine nucleotide-disulphide oxidoreductase family protein	2,0800	9,70E-03
AT2G02960	zinc finger (C3HC4-type RING finger) family protein	-2,2644	9,95E-03
AT5G28030	cysteine synthase, putative / O-acetylserine (thiol)-lyase, putative / O-acetylserine sulfhydrylase, putative	1,8856	1,01E-02
AT2G06025.1	GCN5-related N-acetyltransferase (GNAT) family protein	-2,2877	1,02E-02

AT1G80960	F-box protein-related	-2,0243	1,10E-02
AT5G65080	MAF5 (MADS AFFECTING FLOWERING 5); transcription factor	-2,2682	1,10E-02
AT1G09995.1	unknown protein	-2,4366	1,18E-02
AT1G10095.1	protein prenyltransferase	1,9928	1,19E-02
AT1G62085.1	mitochondrial transcription termination factor family protein / mTERF family protein	-2,1585	1,21E-02
AT1G58270.1	ZW9	-5,0120	1,22E-02
AT4G00970.1	protein kinase family protein	2,0867	1,23E-02
AT5G44360.1	FAD-binding domain-containing protein	1,7594	1,23E-02
AT1G31710.1	copper amine oxidase, putative	-2,6490	1,28E-02
AT4G15400.1	transferase family protein	-2,4428	1,33E-02
AT3G28200.1	peroxidase, putative	-2,3454	1,35E-02
AT5G13380.1	auxin-responsive GH3 family protein	1,6691	1,36E-02
AT1G13460	serine/threonine protein phosphatase 2A (PP2A) regulatory subunit B', putative	-1,9329	1,39E-02
AT1G10110.1	F-box family protein	1,7078	1,42E-02
AT1G60730	aldo/keto reductase family protein	-1,7626	1,42E-02
AT3G45070.1	sulfotransferase family protein	-4,2183	1,47E-02
AT3G53370.1	DNA-binding S1FA family protein	-1,7833	1,49E-02
AT2G31540.1	GDSL-motif lipase/hydrolase family protein	-2,1643	1,51E-02
AT5G05750.1	DNAJ heat shock N-terminal domain-containing protein	-1,8112	1,52E-02
AT3G09160.1	RNA recognition motif (RRM)-containing protein	-1,6438	1,55E-02
AT4G04810.1	methionine sulfoxide reductase domain-containing protein / SeIR domain-containing protein	-2,5055	1,61E-02
AT2G34430.1	LHB1B1 (Photosystem II light harvesting complex gene 1.4); chlorophyll binding	2,1358	1,63E-02
AT2G02240.1	MEE66 (maternal effect embryo arrest 66)	1,4302	1,63E-02
AT3G46370.1	leucine-rich repeat protein kinase, putative	-1,5664	1,63E-02
AT1G21350	electron carrier	-3,3012	1,66E-02
AT2G17560	HMGB4 (HIGH MOBILITY GROUP B 4); transcription factor	-2,1380	1,69E-02
AT4G39180.1	SEC14 (secretion 14)	-2,3395	1,70E-02
AT4G01590	unknown protein	-1,7129	1,70E-02
AT2G17250.1	unknown protein	-1,5552	1,75E-02
AT1G71400.1	disease resistance family protein / LRR family protein	-1,9122	1,79E-02
AT3G55020.1	RabGAP/TBC domain-containing protein	-1,4140	1,80E-02
AT2G25460.1	unknown protein	-3,4877	1,86E-02
AT3G11370.1	DC1 domain-containing protein	-2,3858	1,87E-02
AT4G01380.1	plastocyanin-like domain-containing protein	1,5354	1,88E-02
AT1G78265.2	unknown protein	-1,4463	1,93E-02
AT4G14020.1	rapid alkalinization factor (RALF) family protein	1,8099	1,93E-02
AT4G25810.1	XTR6 (XYLOGLUCAN ENDOTRANSGLYCOSYLASE 6); hydrolase, acting on glycosyl bonds	-1,6909	1,93E-02
AT5G15980.1	pentatricopeptide (PPR) repeat-containing protein	-1,8870	1,93E-02
AT3G57090.1	binding	-1,7608	1,97E-02
AT1G52150	ATHB-15 (INCURVATA 4); DNA binding / transcription factor	-1,5468	1,98E-02
AT2G38620	CDKB1;2 (cyclin-dependent kinase B1;2); kinase	-2,8755	2,01E-02
AT1G30835	unknown protein	-1,9387	2,03E-02
AT1G58602.1	disease resistance protein (CC-NBS-LRR class), putative	1,5144	2,08E-02
AT1G20696	HMGB3 (HIGH MOBILITY GROUP B 3); transcription factor	-1,5139	2,08E-02
AT1G69550.1	disease resistance protein (TIR-NBS class), putative	-2,2729	2,08E-02
AT1G57790.1	F-box family protein	-1,7365	2,12E-02
AT2G26030	F-box family protein	-1,7405	2,14E-02

AT4G23496.1	SP1L5 (SPIRAL1-LIKE5)	1,6356	2,14E-02
AT5G43403.1	unknown protein	-1,8484	2,17E-02
AT2G18670.1	zinc finger (C3HC4-type RING finger) family protein	-1,8833	2,17E-02
AT4G27050	F-box family protein	1,6434	2,18E-02
AT5G48350.1	unknown protein	2,3381	2,19E-02
AT1G51480.1	disease resistance protein (CC-NBS-LRR class), putative	2,0058	2,20E-02
AT1G65370.1	meprin and TRAF homology domain-containing protein / MATH domain-containing protein	-2,6278	2,25E-02
AT2G13960.1	myb family transcription factor	-2,0739	2,25E-02
AT5G39080.1	transferase family protein	-3,1394	2,27E-02
AT4G26320.1	AGP13 (ARABINOGALACTAN PROTEIN 13)	-2,2846	2,28E-02
AT5G02350.1	DC1 domain-containing protein	-1,6163	2,37E-02
AT4G19240.1	unknown protein	2,4371	2,39E-02
AT5G02600.1	heavy-metal-associated domain-containing protein	-1,9063	2,42E-02
AT1G61870.1	pentatricopeptide (PPR) repeat-containing protein	-1,9900	2,43E-02
AT4G20860.1	FAD-binding domain-containing protein	-1,6465	2,45E-02
AT5G42700.1	transcriptional factor B3 family protein	1,5399	2,50E-02
AT5G37350	RIO1 family protein	-1,3890	2,51E-02
AT4G19840.1	ATPP2-A1 (*Arabidopsis thaliana* phloem protein 2-A1)	-1,6718	2,55E-02
AT5G43440	2-oxoglutarate-dependent dioxygenase, putative	-2,1950	2,61E-02
AT5G37990.1	S-adenosylmethionine-dependent methyltransferase	-3,4109	2,61E-02
AT2G20790	unknown protein	-1,6509	2,63E-02
AT3G58940.1	F-box family protein	1,5199	2,68E-02
AT5G01330.1	PDC3 (PYRUVATE DECARBOXYLASE-3); pyruvate decarboxylase	-2,1635	2,71E-02
AT5G25010.1	unknown protein	-2,2563	2,77E-02
AT5G27200.1	ACP5 (ACYL CARRIER PROTEIN 5); acyl carrier	1,9007	2,84E-02
AT5G26300.1	meprin and TRAF homology domain-containing protein / MATH domain-containing protein	1,7398	2,85E-02
AT5G10140.1	FLC (FLOWERING LOCUS C); transcription factor	-1,5370	2,91E-02
AT3G52780	ATPAP20/PAP20; acid phosphatase/ protein serine/threonine phosphatase	-1,7081	2,94E-02
AT1G17240.1	leucine-rich repeat family protein	-2,0162	2,97E-02
AT4G03500.1	ankyrin repeat family protein	-1,7688	3,05E-02
AT5G63180.1	pectate lyase family protein	1,6770	3,12E-02
AT5G46780	VQ motif-containing protein	-1,9200	3,23E-02
AT4G00970.1	protein kinase family protein	1,7132	3,24E-02
AT3G07180	GPI transamidase component PIG-S-related	-1,6178	3,29E-02
AT4G15280.1	UDP-glucoronosyl/UDP-glucosyl transferase family protein	1,4070	3,34E-02
AT3G21770.1	peroxidase 30 (PER30) (P30) (PRXR9)	-1,4489	3,36E-02
AT4G37620.1	nucleic acid binding / ribonuclease H	-2,0134	3,38E-02
AT4G18975	pentatricopeptide (PPR) repeat-containing protein	1,4130	3,42E-02
AT5G24280.1	ATP binding	-1,7795	3,45E-02
AT2G04680.1	DC1 domain-containing protein	-1,7821	3,50E-02
AT3G51030.1	ATTRX1 (*Arabidopsis thaliana* thioredoxin H-type 1); thiol-disulfide exchange intermediate	-1,2798	3,55E-02
AT1G59640	ZCW32 (BIGPETAL, BIGPETALUB); DNA binding / transcription factor	-4,1272	3,56E-02
AT2G43520.1	ATTI2 (*ARABIDOPSIS THALIANA* TRYPSIN INHIBITOR PROTEIN 2); trypsin inhibitor	1,5379	3,56E-02
AT5G59480	haloacid dehalogenase-like hydrolase family protein	-1,7267	3,56E-02
AT4G11830	phospholipase D gamma 2 / PLD gamma 2 (PLDGAMMA2)	-1,3863	3,71E-02
AT5G20750.1	Ulp1 protease family protein	2,6235	3,73E-02
AT5G23395.1	unknown protein	-1,2441	3,73E-02

AT3G29670.1	transferase family protein	-1,8601	3,82E-02
AT4G11210.1	disease resistance-responsive family protein / dirigent family protein	-3,5451	3,84E-02
AT1G30500	CCAAT-binding transcription factor (CBF-B/NF-YA) family protein	-1,6606	3,86E-02
AT5G48335.1	unknown protein	-1,2405	3,87E-02
AT2G31490.1	unknown protein	1,4586	3,87E-02
AT4G16845	VRN2 (REDUCED VERNALIZATION RESPONSE 2); transcription factor	-3,3902	3,96E-02
AT3G26180	CYP71B20 (cytochrome P450, family 71, subfamily B, polypeptide 20); oxygen binding	-1,7984	3,97E-02
AT4G22235	unknown protein	2,0013	4,02E-02
AT1G48690.1	auxin-responsive GH3 family protein	-2,2182	4,11E-02
AT4G19515.1	disease resistance family protein	-1,6977	4,17E-02
AT4G10950.1	GDSL-motif lipase/hydrolase family protein	1,4120	4,17E-02
AT1G12400.1	unknown protein	-1,9005	4,19E-02
AT1G66980.1	protein kinase family protein / glycerophosphoryl diester phosphodiesterase family protein	1,8827	4,20E-02
AT2G35810.1	unknown protein	1,3134	4,35E-02
AT4G27890.1	nuclear movement family protein	1,7663	4,36E-02
AT5G47810.1	phosphofructokinase family protein	-2,3388	4,37E-02
AT3G50440.1	hydrolase	-1,2252	4,38E-02
AT5G43660.1	oxidoreductase, acting on paired donors, with incorporation or reduction of molecular oxygen, 2-oxoglutarate as one donor, and incorporation of one atom each of oxygen into both donors	2,2659	4,42E-02
AT4G02950.1	ubiquitin family protein	1,4742	4,52E-02
AT1G35617.1	unknown protein	1,2972	4,55E-02
AT3G47200	unknown protein	-1,2591	4,57E-02
AT4G22460.1	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein	-1,7099	4,60E-02
AT3G07800.1	thymidine kinase, putative	-1,8550	4,64E-02
AT5G42460.1	F-box family protein	-1,8385	4,67E-02
AT2G23130.1	AGP17 (ARABINOGALACTAN PROTEIN 17)	-2,9692	4,67E-02
AT1G21350.2	electron carrier	-1,5534	4,68E-02
AT3G24515.1	UBC37 (ubiquitin-conjugating enzyme 34); ubiquitin-protein ligase	-2,5174	4,68E-02
AT1G74670.1	gibberellin-responsive protein, putative	2,8030	4,68E-02
AT5G43030.1	DC1 domain-containing protein	-3,2431	4,69E-02
AT4G12030	bile acid:sodium symporter family protein	-1,6147	4,71E-02
AT1G66130.1	oxidoreductase N-terminal domain-containing protein	-2,1604	4,74E-02
AT4G19760.1	glycosyl hydrolase family 18 protein	1,3395	4,75E-02
AT3G24440.1	VRN5 (VIN3-LIKE 1)	-1,4796	4,80E-02
AT1G63130.1	pentatricopeptide (PPR) repeat-containing protein	-2,3385	4,84E-02
AT5G63820.1	unknown protein	1,9385	4,92E-02
AT4G12270.1	copper amine oxidase family protein	2,3717	4,95E-02
AT2G45660.1	AGL20 (AGAMOUS-LIKE 20); transcription factor	1,7297	4,99E-02

App. Table 4. Differentially expressed genes in seedlings - diploid Col-0 (Cy3-label) vs. tetraploid Col-0 (Cy5-label; four biological replicates)

Gene ID	Description	log_2 FC	P-value
AT1G53480.1	unknown protein	4,08E+04	2,99E-04
AT2G01520.1	major latex protein-related / MLP-related	1,46E+04	3,17E-03
AT5G01380.1	transcription factor	-1,52E+04	3,60E-03
AT4G13420.1	HAK5 (High affinity K+ transporter 5); potassium ion transporter	1,75E+04	6,40E-03
AT2G34430.1	LHB1B1 (Photosystem II light harvesting complex gene 1.4); chlorophyll binding	1,33E+04	8,14E-03
AT5G57760.1	unknown protein	1,70E+04	9,39E-03
AT5G49200.1	WD-40 repeat family protein / zfwd4 protein (ZFWD4)	-1,28E+04	1,11E-02
AT5G64870.1	unknown protein	-1,83E+04	1,73E-02
AT2G20800.1	NDB4 (NAD(P)H DEHYDROGENASE B4); NADH dehydrogenase	-1,42E+04	1,77E-02
AT1G29395.1	COR414-TM1 (cold regulated 414 thylakoid membrane 1)	1,03E+04	1,80E-02
AT5G09570.1	unknown protein	-3,21E+04	2,00E-02
AT5G52050.1	MATE efflux protein-related	-1,04E+04	2,22E-02
AT5G62520	SRO5 (SIMILAR TO RCD ONE 5); NAD+ ADP-ribosyltransferase	-1,15E+04	2,26E-02
AT5G52940.1	unknown protein	-1,43E+04	2,53E-02
AT1G01190.1	CYP78A8 (cytochrome P450, family 78, subfamily A, polypeptide 8); oxygen binding	1,03E+04	2,85E-02
AT1G18140.1	LAC1 (Laccase 1); copper ion binding / oxidoreductase	1,08E+04	2,95E-02
AT1G70440.1	SRO3 (SIMILAR TO RCD ONE 3); NAD+ ADP-ribosyltransferase	-1,98E+04	2,96E-02
AT1G28760.1	unknown protein	-1,51E+04	3,67E-02
AT5G24660.1	unknown protein	0.9497	4,64E-02
AT2G18193.1	AAA-type ATPase family protein	-1,19E+04	4,74E-02
AT1G53490.1	DNA binding	1,10E+04	4,85E-02
AT1G73830.1	BEE3 (BR ENHANCED EXPRESSION 3); DNA binding / transcription factor	1,40E+04	4,87E-02

App. Table 5. Differentially expressed genes in seedlings - diploid Col-0 (Cy5-label) vs. tetraploid Col-0 (Cy3-label; four biological replicates)

Gene ID	Description	log_2 FC	P-value
AT1G53480.1	unknown protein	5,20E+04	2,22E-06
AT5G09570.1	unknown protein	-2,66E+04	3,55E-05
AT4G13420.1	HAK5 (High affinity K+ transporter 5); potassium ion transporter	1,81E+04	2,25E-04
AT2G32210.1	unknown protein	-1,86E+04	8,31E-04
AT2G19190.1	FRK1 (FLG22-INDUCED RECEPTOR-LIKE KINASE 1); kinase	1,85E+04	8,39E-04
AT1G80660.1	AHA9 (*Arabidopsis* H(+)-ATPase 9); hydrogen-exporting ATPase, phosphorylative mechanism	-2,38E+04	1,34E-03
AT5G57760.1	unknown protein	1,62E+04	1,62E-03
AT2G20800.1	NDB4 (NAD(P)H DEHYDROGENASE B4); NADH dehydrogenase	-1,54E+04	1,74E-03
AT1G53490.1	DNA binding	1,24E+04	2,31E-03
AT2G47520.1	AP2 domain-containing transcription factor, putative	-1,31E+04	3,98E-03
AT4G28850.1	xyloglucan:xyloglucosyl transferase, putative / xyloglucan endotransglycosylase, putative / endo-	2,81E+04	5,49E-03

	xyloglucan transferase, putative		
AT5G64310.1	AGP1 (ARABINOGALACTAN-PROTEIN 1)	-1,45E+04	6,26E-03
AT4G19680	IRT2 (iron-responsive transporter 2); iron ion transporter/ zinc ion transporter	1,28E+04	6,45E-03
AT5G64870.1	unknown protein	-2,03E+04	7,40E-03
AT5G55490.1	GEX1 (GAMETE EXPRESSED PROTEIN1)	-1,08E+04	7,63E-03
AT1G20180	unknown protein	-1,19E+04	8,18E-03
AT2G28820.1	structural constituent of ribosome	-1,14E+04	8,18E-03
AT2G25460.1	unknown protein	-1,81E+04	8,33E-03
AT1G29395.1	COR414-TM1 (cold regulated 414 thylakoid membrane 1)	1,21E+04	8,39E-03
AT1G32900.1	starch synthase, putative	1,11E+04	8,67E-03
AT4G01390.1	meprin and TRAF homology domain-containing protein / MATH domain-containing protein	1,89E+04	8,89E-03
AT2G34420.1	LHB1B2 (Photosystem II light harvesting complex gene 1.5); chlorophyll binding	1,62E+04	9,24E-03
AT5G42380.1	calmodulin-related protein, putative	-1,36E+04	1,11E-02
AT5G01380.1	transcription factor	-1,43E+04	1,11E-02
AT5G62520	SRO5 (SIMILAR TO RCD ONE 5); NAD+ ADP-ribosyltransferase	-1,31E+04	1,21E-02
AT2G18193.1	AAA-type ATPase family protein	-1,26E+04	1,23E-02
AT1G18140.1	LAC1 (Laccase 1); copper ion binding / oxidoreductase	1,27E+04	1,35E-02
AT2G21640.1	unknown protein	-1,88E+04	1,52E-02
AT4G14060.1	major latex protein-related / MLP-related	1,30E+04	1,72E-02
AT1G63530.1	unknown protein	-1,68E+04	1,84E-02
AT4G38560	unknown protein	2,03E+04	1,85E-02
AT4G19000.1	IWS1 C-terminus family protein	-1,20E+04	1,94E-02
AT3G48580.1	xyloglucan:xyloglucosyl transferase, putative / xyloglucan endotransglycosylase, putative / endo-xyloglucan transferase, putative	-1,13E+04	1,99E-02
AT5G19220.1	ADG2 (ADPG PYROPHOSPHORYLASE 2); glucose-1-phosphate adenylyltransferase	1,03E+04	1,99E-02
AT1G52870	peroxisomal membrane protein-related	1,05E+04	2,03E-02
AT2G18190.1	AAA-type ATPase family protein	-1,57E+04	2,12E-02
AT4G30320.1	allergen V5/Tpx-1-related family protein	1,26E+04	2,44E-02
AT2G42270.1	U5 small nuclear ribonucleoprotein helicase, putative	-1,11E+04	2,58E-02
AT5G64910.1	unknown protein	-0.9628	2,59E-02
AT3G05800.1	transcription factor	0.9920	2,70E-02
AT3G54530.1	unknown protein	-2,03E+04	2,86E-02
AT5G19460.1	ATNUDT20 (*Arabidopsis thaliana* Nudix hydrolase homolog 20); hydrolase	1,08E+04	2,93E-02
AT5G63450.1	CYP94B1 (cytochrome P450, family 94, subfamily B, polypeptide 1); oxygen binding	1,02E+04	2,99E-02
AT5G51500.1	pectinesterase family protein	-0.9635	3,11E-02
AT1G17960.1	threonyl-tRNA synthetase, putative / threonine--tRNA ligase, putative	-0.9393	3,28E-02
AT4G05380.1	AAA-type ATPase family protein	-2,00E+04	3,38E-02
AT5G39530.1	unknown protein	0.9399	3,58E-02
AT1G35720.1	ANNAT1 (ANNEXIN ARABIDOPSIS 1); calcium ion binding / calcium-dependent phospholipid binding	0.9304	3,85E-02
AT3G24780.1	unknown protein	-0.9417	3,88E-02
AT1G69680.1	unknown protein	-1,55E+04	3,93E-02
AT3G08700.1	UBC12 (ubiquitin-conjugating enzyme 12); ubiquitin-protein ligase	-1,02E+04	4,00E-02
AT2G03230.1	unknown protein	-1,31E+04	4,08E-02

AT4G11190.1	disease resistance-responsive family protein / dirigent family protein	0.9205	4,14E-02
AT1G73830.1	BEE3 (BR ENHANCED EXPRESSION 3); DNA binding / transcription factor	1,38E+04	4,21E-02
AT1G45616.1	leucine-rich repeat family protein	-1,34E+04	4,30E-02
AT5G38020.1	S-adenosyl-L-methionine:carboxyl methyltransferase family protein	1,20E+04	4,30E-02
AT3G60160.1	ATMRP9 (*Arabidopsis thaliana* multidrug resistance-associated protein 9)	-0.8719	4,32E-02
AT1G80670.1	transducin family protein / WD-40 repeat family protein	-1,20E+04	4,35E-02
AT3G19240.1	unknown protein	-1,49E+04	4,48E-02
AT4G05050	UBQ11 (UBIQUITIN 11); protein binding	-0.8726	4,57E-02
AT3G50480.1	HR4 (HOMOLOG OF RPW8 4)	2,12E+04	4,63E-02
AT3G16770.1	ATEBP/RAP2.3 (RELATED TO AP2 3); DNA binding / protein binding / transcription factor/ transcriptional activator	1,10E+04	4,69E-02
AT4G40090.1	AGP3 (ARABINOGALACTAN-PROTEIN 3)	1,15E+04	4,78E-02
AT4G26530	fructose-bisphosphate aldolase, putative	1,18E+04	4,99E-02

App. Table 6. Differentially expressed genes in seedlings - diploid Col-0 vs. tetraploid Col-0 (two times four biological replicates, Cy3/Cy5-dye swap)

Gene ID	Description	log_2 FC	P-value
AT4G13420.1	HAK5 (High affinity K+ transporter 5); potassium ion transporter	1,7778	5,11E-08
AT5G57760.1	unknown protein	1,6590	3,68E-07
AT1G53480.1	unknown protein	4,6430	5,58E-07
AT2G20800.1	NDB4 (NAD(P)H DEHYDROGENASE B4); NADH dehydrogenase	-1,4789	8,01E-07
AT5G01380.1	transcription factor	-1,4724	1,08E-06
AT1G53490.1	DNA binding	1,1692	2,31E-06
AT5G09570.1	unknown protein	-2,9358	2,76E-06
AT5G64870.1	unknown protein	-1,9268	5,21E-06
AT1G29395.1	COR414-TM1 (cold regulated 414 thylakoid membrane 1)	1,1199	5,36E-06
AT2G18193.1	AAA-type ATPase family protein	-1,2257	6,13E-06
AT5G62520	SRO5 (SIMILAR TO RCD ONE 5); NAD+ ADP-ribosyltransferase	-1,2269	7,55E-06
AT4G28850.1	xyloglucan:xyloglucosyl transferase, putative / xyloglucan endotransglycosylase, putative / endo-xyloglucan transferase, putative	2,6847	1,09E-05
AT2G19190.1	FRK1 (FLG22-INDUCED RECEPTOR-LIKE KINASE 1); kinase	1,6762	1,16E-05
AT1G18140.1	LAC1 (Laccase 1); copper ion binding / oxidoreductase	1,1724	1,46E-05
AT5G42380.1	calmodulin-related protein, putative	-1,3647	1,50E-05
AT4G19680	IRT2 (iron-responsive transporter 2); iron ion transporter/ zinc ion transporter	1,2251	1,63E-05
AT1G73830.1	BEE3 (BR ENHANCED EXPRESSION 3); DNA binding / transcription factor	1,3919	2,13E-05
AT4G14060.1	major latex protein-related / MLP-related	1,3302	2,13E-05
AT2G23150.1	NRAMP3 (NRAMP metal ion transporter 3); manganese ion transporter/ metal ion transporter	-0,7791	2,20E-05
AT1G80660.1	AHA9 (*Arabidopsis* H(+)-ATPase 9); hydrogen-exporting ATPase, phosphorylative mechanism	-2,1788	2,24E-05
AT5G52940.1	unknown protein	-1,3623	2,42E-05
AT5G48010	pentacyclic triterpene synthase, putative	0,9256	2,97E-05
AT1G70440.1	SRO3 (SIMILAR TO RCD ONE 3); NAD+ ADP-ribosyltransferase	-2,0297	3,63E-05

AT1G28760.1	unknown protein	-1,4336	3,77E-05
AT5G54710.1	ankyrin repeat family protein	0,9805	3,83E-05
AT5G19220.1	ADG2 (ADPG PYROPHOSPHORYLASE 2); glucose-1-phosphate adenylyltransferase	1,0658	4,18E-05
AT1G17960.1	threonyl-tRNA synthetase, putative / threonine--tRNA ligase, putative	-0,8739	4,20E-05
AT1G80670.1	transducin family protein / WD-40 repeat family protein	-1,1553	5,29E-05
AT3G60160.1	ATMRP9 (*Arabidopsis thaliana* multidrug resistance-associated protein 9)	-0,9247	5,34E-05
AT5G39530.1	unknown protein	0,8820	5,65E-05
AT2G34420.1	LHB1B2 (Photosystem II light harvesting complex gene 1.5); chlorophyll binding	1,6008	5,73E-05
AT1G72800.1	nuM1-related	-1,9772	6,53E-05
AT5G38020.1	S-adenosyl-L-methionine:carboxyl methyltransferase family protein	1,2333	6,81E-05
AT4G26530	fructose-bisphosphate aldolase, putative	1,2177	7,03E-05
AT5G64910.1	unknown protein	-0,8976	7,20E-05
AT2G21640.1	unknown protein	-1,7452	7,25E-05
AT1G02205	CER1 (ECERIFERUM 1)	-0,7653	7,48E-05
AT2G32190	unknown protein	-1,4249	7,80E-05
AT5G49200.1	WD-40 repeat family protein / zfwd4 protein (ZFWD4)	-1,3498	8,61E-05
AT2G25460.1	unknown protein	-1,6491	9,47E-05
AT2G42270.1	U5 small nuclear ribonucleoprotein helicase, putative	-1,0643	9,88E-05
AT3G22840.1	ELIP1 (EARLY LIGHT-INDUCABLE PROTEIN); chlorophyll binding	1,0639	1,09E-04
AT5G51500.1	pectinesterase family protein	-0,8746	1,20E-04
AT2G47520.1	AP2 domain-containing transcription factor, putative	-1,1435	1,21E-04
AT2G41880.1	GK-1 (GUANYLATE KINASE 1); guanylate kinase	0,7794	1,22E-04
AT5G43450.1	2-oxoglutarate-dependent dioxygenase, putative	0,7524	1,64E-04
AT1G51820.1	leucine-rich repeat protein kinase, putative	0,9283	1,66E-04
AT4G30320.1	allergen V5/Tpx-1-related family protein	1,1198	1,78E-04
AT1G45616.1	leucine-rich repeat family protein	-1,3508	1,79E-04
AT1G56510.1	disease resistance protein (TIR-NBS-LRR class), putative	0,7685	1,85E-04
AT5G53230.1	unknown protein	-2,1365	1,85E-04
AT4G25790.1	allergen V5/Tpx-1-related family protein	1,2171	2,02E-04
AT2G35290.1	unknown protein	-0,6641	2,04E-04
AT5G24660.1	unknown protein	0,8464	2,08E-04
AT1G20020	ATLFNR2 (LEAF FNR 2); NADPH dehydrogenase/ oxidoreductase	0,7544	2,12E-04
AT5G64310.1	AGP1 (ARABINOGALACTAN-PROTEIN 1)	-1,2689	2,22E-04
AT2G36690.1	oxidoreductase, 2OG-Fe(II) oxygenase family protein	1,2446	2,40E-04
AT1G27020.1	unknown protein	0,9308	2,53E-04
AT5G61560.1	protein kinase family protein	0,9376	2,82E-04
AT1G24090.1	RNase H domain-containing protein	-0,7281	2,83E-04
AT5G04190.1	PKS4 (PHYTOCHROME KINASE SUBSTRATE 4)	1,6176	2,85E-04
AT3G21780.1	UGT71B6 (UDP-glucosyl transferase 71B6); UDP-glycosyltransferase/ abscisic acid glucosyltransferase/ transferase, transferring glycosyl groups	1,1940	2,98E-04
AT5G48000	CYP708A2 (cytochrome P450, family 708, subfamily A, polypeptide 2); oxygen binding	0,6854	3,06E-04
AT5G54100.1	band 7 family protein	-0,8008	3,22E-04
AT1G35720.1	ANNAT1 (ANNEXIN *ARABIDOPSIS* 1); calcium ion binding / calcium-dependent phospholipid binding	0,8494	3,29E-04
AT1G01190.1	CYP78A8 (cytochrome P450, family 78, subfamily A, polypeptide 8); oxygen binding	0,8982	3,60E-04

AT1G09080.1	luminal binding protein 3 (BiP-3) (BP3)	-2,7118	3,67E-04
AT4G11190.1	disease resistance-responsive family protein / dirigent family protein	1,1088	3,68E-04
AT4G35190.1	unknown protein	0,8277	3,76E-04
AT1G64770.1	unknown protein	1,1816	3,81E-04
AT2G28820.1	structural constituent of ribosome	-0,9812	3,88E-04
AT3G27690.1	LHCB2:4 (Photosystem II light harvesting complex gene 2.3); chlorophyll binding	2,1071	4,11E-04
AT1G49150.1	unknown protein	-1,2162	4,19E-04
AT1G21400.1	2-oxoisovalerate dehydrogenase, putative / 3-methyl-2-oxobutanoate dehydrogenase, putative / branched-chain alpha-keto acid dehydrogenase E1 alpha subunit, putative	1,4495	4,34E-04
AT5G38710.1	proline oxidase, putative / osmotic stress-responsive proline dehydrogenase, putative	-0,7026	4,39E-04
AT3G18610.1	ATRANGAP1 (RAN GTPASE-ACTIVATING PROTEIN 1); nucleic acid binding	-1,3507	4,51E-04
AT4G22214.1	unknown protein	2,4334	4,55E-04
AT5G19470.1	ATNUDT24 (*Arabidopsis thaliana* Nudix hydrolase homolog 24); hydrolase	1,0018	4,80E-04
AT3G15720	glycoside hydrolase family 28 protein / polygalacturonase (pectinase) family protein	-0,7568	4,89E-04
AT4G34410.1	AP2 domain-containing transcription factor, putative	-1,5196	4,98E-04
AT4G38560	unknown protein	1,8156	5,04E-04
AT2G19800.1	MIOX2 (MYO-INOSITOL OXYGENASE 2)	0,8087	5,08E-04
AT4G19000.1	IWS1 C-terminus family protein	-1,0558	5,12E-04
AT3G22060.1	receptor protein kinase-related	1,5308	5,19E-04
AT3G50480.1	HR4 (HOMOLOG OF RPW8 4)	2,0668	5,48E-04
AT3G61080.1	fructosamine kinase family protein	0,7477	5,61E-04
AT4G36880.1	cysteine proteinase, putative	-0,8817	5,68E-04
AT3G16690.1	nodulin MtN3 family protein	0,8286	6,02E-04
AT1G75860.1	unknown protein	-1,0277	6,17E-04
AT5G63450.1	CYP94B1 (cytochrome P450, family 94, subfamily B, polypeptide 1); oxygen binding	0,9185	6,28E-04
AT5G52050.1	MATE efflux protein-related	-0,8731	6,59E-04
AT4G01630.1	ATEXPA17 (*ARABIDOPSIS THALIANA* EXPANSIN A17)	-1,6324	6,87E-04
AT1G11655.1	unknown protein	0,8602	6,91E-04
AT1G31290.1	PAZ domain-containing protein / piwi domain-containing protein	-1,5068	6,97E-04
AT3G11020.1	DREB2B (DRE-binding protein 2B); DNA binding / transcription factor/ transcriptional activator	-1,4454	7,01E-04
AT4G02270.1	pollen Ole e 1 allergen and extensin family protein	1,0152	7,05E-04
AT5G03545.1	unknown protein	1,5771	7,10E-04
AT4G23200.1	protein kinase family protein	0,8237	7,36E-04
AT1G22990.1	heavy-metal-associated domain-containing protein / copper chaperone (CCH)-related	-0,7542	7,40E-04
AT5G62330.1	unknown protein	0,8697	7,67E-04
AT1G59660.1	nucleoporin family protein	-2,1289	8,26E-04
AT2G34430.1	LHB1B1 (Photosystem II light harvesting complex gene 1.4); chlorophyll binding	1,1014	8,43E-04
AT4G30150.1	unknown protein	-0,7588	8,65E-04
AT3G08940	LHCB4.2 (LIGHT HARVESTING COMPLEX PSII); chlorophyll binding	1,5177	8,73E-04
AT1G10340	ankyrin repeat family protein	1,4050	8,75E-04
AT3G19240.1	unknown protein	-1,3567	8,78E-04
AT1G02220.1	ANAC003 (*Arabidopsis* NAC domain containing protein 3);	-1,1545	9,07E-04

	transcription factor		
AT1G64370.1	unknown protein	0,8532	9,23E-04
AT3G01490.1	protein kinase, putative	0,7346	9,26E-04
AT1G63530.1	unknown protein	-1,5346	9,47E-04
AT2G42330	D111/G-patch domain-containing protein	-0,7107	1,05E-03
AT1G77840.1	eukaryotic translation initiation factor 5, putative / eIF-5, putative	-0,5443	1,08E-03
AT3G25690.1	CHUP1 (CHLOROPLAST UNUSUAL POSITIONING 1)	0,5525	1,13E-03
AT4G16990	disease resistance protein (TIR-NBS class), putative	0,7374	1,14E-03
AT4G28270.1	zinc finger (C3HC4-type RING finger) family protein	0,6009	1,15E-03
AT4G30430.1	TET9 (TETRASPANIN9)	-1,5603	1,21E-03
AT1G20140.1	ASK4 (*ARABIDOPSIS* SKP1-LIKE 4); ubiquitin-protein ligase	-0,7940	1,24E-03
AT1G10070	ATBCAT-2; branched-chain-amino-acid transaminase/ catalytic	1,0991	1,26E-03
AT1G69680.1	unknown protein	-1,3040	1,26E-03
AT4G37930.1	SHM1 (SERINE HYDROXYMETHYLTRANSFERASE 1); glycine hydroxymethyltransferase	0,9887	1,31E-03
AT2G42750.1	DNAJ heat shock N-terminal domain-containing protein	0,6425	1,32E-03
AT1G43160.1	RAP2.6 (related to AP2 6); DNA binding / transcription factor	0,9746	1,42E-03
AT4G18810.1	transcriptional repressor	0,6409	1,42E-03
AT5G13200.1	GRAM domain-containing protein / ABA-responsive protein-related	-0,8655	1,45E-03
AT5G10230.1	ANN7 (ANN7, ANNEXIN *ARABIDOPSIS* 7); calcium ion binding / calcium-dependent phospholipid binding	1,3482	1,49E-03
AT1G56510.1	disease resistance protein (TIR-NBS-LRR class), putative	0,6820	1,52E-03
AT3G01600.1	ANAC044 (*Arabidopsis* NAC domain containing protein 44); transcription factor	-0,5372	1,53E-03
AT1G55110.1	*ARABIDOPSIS THALIANA* INDETERMINATE(ID)-DOMAIN 7/ATIDD7; nucleic acid binding / transcription factor/ zinc ion binding	0,6347	1,53E-03
AT1G05250.1	peroxidase, putative	0,6395	1,57E-03
AT1G32900.1	starch synthase, putative	0,9374	1,58E-03
AT2G19500.1	CKX2 (CYTOKININ OXIDASE 2); cytokinin dehydrogenase	-0,9315	1,63E-03
AT5G04230.1	PAL3 (PHENYL ALANINE AMMONIA-LYASE 3); phenylalanine ammonia-lyase	0,7424	1,64E-03
AT2G39980.1	transferase family protein	1,1686	1,64E-03
AT1G69490.1	NAP (NAC-LIKE, ACTIVATED BY AP3/PI); transcription factor	1,3195	1,66E-03
AT5G17230	PSY (PHYTOENE SYNTHASE); geranylgeranyl-diphosphate geranylgeranyltransferase	0,6668	1,66E-03
AT5G66480.1	unknown protein	-0,9227	1,67E-03
AT4G26860.1	alanine racemase family protein	0,7778	1,70E-03
AT3G07490.1	AGD11 (ARF-GAP DOMAIN 11); calcium ion binding	0,8623	1,70E-03
AT5G56080.1	nicotianamine synthase, putative	1,1809	1,77E-03
AT3G58270	meprin and TRAF homology domain-containing protein / MATH domain-containing protein	-0,9746	1,81E-03
AT1G55020.1	LOX1 (Lipoxygenase 1); lipoxygenase	-0,8080	1,85E-03
AT1G72660	developmentally regulated GTP-binding protein, putative	-1,4991	1,88E-03
AT5G53590.1	auxin-responsive family protein	-0,7421	1,90E-03
AT5G60250.1	zinc finger (C3HC4-type RING finger) family protein	-1,8817	1,95E-03
AT5G57625.1	allergen V5/Tpx-1-related family protein	0,9263	1,95E-03
AT5G60680.1	unknown protein	0,8205	1,99E-03
AT4G35770	SEN1 (DARK INDUCIBLE 1)	1,4648	2,18E-03

AT3G61400.1	2-oxoglutarate-dependent dioxygenase, putative	-0,9654	2,19E-03
AT3G58150.1	unknown protein	-0,9879	2,25E-03
AT2G03230.1	unknown protein	-1,2739	2,25E-03
AT5G67400.1	peroxidase 73 (PER73) (P73) (PRXR11)	0,7720	2,25E-03
AT5G44580.1	unknown protein	1,4414	2,39E-03
AT4G33470.1	HDA14 (histone deacetylase 14); histone deacetylase	0,5889	2,40E-03
AT1G74650.1	AtMYB31/AtY13 (myb domain protein 31); DNA binding / transcription factor	1,0038	2,44E-03
AT5G43750.1	unknown protein	1,1120	2,46E-03
AT5G51130.1	unknown protein	-0,6412	2,52E-03
AT1G14960.1	major latex protein-related / MLP-related	0,8887	2,55E-03
AT1G49930.1	unknown protein	-1,1967	2,59E-03
AT1G53790.1	F-box family protein	-0,8570	2,61E-03
AT3G08810.1	kelch repeat-containing F-box family protein	-1,5362	2,63E-03
AT4G27300.1	S-locus protein kinase, putative	0,6179	2,69E-03
AT3G05500.1	rubber elongation factor (REF) family protein	-0,5383	2,75E-03
AT5G18180.1	Gar1 RNA-binding region family protein	-0,5651	2,78E-03
AT5G50450.1	zinc finger (MYND type) family protein	0,6959	2,79E-03
AT3G02730.1	ATF1/TRXF1 (THIOREDOXIN F-TYPE 1); thiol-disulfide exchange intermediate	0,9666	2,81E-03
AT3G21150.1	zinc finger (B-box type) family protein	0,5776	2,86E-03
AT5G46110	APE2 (ACCLIMATION OF PHOTOSYNTHESIS TO ENVIRONMENT); antiporter/ triose-phosphate transporter	1,0245	3,00E-03
AT1G64590.1	short-chain dehydrogenase/reductase (SDR) family protein	0,6798	3,10E-03
AT3G60330.1	AHA7 (*ARABIDOPSIS* H(+)-ATPASE 7); hydrogen-exporting ATPase, phosphorylative mechanism	0,8079	3,12E-03
AT1G26770	ATEXPA10 (*ARABIDOPSIS THALIANA* EXPANSIN A10)	-0,7070	3,15E-03
AT5G42680.1	unknown protein	0,8895	3,16E-03
AT3G22910.1	calcium-transporting ATPase, plasma membrane-type, putative / Ca(2+)-ATPase, putative (ACA13)	-1,3064	3,22E-03
AT5G39580	peroxidase, putative	1,1353	3,23E-03
AT2G18180.1	SEC14 cytosolic factor, putative / phosphoglyceride transfer protein, putative	-1,1275	3,29E-03
AT1G70130.1	lectin protein kinase, putative	-0,6108	3,35E-03
AT1G23770.1	F-box family protein	-1,0657	3,38E-03
AT1G18730	unknown protein	0,9435	3,43E-03
AT2G20520.1	FLA6 (FLA6)	1,5204	3,44E-03
AT5G13210.1	unknown protein	-1,1076	3,48E-03
AT3G46270.1	receptor protein kinase-related	0,7727	3,49E-03
AT2G02350.1	SKIP3 (SKP1 INTERACTING PARTNER 3)	-0,5521	3,50E-03
AT1G13340.1	unknown protein	-0,8408	3,62E-03
AT5G55490.1	GEX1 (GAMETE EXPRESSED PROTEIN1)	-0,8671	3,68E-03
AT4G01390.1	meprin and TRAF homology domain-containing protein / MATH domain-containing protein	1,5617	3,68E-03
AT4G22217.1	unknown protein	2,0732	3,78E-03
AT1G53540.1	17.6 kDa class I small heat shock protein (HSP17.6C-CI) (AA 1-156)	-2,1800	3,80E-03
AT3G01500	CA1 (CARBONIC ANHYDRASE 1); carbonate dehydratase/ zinc ion binding	1,5264	3,83E-03
AT3G21680.1	unknown protein	-0,5506	3,84E-03
AT4G21490.1	pyridine nucleotide-disulphide oxidoreductase family protein	-0,7975	3,85E-03
AT5G23980.1	ATFRO4/FRO4 (FERRIC REDUCTION OXIDASE 4); ferric-chelate reductase	1,0193	3,88E-03
AT1G57770.1	amine oxidase family	0,9878	3,95E-03

AT3G03110.1	XPO1B (exportin 1B); protein transporter	-0,8047	3,95E-03
AT5G04960.1	pectinesterase family protein	0,8397	3,99E-03
AT5G43330.1	malate dehydrogenase, cytosolic, putative	-0,8314	4,00E-03
AT3G56710.1	SIB1 (SIGMA FACTOR BINDING PROTEIN 1); binding	-0,7820	4,00E-03
AT1G14280.1	PKS2 (PHYTOCHROME KINASE SUBSTRATE 2)	0,9119	4,04E-03
AT5G17020.1	XPO1A (exportin 1A); protein transporter	-0,6270	4,25E-03
AT1G04310.1	ERS2 (ETHYLENE RESPONSE SENSOR 2); receptor	0,5787	4,25E-03
AT3G52340	SPP2 (SPP2); sucrose-phosphatase	0,5179	4,28E-03
AT4G32785.1	unknown protein	0,5392	4,31E-03
AT1G25470	AP2 domain-containing transcription factor, putative	-0,6095	4,35E-03
AT1G32080.1	membrane protein, putative	0,7735	4,43E-03
AT1G01790.1	KEA1 (K EFFLUX ANTIPORTER 1); potassium:hydrogen antiporter	0,7431	4,46E-03
AT1G71150.1	unknown protein	-0,7986	4,54E-03
AT2G34350.1	nodulin-related	0,9336	4,62E-03
AT1G34760	GRF11 (General regulatory factor 11); amino acid binding / protein phosphorylated amino acid binding	0,6701	4,62E-03
AT1G72280.1	AERO1 (*ARABIDOPSIS* ENDOPLASMIC RETICULUM OXIDOREDUCTINS 1)	-0,5842	4,66E-03
AT3G04800.1	ATTIM23-3 (*Arabidopsis thaliana* translocase inner membrane subunit 23-3); protein translocase	-0,5599	4,71E-03
AT5G61390.1	exonuclease family protein	-1,1029	4,73E-03
AT3G12700.1	aspartyl protease family protein	-0,7417	4,74E-03
AT1G50400.1	porin family protein	-0,9055	4,74E-03
AT3G49400.1	transducin family protein / WD-40 repeat family protein	-0,5642	4,76E-03
AT1G18830.1	transducin family protein / WD-40 repeat family protein	-2,9332	4,78E-03
AT4G03430.1	STA1 (STABILIZED1); RNA splicing factor, transesterification mechanism	-0,7208	4,79E-03
AT3G01650.1	copine-related	-0,8204	4,80E-03
AT1G72930	TIR (TOLL/INTERLEUKIN-1 RECEPTOR-LIKE); transmembrane receptor	1,0015	4,80E-03
AT1G60030.1	xanthine/uracil permease family protein	-0,6847	4,82E-03
AT4G10340.1	LHCB5 (LIGHT HARVESTING COMPLEX OF PHOTOSYSTEM II 5); chlorophyll binding	0,9256	4,82E-03
AT1G35730.1	APUM9 (*ARABIDOPSIS* PUMILIO 9); RNA binding	-0,6633	4,89E-03
AT3G07720.1	kelch repeat-containing protein	0,6581	4,95E-03
AT5G41080	glycerophosphoryl diester phosphodiesterase family protein	0,9014	4,97E-03
AT1G62380.1	ACO2 (ACC OXIDASE 2)	0,6689	4,98E-03
AT5G40880.1	WD-40 repeat family protein / zfwd3 protein (ZFWD3)	-0,8725	5,08E-03
AT4G08620.1	SULTR1;1 (sulfate transporter 1;1); sulfate transporter	0,8329	5,10E-03
AT5G56870.1	beta-galactosidase, putative / lactase, putative	1,2741	5,11E-03
AT3G02840.1	immediate-early fungal elicitor family protein	-1,2379	5,12E-03
AT1G80470.1	unknown protein	-0,9570	5,17E-03
AT5G01540.1	lectin protein kinase, putative	0,7339	5,20E-03
AT5G63900.1	PHD finger family protein	-1,4252	5,21E-03
AT5G43980.1	receptor-like protein kinase-related	-1,2437	5,27E-03
AT5G05490	DIF1/SYN1	-0,5849	5,31E-03
AT3G26280.1	CYP71B4 (cytochrome P450, family 71, subfamily B, polypeptide 4); oxygen binding	0,6712	5,32E-03
AT5G64190.1	unknown protein	-1,5346	5,36E-03
AT5G55270.1	unknown protein	-0,7710	5,41E-03
AT1G62780.1	unknown protein	0,6275	5,51E-03
AT5G02030.1	LSN (LARSON, VAAMANA); DNA binding / transcription factor	0,4929	5,56E-03

AT5G49360.1	BXL1 (BETA-XYLOSIDASE 1); hydrolase, hydrolyzing O-glycosyl compounds	1,2490	5,64E-03
AT3G19470	F-box family protein	-0,8364	5,75E-03
AT5G16360.1	NC domain-containing protein	-0,6720	5,81E-03
AT4G23290	protein kinase family protein	0,7895	5,84E-03
AT4G25410.1	basix helix-loop-helix (bHLH) family protein	0,5424	5,85E-03
AT2G38340.1	AP2 domain-containing transcription factor, putative (DRE2B)	-1,0691	5,87E-03
AT2G17740.1	DC1 domain-containing protein	1,0232	5,88E-03
AT1G06450.1	CCR4-NOT transcription complex protein, putative	-0,7319	5,90E-03
AT4G01890.1	glycoside hydrolase family 28 protein / polygalacturonase (pectinase) family protein	0,8799	5,94E-03
AT5G38990.1	protein kinase family protein	0,7303	6,04E-03
AT3G50330.1	basic helix-loop-helix (bHLH) family protein	1,2522	6,11E-03
AT3G03640.1	GLUC (Beta-glucosidase homolog); hydrolase, hydrolyzing O-glycosyl compounds	-0,8268	6,24E-03
AT1G79380.1	copine-related	0,5723	6,27E-03
AT5G44390.1	FAD-binding domain-containing protein	0,7802	6,28E-03
AT5G14800	P5CR (PYRROLINE-5- CARBOXYLATE (P5C) REDUCTASE); pyrroline-5-carboxylate reductase	-0,6150	6,29E-03
AT1G61470.1	CCR4-NOT transcription complex protein, putative	-0,7173	6,29E-03
AT3G12510.1	unknown protein	-1,1049	6,31E-03
AT5G52450.1	MATE efflux protein-related	0,6359	6,43E-03
AT5G14780.1	FDH (FORMATE DEHYDROGENASE); NAD binding / cofactor binding / oxidoreductase, acting on the CH-OH group of donors, NAD or NADP as acceptor	0,7115	6,43E-03
AT5G42510.1	disease resistance-responsive family protein	0,7463	6,46E-03
AT3G54530.1	unknown protein	-1,6445	6,48E-03
AT3G29810.1	phytochelatin synthetase family protein / COBRA cell expansion protein COBL2	-0,9612	6,57E-03
AT2G30930.1	unknown protein	0,6599	6,61E-03
AT5G64060.1	ANAC103 (*Arabidopsis* NAC domain containing protein 103); transcription factor	-1,4441	6,76E-03
AT3G50940.1	AAA-type ATPase family protein	-0,9709	6,96E-03
AT4G05380.1	AAA-type ATPase family protein	-1,6594	6,98E-03
AT4G27580.1	unknown protein	-0,5488	7,06E-03
AT4G32050.1	neurochondrin family protein	-0,6017	7,14E-03
AT3G12345.1	unknown protein	1,0085	7,15E-03
AT4G38780.1	splicing factor, putative	-0,9479	7,24E-03
AT1G53680.1	ATGSTU28 (*Arabidopsis thaliana* Glutathione S-transferase (class tau) 28); glutathione transferase	0,7798	7,26E-03
AT1G19660	wound-responsive family protein	0,5535	7,38E-03
AT2G40340.1	AP2 domain-containing transcription factor, putative (DRE2B)	-1,1458	7,42E-03
AT5G15870.1	glycosyl hydrolase family 81 protein	-0,6507	7,45E-03
AT1G56600.1	ATGOLS2 (*ARABIDOPSIS THALIANA* GALACTINOL SYNTHASE 2); transferase, transferring glycosyl groups / transferase, transferring hexosyl groups	0,7651	7,45E-03
AT4G38810	calcium-binding EF hand family protein	0,5254	7,48E-03
AT5G47990.1	CYP705A5 (cytochrome P450, family 705, subfamily A, polypeptide 5); oxygen binding	0,6867	7,54E-03
AT1G69160.1	unknown protein	0,7193	7,73E-03
AT5G57685.1	unknown protein	0,6090	7,73E-03
AT1G02770.1	unknown protein	-0,9348	7,76E-03
AT1G43670.1	fructose-1,6-bisphosphatase, putative / D-fructose-1,6-	0,8890	7,95E-03

	bisphosphate 1-phosphohydrolase, putative / FBPase, putative		
AT1G52870	peroxisomal membrane protein-related	0,8621	7,96E-03
AT3G13380.1	BRL3 (BRI1-LIKE 3); protein binding / protein kinase	-0,5669	8,21E-03
AT3G03950	ECT1; protein binding	-0,5843	8,28E-03
AT5G46050.1	ATPTR3/PTR3 (PEPTIDE TRANSPORTER PROTEIN 3); transporter	0,6573	8,29E-03
AT1G75100.1	JAC1 (J-DOMAIN PROTEIN REQUIRED FOR CHLOROPLAST ACCUMULATION RESPONSE 1); heat shock protein binding	0,8724	8,47E-03
AT3G28300.1	AT14A	1,4740	8,50E-03
AT5G20010.1	RAN-1 (Ras-related GTP-binding nuclear protein 1); GTP binding	-0,9551	8,57E-03
AT4G37700.1	unknown protein	0,7223	8,64E-03
AT3G26230.1	CYP71B24 (cytochrome P450, family 71, subfamily B, polypeptide 24); oxygen binding	0,8501	8,70E-03
AT1G55480.1	binding / protein binding	0,8186	8,80E-03
AT3G22420	WNK2 (WITH NO K 2); kinase	0,6075	8,82E-03
AT3G01440.1	oxygen evolving enhancer 3 (PsbQ) family protein	0,9759	8,85E-03
AT5G05600.1	oxidoreductase, 2OG-Fe(II) oxygenase family protein	0,7332	8,86E-03
AT4G26010.1	peroxidase, putative	0,8170	8,98E-03
AT1G66600.1	WRKY63 (WRKY DNA-binding protein 63); transcription factor	-1,0785	8,99E-03
AT5G48657	defense protein-related	0,7379	9,14E-03
AT5G23660.1	MTN3 (*ARABIDOPSIS* HOMOLOG OF MEDICAGO TRUNCATULA MTN3)	0,7425	9,19E-03
AT5G39380.1	calmodulin-binding protein-related	-0,5548	9,24E-03
AT1G06350.1	fatty acid desaturase family protein	-0,9836	9,24E-03
AT4G10120	ATSPS4F; sucrose-phosphate synthase/ transferase, transferring glycosyl groups	0,7621	9,28E-03
AT1G29100.1	copper-binding family protein	-0,6805	9,29E-03
AT2G05380.1	GRP3S (GLYCINE-RICH PROTEIN 3 SHORT ISOFORM)	1,3291	9,44E-03
AT4G22470.1	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein	1,1907	9,54E-03
AT1G74710.1	ICS1 (ISOCHORISMATE SYNTHASEI); isochorismate synthase	1,5155	9,61E-03
AT4G12040	zinc finger (AN1-like) family protein	-0,5534	9,66E-03
AT3G14590.1	unknown protein	-0,7128	9,69E-03
AT5G57510.1	unknown protein	-0,9246	9,75E-03
AT1G17020.1	SRG1 (SENESCENCE-RELATED GENE 1); oxidoreductase, acting on paired donors, with incorporation or reduction of molecular oxygen, 2-oxoglutarate as one donor, and incorporation of one atom each of oxygen into both donors	-0,6761	9,76E-03
AT1G51620.1	protein kinase family protein	0,7823	9,77E-03
AT4G17695.1	KAN3 (KANADI 3); DNA binding / transcription factor	0,5615	9,84E-03
AT3G16390.1	jacalin lectin family protein	1,3643	9,85E-03
AT1G16530.1	LOB domain protein 3 / lateral organ boundaries domain protein 3 (LBD3)	-0,6982	9,85E-03
AT1G62980.1	ATEXPA18 (*ARABIDOPSIS THALIANA* EXPANSIN A18)	0,8986	9,85E-03
AT4G01360.1	unknown protein	-0,7700	9,87E-03
AT1G33600.1	leucine-rich repeat family protein	0,7282	9,89E-03
AT5G64710	unknown protein	-0,5855	9,94E-03
AT1G32060.1	PRK (PHOSPHORIBULOKINASE); ATP binding / phosphoribulokinase/ protein binding	0,8418	1,00E-02
AT1G25230.1	purple acid phosphatase family protein	0,7982	1,00E-02

AT5G60660.1	PIP2;4/PIP2F (plasma membrane intrinsic protein 2;4); water channel	1,0072	1,01E-02
AT3G21080.1	ABC transporter-related	1,7261	1,01E-02
AT1G76020.1	unknown protein	0,6695	1,02E-02
AT5G58290.1	RPT3 (root phototropism 3); ATPase	-0,5443	1,02E-02
AT2G34720.1	CCAAT-binding transcription factor (CBF-B/NF-YA) family protein	0,5768	1,02E-02
AT1G05660.1	polygalacturonase, putative / pectinase, putative	-0,7637	1,02E-02
AT5G47070.1	protein kinase, putative	-0,4555	1,03E-02
AT1G74710	ICS1 (ISOCHORISMATE SYNTHASEI); isochorismate synthase	1,5332	1,03E-02
ATMG00516.1	unknown protein	-1,6776	1,04E-02
AT5G04120.1	phosphoglycerate/bisphosphoglycerate mutase family protein	1,9704	1,06E-02
AT5G18400	unknown protein	-0,6341	1,06E-02
AT5G61650.1	CYCP4;2 (CYCLIN P4;2); cyclin-dependent protein kinase	0,8674	1,06E-02
AT3G51910.1	AT-HSFA7A (*Arabidopsis thaliana* heat shock transcription factor A7A); DNA binding / transcription factor	1,3232	1,07E-02
AT2G18780.1	F-box family protein	-0,7678	1,07E-02
AT3G15360.1	ATHM4 (*Arabidopsis* thioredoxin M-type 4); thiol-disulfide exchange intermediate	0,8030	1,07E-02
AT5G63160.1	BT1 (BTB and TAZ domain protein 1); protein binding / transcription regulator	1,1460	1,07E-02
AT1G55670.1	PSAG	1,0765	1,08E-02
AT4G24420.1	RNA recognition motif (RRM)-containing protein	-0,8879	1,09E-02
AT4G10060.1	unknown protein	0,5578	1,09E-02
AT5G19460.1	ATNUDT20 (*Arabidopsis thaliana* Nudix hydrolase homolog 20); hydrolase	0,9102	1,10E-02
AT5G49520.1	WRKY48 (WRKY DNA-binding protein 48); transcription factor	-0,5789	1,11E-02
ATMG01320	unknown protein	-0,7683	1,11E-02
AT4G14690.1	ELIP2 (EARLY LIGHT-INDUCIBLE PROTEIN 2); chlorophyll binding	1,7862	1,13E-02
AT3G46780.1	PTAC16 (PLASTID TRANSCRIPTIONALLY ACTIVE18)	1,0114	1,13E-02
AT1G42970.1	GAPB (GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE B SUBUNIT); glyceraldehyde-3-phosphate dehydrogenase	0,8297	1,15E-02
AT4G13260.1	YUC2 (YUCCA2); monooxygenase/ oxidoreductase	-0,6792	1,15E-02
AT3G52720.1	carbonic anhydrase family protein	0,5810	1,17E-02
AT2G29290.1	tropinone reductase, putative / tropine dehydrogenase, putative	0,7289	1,17E-02
AT2G18190.1	AAA-type ATPase family protein	-1,2468	1,17E-02
AT5G18270	ANAC087; transcription factor	-0,7121	1,18E-02
AT2G30570	photosystem II reaction center W (PsbW) protein-related	0,9580	1,18E-02
AT5G52670.1	heavy-metal-associated domain-containing protein	0,6279	1,18E-02
AT5G06060.1	tropinone reductase, putative / tropine dehydrogenase, putative	0,5777	1,21E-02
AT1G64540.1	F-box family protein	-0,8709	1,24E-02
AT5G11100.1	C2 domain-containing protein	-0,4685	1,25E-02
AT1G09530	PIF3 (PHYTOCHROME INTERACTING FACTOR 3); DNA binding / transcription factor/ transcription regulator	0,8112	1,26E-02
AT1G61380.1	S-locus protein kinase, putative	0,6108	1,27E-02
AT5G36150.1	ATPEN3 (PUTATIVE PENTACYCLIC TRITERPENE SYNTHASE 3); catalytic/ lupeol synthase	1,7907	1,29E-02
ATMG01080.1	unknown protein	-1,0903	1,30E-02
AT5G05580.1	FAD8 (FATTY ACID DESATURASE 8); omega-3 fatty acid	0,7042	1,31E-02

	desaturase		
AT2G38390.1	peroxidase, putative	-0,5631	1,31E-02
AT2G25940.1	ALPHA-VPE (ALPHA-VACUOLAR PROCESSING ENZYME); cysteine-type endopeptidase	-0,4596	1,32E-02
AT4G34580.1	transporter	0,8722	1,33E-02
AT3G45970.1	ATEXLA1 (*ARABIDOPSIS THALIANA* EXPANSIN-LIKE A1)	-0,5762	1,33E-02
AT5G09420.1	chloroplast outer membrane translocon subunit, putative	-0,5175	1,33E-02
AT4G35940.1	unknown protein	-0,4429	1,35E-02
ATMG00220	unknown protein	-1,1908	1,36E-02
AT5G21940.1	unknown protein	0,7855	1,36E-02
AT5G01880.1	zinc finger (C3HC4-type RING finger) family protein	0,5853	1,36E-02
AT3G50610.1	unknown protein	-0,8843	1,36E-02
AT1G03130.1	PSAD-2 (photosystem I subunit D-2)	1,0521	1,37E-02
AT1G22400.1	UGT85A1 (UDP-glucosyl transferase 85A1); UDP-glycosyltransferase/ transferase, transferring glycosyl groups / transferase, transferring hexosyl groups	1,4642	1,37E-02
AT2G32460	AtM1/AtMYB101/MYB101 (myb domain protein 101); DNA binding / transcription factor	-0,4371	1,38E-02
AT3G15630.1	unknown protein	1,0661	1,38E-02
AT3G59340.1	unknown protein	0,6781	1,38E-02
AT1G71140.1	MATE efflux family protein	-0,8166	1,41E-02
AT2G42650.1	60S ribosomal protein-related	-0,4694	1,42E-02
AT4G30270.1	MERI5B (MERISTEM-5); hydrolase, acting on glycosyl bonds	-0,6950	1,42E-02
AT3G26650.1	GAPA (GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE A SUBUNIT); glyceraldehyde-3-phosphate dehydrogenase	0,7082	1,44E-02
AT1G05650.1	polygalacturonase, putative / pectinase, putative	-0,8169	1,46E-02
AT1G08630	THA1 (THREONINE ALDOLASE 1); aldehyde-lyase/ threonine aldolase	1,1864	1,47E-02
AT1G14480.1	protein binding	-0,4340	1,49E-02
AT2G35260.1	unknown protein	0,7430	1,51E-02
AT1G29390	COR314-TM2 (cold regulated 314 thylakoid membrane 2)	0,5546	1,52E-02
AT1G75190	unknown protein	0,5409	1,52E-02
AT5G37990.1	S-adenosylmethionine-dependent methyltransferase	0,5953	1,53E-02
AT4G21800	ATP-binding family protein	-0,4383	1,54E-02
AT2G43535.1	trypsin inhibitor, putative	0,5646	1,54E-02
AT2G21530.1	forkhead-associated domain-containing protein / FHA domain-containing protein	0,7180	1,56E-02
AT4G15530	PPDK (PYRUVATE ORTHOPHOSPHATE DIKINASE); kinase/ pyruvate, phosphate dikinase	0,6454	1,58E-02
AT1G18870.1	ICS2; isochorismate synthase	0,8405	1,59E-02
AT5G21130.1	unknown protein	0,6575	1,59E-02
AT5G13610.1	unknown protein	-0,5422	1,62E-02
AT2G05540.1	glycine-rich protein	0,8172	1,63E-02
AT3G52780	ATPAP20/PAP20; acid phosphatase/ protein serine/threonine phosphatase	-0,7013	1,66E-02
AT3G19760.1	eukaryotic translation initiation factor 4A, putative / eIF-4A, putative / DEAD box RNA helicase, putative	-0,4873	1,66E-02
AT4G01920.1	DC1 domain-containing protein	-0,4489	1,68E-02
AT2G36050.1	ATOFP15/OFP15 (*Arabidopsis thaliana* ovate family protein 15)	0,7398	1,68E-02
AT2G18370.1	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein	-0,6497	1,69E-02

AT2G05270.1	unknown protein	-0,6451	1,69E-02
AT1G18400.1	BEE1 (BR ENHANCED EXPRESSION 1); transcription factor	0,6719	1,71E-02
AT3G19920.1	unknown protein	-1,0117	1,74E-02
AT1G18270.1	ketose-bisphosphate aldolase class-II family protein	0,4870	1,74E-02
AT2G15890	MEE14 (maternal effect embryo arrest 14)	0,5749	1,74E-02
AT1G60470.1	ATGOLS4 (*ARABIDOPSIS THALIANA* GALACTINOL SYNTHASE 4); transferase, transferring glycosyl groups / transferase, transferring hexosyl groups	1,2875	1,75E-02
AT5G26667.1	uridylate kinase / uridine monophosphate kinase / UMP kinase (PYR6)	0,5664	1,75E-02
AT1G28140.1	unknown protein	0,5224	1,76E-02
AT2G04039	unknown protein	0,8317	1,76E-02
AT4G01110.1	unknown protein	0,7696	1,77E-02
AT2G21960.1	unknown protein	0,8233	1,81E-02
AT1G25240.1	epsin N-terminal homology (ENTH) domain-containing protein / clathrin assembly protein-related	1,1157	1,84E-02
AT1G18940.1	nodulin family protein	0,5360	1,85E-02
AT5G19890.1	peroxidase, putative	1,1086	1,86E-02
AT1G66100.1	thionin, putative	1,3544	1,87E-02
AT5G45650.1	subtilase family protein	-0,6148	1,89E-02
AT3G20500.1	ATPAP18/PAP18 (purple acid phosphatase 18); acid phosphatase/ protein serine/threonine phosphatase	-0,7511	1,89E-02
AT5G59490.1	haloacid dehalogenase-like hydrolase family protein	-0,4707	1,90E-02
AT2G46610	arginine/serine-rich splicing factor, putative	-0,8613	1,91E-02
AT2G16660.1	nodulin family protein	-0,8758	1,92E-02
AT3G53470	unknown protein	0,6906	1,95E-02
AT5G06990.1	unknown protein	0,8602	1,95E-02
AT3G26210.1	CYP71B23 (cytochrome P450, family 71, subfamily B, polypeptide 23); oxygen binding	0,5054	1,96E-02
AT5G13410.1	immunophilin / FKBP-type peptidyl-prolyl cis-trans isomerase family protein	0,5236	1,96E-02
AT5G64990.1	AtRABH1a (*Arabidopsis* Rab GTPase homolog H1a); GTP binding	-0,7075	1,98E-02
AT5G24600.1	unknown protein	-0,8694	1,98E-02
AT3G60280.1	UCC3 (UCLACYANIN 3); copper ion binding	0,7841	2,03E-02
AT3G53980	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein	1,4646	2,05E-02
AT1G60970.1	clathrin adaptor complex small chain family protein	-0,6375	2,07E-02
AT2G25780.1	unknown protein	0,6372	2,07E-02
AT3G44690.1	unknown protein	-0,6499	2,08E-02
AT5G44575.1	unknown protein	1,1765	2,09E-02
AT3G19850.1	phototropic-responsive NPH3 family protein	0,8337	2,09E-02
AT3G47470.1	LHCA4 (Photosystem I light harvesting complex gene 4); chlorophyll binding	0,9811	2,13E-02
AT2G41640.1	unknown protein	-0,6498	2,13E-02
AT5G22540.1	unknown protein	-1,1448	2,14E-02
AT5G22410.1	peroxidase, putative	1,3743	2,15E-02
AT1G16720.1	oxidoreductase/ transcriptional repressor	0,7191	2,16E-02
AT5G02760.1	protein phosphatase 2C family protein / PP2C family protein	1,0981	2,16E-02
AT1G79720.1	aspartyl protease family protein	0,6317	2,18E-02
AT2G20500.1	unknown protein	-0,7548	2,18E-02
AT2G32540.1	ATCSLB04 (Cellulose synthase-like B4); transferase/ transferase, transferring glycosyl groups	0,7990	2,19E-02

AT5G23120.1	HCF136 (High chlorophyll fluorescence 136)	0,7924	2,20E-02
AT4G36010	pathogenesis-related thaumatin family protein	-1,0291	2,20E-02
AT2G31230.1	ATERF15 (ETHYLENE-RESPONSIVE ELEMENT BINDING FACTOR 15); DNA binding / transcription factor/ transcriptional activator	0,7377	2,28E-02
AT5G14730.1	unknown protein	-0,5533	2,28E-02
AT1G35420.1	dienelactone hydrolase family protein	0,5835	2,31E-02
AT5G03880.1	unknown protein	0,6341	2,31E-02
ATMG00480	unknown protein	-1,7708	2,33E-02
AT1G17190.1	ATGSTU26 (*Arabidopsis thaliana* Glutathione S-transferase (class tau) 26); glutathione transferase	0,6891	2,34E-02
AT5G50260.1	cysteine proteinase, putative	-0,6077	2,34E-02
AT1G24430.1	transferase family protein	-0,6425	2,34E-02
AT5G10300.1	hydrolase, alpha/beta fold family protein	-0,5557	2,36E-02
AT5G42250.1	alcohol dehydrogenase, putative	0,6434	2,36E-02
AT1G22160.1	senescence-associated protein-related	0,4638	2,38E-02
AT1G17050.1	SPS2 (Solanesyl diphosphate synthase 2); dimethylallyltranstransferase	0,7170	2,40E-02
AT4G22280	F-box family protein	-0,5826	2,43E-02
AT1G76570.1	chlorophyll A-B binding family protein	0,6023	2,43E-02
AT1G25560.1	AP2 domain-containing transcription factor, putative	0,7528	2,44E-02
AT2G43660	glycosyl hydrolase family protein 17	-0,7441	2,45E-02
AT1G52510	hydrolase, alpha/beta fold family protein	0,8392	2,45E-02
AT5G28840.1	GME (GDP-D-MANNOSE 3',5'-EPIMERASE); GDP-mannose 3,5-epimerase/ NAD binding / catalytic	0,4630	2,47E-02
AT4G33020.1	ZIP9 (ZINC TRANSPORTER 9 PRECURSOR); cation transporter	0,5526	2,52E-02
AT3G52450.1	U-box domain-containing protein	-0,4935	2,54E-02
AT4G04570.1	protein kinase family protein	0,4220	2,55E-02
AT4G25820.1	XTR9 (XYLOGLUCAN ENDOTRANSGLYCOSYLASE 9); hydrolase, acting on glycosyl bonds	0,9503	2,58E-02
AT3G48420.1	haloacid dehalogenase-like hydrolase family protein	0,8529	2,60E-02
AT1G32870.1	ANAC013 (*Arabidopsis* NAC domain containing protein 13); transcription factor	-0,6526	2,61E-02
AT4G12830.1	hydrolase, alpha/beta fold family protein	0,9583	2,62E-02
AT5G07200.1	YAP169 (Gibberellin 20 oxidase 3); gibberellin 20-oxidase	-0,5593	2,64E-02
AT3G50030.1	binding	-0,6332	2,66E-02
AT4G02770.1	PSAD-1 (photosystem I subunit D-1)	1,1155	2,67E-02
AT4G22240.1	plastid-lipid associated protein PAP, putative	0,6016	2,69E-02
AT4G26320.1	AGP13 (ARABINOGALACTAN PROTEIN 13)	0,7892	2,69E-02
AT1G73330.1	ATDR4 (*Arabidopsis thaliana* drought-repressed 4)	0,6012	2,70E-02
AT3G63160.1	unknown protein	0,8796	2,70E-02
AT1G16370.1	transporter-related	0,9309	2,75E-02
AT3G62310.1	RNA helicase, putative	-0,5089	2,76E-02
AT1G01810.1	unknown protein	-0,4729	2,77E-02
AT3G19760.1	eukaryotic translation initiation factor 4A, putative / eIF-4A, putative / DEAD box RNA helicase, putative	-0,4705	2,77E-02
AT1G23550.1	SRO2 (SIMILAR TO RCD ONE 2); NAD+ ADP-ribosyltransferase	-1,0349	2,87E-02
AT5G46295.1	unknown protein	1,2839	2,88E-02
AT3G28500.1	60S acidic ribosomal protein P2 (RPP2C)	-0,4746	2,88E-02
AT5G23030.1	TET12 (TETRASPANIN12)	1,0379	2,89E-02
AT5G21930	PAA2 (P-type ATPase of *Arabidopsis* 2); ATPase, coupled to transmembrane movement of ions, phosphorylative mechanism / copper ion transporter	0,5599	2,89E-02

AT2G03710	SEP4 (SEPALLATA4); DNA binding / transcription factor	0,5988	2,90E-02
AT1G09340.1	catalytic/ coenzyme binding	1,0515	2,91E-02
AT5G11950	unknown protein	0,4492	2,93E-02
AT2G16570.1	ATASE (GLN PHOSPHORIBOSYL PYROPHOSPHATE AMIDOTRANSFERASE 1); amidophosphoribosyltransferase	-0,6325	2,95E-02
AT1G65960.1	GAD2 (GLUTAMATE DECARBOXYLASE 2); calmodulin binding	0,6678	2,96E-02
AT3G13270.1	unknown protein	0,8630	2,96E-02
AT1G11860	aminomethyltransferase, putative	0,5995	2,97E-02
AT1G56580.1	unknown protein	-0,4721	2,97E-02
AT1G15410.1	aspartate-glutamate racemase family	0,5572	3,07E-02
ATMG01320.1	unknown protein	-1,0821	3,08E-02
AT5G19600.1	SULTR3;5 (SULTR3;5); sulfate transporter	-0,8111	3,09E-02
AT1G30380.1	PSAK (PHOTOSYSTEM I SUBUNIT K)	0,9583	3,09E-02
AT2G35990.1	unknown protein	0,5982	3,09E-02
AT1G31190.1	inositol monophosphatase family protein	0,6201	3,09E-02
AT1G26800.1	zinc finger (C3HC4-type RING finger) family protein	1,6105	3,10E-02
AT4G19170.1	NCED4 (NINE-CIS-EPOXYCAROTENOID DIOXYGENASE 4)	1,3854	3,13E-02
AT4G38950.1	kinesin motor family protein	0,5595	3,14E-02
AT5G01210.1	transferase family protein	0,7568	3,15E-02
AT4G34560.1	unknown protein	0,5598	3,15E-02
AT4G33010.1	glycine dehydrogenase (decarboxylating), putative / glycine decarboxylase, putative / glycine cleavage system P-protein, putative	0,8546	3,15E-02
AT1G12320.1	unknown protein	-0,8627	3,16E-02
AT5G63750.1	IBR domain-containing protein	-0,6556	3,17E-02
AT5G60610.1	F-box family protein	-0,6785	3,19E-02
AT1G75160.1	unknown protein	0,6409	3,19E-02
AT2G44940.1	AP2 domain-containing transcription factor TINY, putative	0,4803	3,19E-02
AT4G18300.1	eIF4-gamma/eIF5/eIF2-epsilon domain-containing protein	-0,6044	3,21E-02
AT3G28210.1	PMZ; zinc ion binding	-0,7262	3,22E-02
AT2G22990	SNG1 (SINAPOYLGLUCOSE 1); serine carboxypeptidase	0,8195	3,23E-02
AT1G29920.1	CAB2 (Chlorophyll a/b-binding protein 2); chlorophyll binding	1,5971	3,28E-02
AT1G02750.1	unknown protein	-0,6991	3,29E-02
AT3G10340.1	phenylalanine ammonia-lyase, putative	0,5556	3,30E-02
AT5G40010.1	AATP1 (AAA-ATPASE 1); ATP binding / ATPase	-0,5432	3,32E-02
AT3G45070.1	sulfotransferase family protein	0,8712	3,34E-02
AT5G02020	unknown protein	1,0658	3,34E-02
AT5G46240.1	KAT1 (K+ ATPase 1); cyclic nucleotide binding / inward rectifier potassium channel	-0,4853	3,38E-02
AT1G67150	unknown protein	-0,7870	3,40E-02
AT3G20820.1	leucine-rich repeat family protein	0,5326	3,40E-02
AT1G47540	trypsin inhibitor, putative	-0,7855	3,40E-02
AT1G60870.1	MEE9 (maternal effect embryo arrest 9)	0,5822	3,46E-02
AT2G40180.1	protein phosphatase 2C, putative / PP2C, putative	-0,6353	3,47E-02
AT5G09660	PMDH2 (PEROXISOMAL NAD-MALATE DEHYDROGENASE 2); malate dehydrogenase	0,7552	3,50E-02
AT1G73650	unknown protein	0,7948	3,50E-02
AT3G27280	ATPHB4 (PROHIBITIN 4)	-0,4801	3,53E-02
AT2G19310.1	unknown protein	0,9746	3,55E-02
AT1G79360.1	transporter-related	-0,4371	3,58E-02

AT5G43620.1	S-locus protein-related	-0,8966	3,60E-02
AT5G36180.1	SCPL1 (serine carboxypeptidase-like 1); serine carboxypeptidase	0,9966	3,60E-02
AT2G41870.1	remorin family protein	0,7079	3,62E-02
AT2G19610.1	zinc finger (C3HC4-type RING finger) family protein	-0,7676	3,64E-02
AT5G50100.1	unknown protein	0,5993	3,64E-02
AT5G66700.1	HB53 (homeobox-8); DNA binding / transcription factor	0,5126	3,68E-02
AT5G49850.1	jacalin lectin family protein	-0,8077	3,82E-02
AT3G29370.1	unknown protein	0,7323	3,83E-02
AT5G04360.1	ATPU1 (PULLULANASE 1); alpha-amylase/ limit dextrinase	0,4353	3,86E-02
AT3G26510	octicosapeptide/Phox/Bem1p (PB1) domain-containing protein	0,7249	3,86E-02
AT1G74810.1	anion exchange family protein	1,4072	3,86E-02
AT1G12080	unknown protein	0,5040	3,87E-02
AT1G06160.1	ethylene-responsive factor, putative	0,5753	3,92E-02
AT5G02940.1	unknown protein	0,6178	3,93E-02
AT3G56080.1	dehydration-responsive protein-related	0,4266	3,95E-02
AT4G30670.1	unknown protein	0,5395	4,00E-02
AT1G05730.1	unknown protein	-0,6370	4,01E-02
AT3G47250	unknown protein	0,4851	4,01E-02
AT5G44260.1	zinc finger (CCCH-type) family protein	0,7750	4,03E-02
AT1G13420.1	sulfotransferase family protein	1,0018	4,05E-02
AT1G63750	ATP binding / nucleoside-triphosphatase/ nucleotide binding / protein binding	0,5744	4,11E-02
AT5G64000.1	SAL2; 3'(2'),5'-bisphosphate nucleotidase/ inositol or phosphatidylinositol phosphatase	0,9570	4,11E-02
AT1G29520.1	AWPM-19-like membrane family protein	0,4836	4,17E-02
AT1G70490	ADP-ribosylation factor, putative	-0,7103	4,20E-02
AT2G25700.1	ASK3 (*ARABIDOPSIS* SKP1-LIKE 3); protein binding / ubiquitin-protein ligase	-0,5697	4,20E-02
AT1G66270	beta-glucosidase (PSR3.2)	1,0212	4,23E-02
AT1G68080	oxidoreductase, acting on paired donors, with incorporation or reduction of molecular oxygen, 2-oxoglutarate as one donor, and incorporation of one atom each of oxygen into both donors	-0,4359	4,29E-02
AT5G57530.1	xyloglucan:xyloglucosyl transferase, putative / xyloglucan endotransglycosylase, putative / endo-xyloglucan transferase, putative	1,2444	4,31E-02
AT1G27890.1	CCR4-NOT transcription complex protein, putative	-0,6588	4,33E-02
AT1G58270.1	ZW9	0,7477	4,33E-02
AT5G20250	DIN10 (DARK INDUCIBLE 10); hydrolase, hydrolyzing O-glycosyl compounds	0,9506	4,35E-02
AT1G01140	CIPK9 (CBL-INTERACTING PROTEIN KINASE 9); kinase	0,5500	4,36E-02
AT1G70700	unknown protein	0,8496	4,38E-02
AT1G65260.1	PTAC4 (PLASTID TRANSCRIPTIONALLY ACTIVE4)	0,5375	4,39E-02
AT4G03280	PETC (PHOTOSYNTHETIC ELECTRON TRANSFER C)	0,9358	4,40E-02
AT2G02955.1	MEE12 (maternal effect embryo arrest 12); zinc ion binding	-0,5579	4,40E-02
AT1G04180.1	flavin-containing monooxygenase family protein / FMO family protein	-0,4806	4,43E-02
AT5G19120.1	pepsin A	1,4663	4,45E-02
AT3G61470.1	LHCA2 (Photosystem I light harvesting complex gene 2); chlorophyll binding	0,9337	4,50E-02
AT3G51540.1	unknown protein	0,4359	4,50E-02
AT1G67650.1	binding	-0,6691	4,55E-02

AT3G12340.1	FK506 binding / peptidyl-prolyl cis-trans isomerase	-0,4077	4,56E-02
AT1G22380.1	transcription factor/ transferase, transferring glycosyl groups	-0,9130	4,61E-02
AT1G02640.1	BXL2 (BETA-XYLOSIDASE 2); hydrolase, hydrolyzing O-glycosyl compounds	0,6779	4,61E-02
AT4G09350.1	DNAJ heat shock N-terminal domain-containing protein	0,9089	4,65E-02
AT3G43930	unknown protein	0,5300	4,67E-02
AT1G74110.1	CYP78A10 (cytochrome P450, family 78, subfamily A, polypeptide 10); oxygen binding	0,5036	4,67E-02
AT5G41610	ATCHX18 (cation/hydrogen exchanger 18); monovalent cation:proton antiporter	-0,6601	4,67E-02
AT1G03170.1	unknown protein	-0,7172	4,68E-02
AT3G54890	LHCA1; chlorophyll binding	0,9797	4,73E-02
AT5G48570.1	peptidyl-prolyl cis-trans isomerase, putative / FK506-binding protein, putative	-0,5845	4,75E-02
AT5G48175.1	unknown protein	0,8526	4,75E-02
AT3G19760.1	eukaryotic translation initiation factor 4A, putative / eIF-4A, putative / DEAD box RNA helicase, putative	-0,4499	4,76E-02
AT5G16320.1	FRL1 (FRIGIDA LIKE 1)	-0,4462	4,76E-02
AT5G27510.1	protein kinase family protein	-0,5102	4,78E-02
AT4G09030.1	AGP10 (Arabinogalactan protein 10)	-0,5804	4,81E-02
AT2G43560.1	immunophilin / FKBP-type peptidyl-prolyl cis-trans isomerase family protein	0,7764	4,81E-02
AT1G67000.1	kinase	-0,3930	4,82E-02
AT2G45750.1	dehydration-responsive family protein	0,8749	4,84E-02
AT1G10550.1	XTH33 (xyloglucan:xyloglucosyl transferase 33); hydrolase, acting on glycosyl bonds	0,8001	4,84E-02
AT5G41360.1	XPB2 (*ARABIDOPSIS* HOMOLOG OF XERODERMA PIGMENTOSUM COMPLEMENTATION GROUP B 2); ATP-dependent helicase	-0,4901	4,88E-02
AT3G50240.1	KICP-02; microtubule motor	0,4350	4,89E-02
AT3G45680.1	proton-dependent oligopeptide transport (POT) family protein	0,4417	4,98E-02
AT5G48657.1	defense protein-related	0,6663	4,99E-02
AT4G30650.1	hydrophobic protein, putative / low temperature and salt responsive protein, putative	0,7745	5,00E-02

App. Table 7 Enrichment/depletion of gene ontology (GO) functional groups among the differentially expressed genes in seedlings of tetraploid Col-0 plants.

GO	Description	GOcategory	FDR	Over/Under
GO:0009535	chloroplast thylakoid membrane	cellular component	8.43E-009	over
GO:0009579	thylakoid	cellular component	8.43E-009	over
GO:0044435	plastid part	cellular component	8.43E-009	over
GO:0042651	thylakoid membrane	cellular component	8.43E-009	over
GO:0016168	chlorophyll binding	molecular function	8.43E-009	over
GO:0055035	plastid thylakoid membrane	molecular function	8.43E-009	over
GO:0044434	chloroplast part	cellular component	8.43E-009	over
GO:0031976	plastid thylakoid	cellular	8.43E-009	over

		component		
GO:0044436	thylakoid part	cellular component	8.43E-009	over
GO:0031984	organelle subcompartment	cellular component	9.18E-009	over
GO:0009534	chloroplast thylakoid	cellular component	9.18E-009	over
GO:0015979	photosynthesis	biological process	9.64E-009	over
GO:0031090	organelle membrane	cellular component	9.64E-009	over
GO:0030076	light-harvesting complex	cellular component	8.28E-008	over
GO:0046906	tetrapyrolle binding	molecular function	3.10E-007	over
GO:0009536	plastid	cellular component	9.92E-006	over
GO:0009507	chloroplast	cellular component	2.50E-005	over
GO:0016491	oxidoreductase activity	molecular function	1.16E-004	over
GO:0019684	photosynthesis, light reaction	biological process	1.91E-004	over
GO:0044464	cell part	cellular component	2.25E-004	over
GO:0005623	cell	cellular component	2.25E-004	over
GO:0009521	photosystem	cellular component	2.65E-004	over
GO:0016020	membrane	cellular component	6.56E-004	over
GO:0010287	plastoglobule	cellular component	6.56E-004	over
GO:0009522	photosystem I	cellular component	0	over
GO:0009765	photosynthesis, light harvesting	biological process	0	over
GO:0010319	stromule	cellular component	0	over
GO:0009266	response to temperature stimulus	biological process	0	over
GO:0044446	intracellular organelle part	cellular component	0	over
GO:0044422	organelle part	cellular component	0	over
GO:0006091	generation of precursor metabolites and energy	biological process	0	over
GO:0005737	cytoplasm	cellular component	0	over
GO:0044425	membrane part	cellular component	0.01	over
GO:0006811	ion transport	biological process	0.01	over

GO:0003824	catalytic activity	molecular function	0.01	over
GO:0060090	molecular adaptor activity	molecular function	0.01	under
GO:0030533	triplet codon-amino acid adaptor activity	molecular function	0.01	under
GO:0006412	translation	biological process	0.01	under
GO:0006414	translational elongation	biological process	0.01	under
GO:0043283	biopolymer metabolic process	biological process	0.01	under
GO:0009987	cellular process	biological process	0.01	under
GO:0009059	macromolecule biosynthetic process	biological process	0.01	under
GO:0043284	biopolymer biosynthetic process	biological process	0.01	under
GO:0044260	cellular macromolecule metabolic process	biological process	0.01	under
GO:0044267	cellular protein metabolic process	biological process	0.01	under
GO:0019538	protein metabolic process	biological process	0.01	under
GO:0006259	DNA metabolic process	biological process	0.01	under
GO:0043170	macromolecule metabolic process	biological process	0.01	under
GO:0044237	cellular metabolic process	biological process	0.01	under
GO:0006139	nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	biological process	0.01	under
GO:0044238	primary metabolic process	biological process	0.01	under
GO:0007275	multicellular organismal development	biological process	0.01	under
GO:0048316	seed development	biological process	0.01	under
GO:0022621	shoot system development	biological process	0.01	under
GO:0000003	reproduction	biological process	0.01	under
GO:0048367	shoot development	biological process	0.01	under
GO:0003006	reproductive developmental process	biological process	0.01	under
GO:0048608	reproductive structure develpment	biological process	0.01	under
GO:0032501	multicellular organismal process	biological process	0.01	under
GO:0022414	reproductive process	biological process	0.01	under

GO:0009790	embryonic develoment ending in seed dormancy	biological process	0.01	under
GO:0005794	Golgi apparatus	cellular component	0.01	under
GO:0009790	embryonic development	biological process	0.01	under
GO:0048856	anatomical structure development	biological process	0.01	under
GO:0044249	cellular biosynthetic process	biological process	0.01	under
GO:0048827	phyllome development	biological process	0.01	under
GO:0043232	intracellular non-membrane bound organelle	cellular component	0.01	under
GO:0043228	non-membrane-bound organelle	cellular component	0.01	under
GO:0048366	leaf development	biological process	0.02	under
GO:0008909	isochorismate synthase activity	molecular function	0.02	over
GO:0050486	intramolecular transferase activity, transferring hydroxy groups	molecular function	0.02	over
GO:0019464	glycine decarboxylation via glycine cleavage system	biological process	0.02	over
GO:0003950	NAD+ ADP-ribosyltransferase activity	molecular function	0.02	over
GO:0006950	response to stress	biological process	0.02	over
GO:0016043	cellular component organization and biogenesis	cellular component	0.02	under
GO:0043231	intracellular membrane-bound organelle	cellular component	0.02	over
GO:0043227	membrane-bound organelle	cellular component	0.02	over
GO:0050896	response to stimulus	biological process	0.03	over
GO:0016773	phosphotransferase activity,alcohol group as acceptor	molecular function	0.03	under
GO:0048869	cellular developmental process	biological process	0.03	under
GO:0009063	amino acid catabolic process	biological process	0.03	over
GO:0009310	amine catabolic process	biological process	0.03	over
GO:0030529	ribonucleoprotein complex	cellular component	0.03	under
GO:0004674	protein serine/threonine kinase activity	molecular function	0.03	under
GO:0006393	RNA processing	biological process	0.03	under
GO:0006974	response to DNA damage stimulus	biological process	0.03	under

GO:0006996	organelle organization and biogenesis	biological process	0.03	under
GO:0016638	oxidoreductase activity, acting on the CH-NH2 group of donors	molecular function	0.03	over
GO:0005527	macrolide binding	molecular function	0.03	over
GO:0005528	FK506 binding	molecular function	0.03	over
GO:0009791	post-embryonic development	biological process	0.03	under
GO:0010016	shoot morphogenesis	biological process	0.03	under
GO:0004091	carboxylesterase activity	molecular function	0.03	under
GO:0040029	regulation of gene expression,epigenetic	molecular function	0.04	under
GO:0006281	DNA repair	biological process	0.04	under
GO:0010076	maintenance of floral meristem identity	biological process	0.04	over
GO:0044270	nitrogen compound catabolic process	biological process	0.04	over
GO:0009628	response to abiotic stimulus	biological process	0.04	over
GO:0022890	inorganic cation transmembrane transporter activity	molecular function	0.04	over
GO:0007568	aging	biological process	0.04	over
GO:0005622	intracellular	cellular component	0.04	over
GO:0008144	drug binding	molecular function	0.04	over
GO:0004601	peroxidase activity	molecular function	0.04	over
GO:0016684	oxidoreductase activity, acting on peroxide as acceptor	molecular function	0.04	over
GO:0032502	developmental process	biological process	0.04	under
GO:0030154	cell differentiation	biological process	0.04	under

App. Table 8. Gene expression differences between consecutive (2nd & 3rd) generations of Col-0 tetraploids (4 arrays; Col-0 lines 1326-12,-19,-26 and -28)

Gene ID	Description	log_2FC	P-value
AT3G54890	LHCA1		0.00878
AT4G25100	FSD1 (FE SUPEROXIDE DISMUTASE 1); iron superoxide dismutase		0.03268
AT5G54270.1	LHCB3 (LIGHT-HARVESTING CHLOROPHYLL BINDING PROTEIN 3)		0.01743
AT5G56080.1	nicotianamine synthase, putative		0.02246

App. Table 9. Differentially expressed genes in leaves - diploid vs. tetraploid Col-0 (12 biological replicates; 4 arrays with Col-0 line 1326-12,-19,-26,-28; 4 arrays each with individual plants of line 1326-26 and 1326-28 respectively)

Gene ID	Description	log_2 FC	P-value
AT1G53480.1	unknown protein	4,7144	0,00E+00
AT5G66580.1	unknown protein	2,3027	3,93E-06
AT1G53490.1	DNA binding	1,0068	1,11E-05
AT2G38120.1	AUX1 (AUXIN RESISTANT 1); amino acid permease/ transporter	1,5212	1,60E-05
AT4G32280.1	*IAA*29 (indoleacetic acid-induced protein 29); transcription factor	3,4778	1,80E-05
AT5G44210.1	ATERF-9/ATERF9/ERF9 (ERF domain protein 9); DNA binding / transcription factor/ transcriptional repressor	1,8177	2,02E-05
AT1G09250.1	transcription factor	1,4376	2,02E-05
AT5G39860.1	PRE1 (PACLOBUTRAZOL RESISTANCE1); DNA binding / transcription factor	2,7093	2,67E-05
AT5G48900.1	pectate lyase family protein	2,2687	7,07E-05
AT1G04180.1	flavin-containing monooxygenase family protein / FMO family protein	2,1378	8,44E-05
AT1G15580.1	*IAA*5 (indoleacetic acid-induced protein 5); transcription factor	2,1170	9,29E-05
AT5G43810.1	ZLL (ZWILLE)	0,8701	9,94E-05
AT5G64770.1	unknown protein	1,3964	1,01E-04
AT1G06080.1	ADS1 (DELTA 9 DESATURASE 1); oxidoreductase	4,1604	1,14E-04
AT2G40610.1	ATEXPA8 (*ARABIDOPSIS THALIANA* EXPANSIN A8)	3,3388	1,15E-04
AT5G18010.1	auxin-responsive protein, putative	2,6351	1,45E-04
AT1G51820.1	leucine-rich repeat protein kinase, putative	-2,7394	1,51E-04
AT2G32870.1	meprin and TRAF homology domain-containing protein / MATH domain-containing protein	1,5303	1,78E-04
AT5G18050.1	auxin-responsive protein, putative	2,6458	1,92E-04
AT1G26945.1	transcription regulator	2,3806	2,05E-04
AT3G62100.1	*IAA*30 (indoleacetic acid-induced protein 30); transcription factor	1,0521	2,76E-04
AT1G02350.1	protoporphyrinogen oxidase-related	1,8412	2,76E-04
AT5G49100.1	unknown protein	1,4588	2,87E-04
AT4G19810.1	glycosyl hydrolase family 18 protein	-0,8537	2,88E-04
AT5G18080.1	auxin-responsive protein, putative	1,9974	3,08E-04
AT3G21330.1	basic helix-loop-helix (bHLH) family protein	2,1218	3,25E-04
AT5G03180.1	zinc finger (C3HC4-type RING finger) family protein	0,9407	4,67E-04
AT5G50335.1	unknown protein	2,1001	5,06E-04
AT5G02540.1	short-chain dehydrogenase/reductase (SDR) family protein	2,6771	5,58E-04
AT5G66590.1	allergen V5/Tpx-1-related family protein	2,2461	5,81E-04
AT4G28720.1	flavin-containing monooxygenase family protein / FMO family protein	1,3185	6,13E-04
AT2G31190.1	unknown protein	0,7642	6,34E-04
AT3G26445.1	glycosyl transferase family 17 protein	0,8858	6,50E-04
AT4G18270.1	ATTRANS11 (*Arabidopsis thaliana* translocase 11); catalytic	1,3791	6,79E-04
AT4G22190.1	unknown protein	0,8118	8,24E-04
AT5G46500.1	unknown protein	-1,1751	8,30E-04
AT3G28420.1	unknown protein	1,0497	1,02E-03
AT1G17140	tropomyosin-related	1,2648	1,13E-03
AT2G07708.1	unknown protein	-0,6220	1,20E-03
AT3G31900.1	unknown protein	-0,6180	1,21E-03
AT4G13260.1	YUC2 (YUCCA2); monooxygenase/ oxidoreductase	1,1334	1,22E-03
AT1G79910	unknown protein	0,6994	1,30E-03

AT3G50340.1	unknown protein	1,2611	1,31E-03
AT4G04340	early-responsive to dehydration protein-related / ERD protein-related	1,0816	1,35E-03
AT5G48530.1	unknown protein	-0,8495	1,36E-03
AT5G50090	unknown protein	0,5491	1,37E-03
AT4G22760.1	pentatricopeptide (PPR) repeat-containing protein	0,7588	1,42E-03
AT3G46110	signal transducer	-1,2938	1,44E-03
AT1G79750.1	ATNADP-ME4 (NADP-MALIC ENZYME 4); malate dehydrogenase (oxaloacetate-decarboxylating) (NADP+)/ malic enzyme/ oxidoreductase, acting on NADH or NADPH, NAD or NADP as acceptor	-0,4918	1,45E-03
AT1G76610.1	unknown protein	1,6718	1,46E-03
AT4G18530.1	unknown protein	0,6942	1,66E-03
AT3G08840.1	D-alanine--D-alanine ligase family	0,5710	1,72E-03
AT5G44260.1	zinc finger (CCCH-type) family protein	2,3786	1,82E-03
AT1G79040.1	PSBR (photosystem II subunit R)	1,1870	2,01E-03
AT5G12050.1	unknown protein	2,4506	2,14E-03
AT5G48480.1	unknown protein	0,7080	2,17E-03
AT1G05970	nucleotide binding	0,5957	2,17E-03
AT1G23080	PIN7 (PIN-FORMED 7); auxin:hydrogen symporter/ transporter	1,4948	2,22E-03
AT2G28870.1	unknown protein	0,8405	2,26E-03
AT4G30100.1	tRNA-splicing endonuclease positive effector-related	-0,5610	2,31E-03
AT4G30180.1	transcription factor/ transcription regulator	2,1890	2,35E-03
AT3G26750.1	unknown protein	0,6887	2,37E-03
AT3G16370.1	GDSL-motif lipase/hydrolase family protein	1,5031	2,39E-03
AT2G41010.1	ATCAMBP25 (*ARABIDOPSIS THALIANA* CALMODULIN (CAM)-BINDING PROTEIN OF 25 KDA); calmodulin binding	-1,1600	2,46E-03
AT2G21830.1	DC1 domain-containing protein	1,0524	2,49E-03
AT1G78170.1	unknown protein	1,0577	2,53E-03
AT1G50280.1	phototropic-responsive NPH3 family protein	1,2820	2,54E-03
ATMG01010	unknown protein	-0,5647	2,59E-03
AT2G47560.1	zinc finger (C3HC4-type RING finger) family protein	1,0397	2,71E-03
AT2G36980.1	pentatricopeptide (PPR) repeat-containing protein	0,6058	2,72E-03
AT4G00220.1	LOB domain protein 30 / lateral organ boundaries domain protein 30 (LBD30)	0,6287	2,78E-03
AT5G59350.1	unknown protein	1,1365	2,93E-03
AT2G37640.1	ATEXPA3 (*ARABIDOPSIS THALIANA* EXPANSIN A3)	1,5698	2,96E-03
AT5G25980	TGG2 (GLUCOSIDE GLUCOHYDROLASE 2); hydrolase, hydrolyzing O-glycosyl compounds	-0,8238	3,14E-03
AT4G14550.1	*IAA*14 (SOLITARY ROOT); transcription factor	1,9129	3,18E-03
AT1G67480	kelch repeat-containing F-box family protein	0,5491	3,25E-03
AT1G01725.1	unknown protein	0,6727	3,65E-03
AT3G15356.1	legume lectin family protein	-2,2357	3,65E-03
AT1G65920.1	regulator of chromosome condensation (RCC1) family protein / zinc finger protein-related	1,2231	3,66E-03
AT3G26520.1	TIP2 (TONOPLAST INTRINSIC PROTEIN 2); water channel	1,3525	3,71E-03
AT3G55760	unknown protein	0,7136	3,79E-03
AT5G01340.1	mitochondrial substrate carrier family protein	-1,0252	3,93E-03
AT4G12750.1	DNA binding / sequence-specific DNA binding / transcription factor	0,5941	4,17E-03
AT2G35700.1	AP2 domain-containing transcription factor, putative	0,7489	4,18E-03
AT1G54200.1	unknown protein	1,0231	4,41E-03
AT3G59430	unknown protein	0,5231	4,43E-03

AT3G26580.1	binding	0,8636	4,79E-03
AT5G28300.1	trihelix DNA-binding protein, putative	0,9067	4,88E-03
AT3G54860.1	ATVPS33 (*Arabidopsis thaliana* vacuolar protein sorting 33); protein transporter	0,5448	4,98E-03
AT2G07708	unknown protein	-0,5612	4,99E-03
AT3G63450	RNA binding	1,6617	5,01E-03
AT1G68560.1	ATXYL1/XYL1 (ALPHA-XYLOSIDASE 1); alpha-N-arabinofuranosidase/ hydrolase, hydrolyzing O-glycosyl compounds / xylan 1,4-beta-xylosidase	1,2641	5,15E-03
AT2G40640	unknown protein	0,7767	5,28E-03
AT3G48710.1	GTP binding / RNA binding	0,5924	5,31E-03
AT3G16510.1	C2 domain-containing protein	-1,1964	5,39E-03
AT5G42540.1	XRN2 (EXORIBONUCLEASE 2); 5'-3' exonuclease/ nucleic acid binding	-0,6753	5,46E-03
AT5G18030.1	auxin-responsive protein, putative	2,1813	5,58E-03
AT1G35350.1	unknown protein	-1,4047	5,71E-03
AT2G37100.1	protamine P1 family protein	0,6305	5,72E-03
AT1G03270.1	unknown protein	0,5455	5,81E-03
AT4G40060.1	ATHB16 (*ARABIDOPSIS THALIANA* HOMEOBOX PROTEIN 13); transcription factor	1,0064	6,06E-03
AT3G23550.1	MATE efflux family protein	-2,6662	6,21E-03
AT1G08090.1	ATNRT2:1 (*Arabidopsis thaliana* high affinity nitrate transporter 2.1); nitrate transporter	-1,4876	6,27E-03
AT5G42030.1	ABIL4 (ABL INTERACTOR-LIKE PROTEIN 4)	0,7200	6,30E-03
AT5G08760.1	unknown protein	-1,2415	6,73E-03
AT1G33760.1	AP2 domain-containing transcription factor, putative	-1,6951	6,88E-03
AT3G53210.1	nodulin MtN21 family protein	1,4675	6,92E-03
AT4G14510.1	unknown protein	0,6359	6,94E-03
AT1G20870.1	unknown protein	0,9167	7,07E-03
AT5G11550.1	binding	1,5969	7,21E-03
AT3G23030.1	*IAA*2 (indoleacetic acid-induced protein 2); transcription factor	0,9833	7,34E-03
AT3G16370.1	GDSL-motif lipase/hydrolase family protein	0,9240	7,37E-03
AT5G57887.1	unknown protein	1,1876	7,41E-03
AT1G70550	unknown protein	0,7593	7,45E-03
AT1G11220.1	unknown protein	0,9639	7,95E-03
AT5G02760.1	protein phosphatase 2C family protein / PP2C family protein	3,1337	8,54E-03
AT3G14210.1	ESM1 (EPITHIOSPECIFIER MODIFIER 1); carboxylic ester hydrolase	1,3426	8,60E-03
AT5G18020.1	auxin-responsive protein, putative	1,8587	8,63E-03
AT5G17890.1	LIM domain-containing protein / disease resistance protein-related	-0,7257	8,76E-03
AT1G02160	unknown protein	0,7635	9,23E-03
AT4G30400.1	zinc finger (C3HC4-type RING finger) family protein	0,6059	9,24E-03
AT1G22400.1	UGT85A1 (UDP-glucosyl transferase 85A1); UDP-glycosyltransferase/ transferase, transferring glycosyl groups / transferase, transferring hexosyl groups	-2,1516	9,31E-03
AT5G55610	unknown protein	-0,6801	9,50E-03
AT4G25410.1	basix helix-loop-helix (bHLH) family protein	0,7462	9,65E-03
AT5G54530.1	unknown protein	0,9342	9,67E-03
AT1G06350.1	fatty acid desaturase family protein	1,1437	9,87E-03
AT1G09710.1	DNA binding	0,8371	1,04E-02
AT1G03370.1	C2 domain-containing protein / GRAM domain-containing protein	-0,8682	1,04E-02
AT5G06460.1	ATUBA2 (*Arabidopsis thaliana* ubiquitin activating enzyme	0,4260	1,09E-02

	2); ubiquitin activating enzyme		
AT4G38750.1	unknown protein	0,4722	1,11E-02
AT4G13540.1	unknown protein	1,0834	1,13E-02
AT1G78480.1	prenyltransferase/squalene oxidase repeat-containing protein	0,9088	1,14E-02
AT4G21215	unknown protein	1,2692	1,16E-02
AT1G29460.1	auxin-responsive protein, putative	2,2818	1,18E-02
AT3G16370.1	GDSL-motif lipase/hydrolase family protein	1,2385	1,21E-02
AT2G28720.1	histone H2B, putative	1,1909	1,23E-02
AT2G27960.1	CKS1 (CDK-SUBUNIT 1); cyclin-dependent protein kinase	0,5682	1,23E-02
AT4G34980.1	SLP2 (subtilisin-like serine protease 2); subtilase	0,9098	1,25E-02
AT3G16800	protein phosphatase 2C, putative / PP2C, putative	1,5108	1,26E-02
AT2G22122.1	unknown protein	1,7152	1,30E-02
AT4G11830	phospholipase D gamma 2 / PLD gamma 2 (PLDGAMMA2)	-0,7214	1,31E-02
AT5G64290.1	DCT/DIT2.1 (DICARBOXYLATE TRANSPORT); oxoglutarate:malate antiporter	-0,6672	1,32E-02
AT1G49560.1	myb family transcription factor	0,9982	1,34E-02
AT3G07980.1	MAPKKK6 (MAP3K EPSILON PROTEIN KINASE 2); kinase	-0,5023	1,34E-02
AT3G26050.1	unknown protein	0,7431	1,35E-02
AT1G13930.1	unknown protein	1,2064	1,35E-02
AT1G50560.1	CYP705A25 (cytochrome P450, family 705, subfamily A, polypeptide 25); oxygen binding	-0,5179	1,37E-02
AT5G10840.1	endomembrane protein 70, putative	-0,7425	1,40E-02
AT1G05460.1	SDE3 (SILENCING DEFECTIVE)	0,6734	1,41E-02
AT1G63930.1	unknown protein	-1,0415	1,42E-02
AT4G30270.1	MERI5B (MERISTEM-5); hydrolase, acting on glycosyl bonds	-2,2341	1,43E-02
AT2G02760.1	ATUBC2 (UBIQUITING-CONJUGATING ENZYME 2); ubiquitin-protein ligase	0,6014	1,47E-02
AT3G14170.1	unknown protein	0,7299	1,49E-02
AT4G02200	drought-responsive family protein	-0,8093	1,52E-02
AT5G60850.1	OBP4 (OBF BINDING PROTEIN 4); DNA binding / transcription factor	1,0837	1,53E-02
AT1G32090.1	early-responsive to dehydration protein-related / ERD protein-related	1,1922	1,60E-02
AT2G46530	ARF11 (AUXIN RESPONSE FACTOR 11); transcription factor	1,1489	1,60E-02
AT3G09800	clathrin adaptor complex small chain family protein	0,7169	1,62E-02
AT4G24790.1	ATP binding / DNA-directed DNA polymerase	0,8997	1,62E-02
AT2G43290.1	MSS3 (MULTICOPY SUPPRESSORS OF SNF4 DEFICIENCY IN YEAST 3); calcium ion binding	-0,7580	1,65E-02
AT4G26750.1	hydroxyproline-rich glycoprotein family protein	-0,7069	1,70E-02
AT2G01150.1	RHA2B (RING-H2 FINGER PROTEIN 2B); protein binding / zinc ion binding	1,2032	1,70E-02
AT3G62090	PIL2 (PHYTOCHROME INTERACTING FACTOR 3-LIKE 2); transcription factor	1,5723	1,71E-02
AT3G49900.1	BTB/POZ domain-containing protein	0,6094	1,71E-02
AT2G45460.1	forkhead-associated domain-containing protein / FHA domain-containing protein	0,6417	1,74E-02
AT1G77122.1	unknown protein	0,5505	1,78E-02
AT3G63445.2	unknown protein	1,0228	1,80E-02
AT1G21680.1	unknown protein	0,9423	1,82E-02
AT1G29395.1	COR414-TM1 (cold regulated 414 thylakoid membrane 1)	1,7230	1,83E-02
AT5G62170.1	unknown protein	1,0179	1,84E-02
AT1G29500.1	auxin-responsive protein, putative	1,7362	1,85E-02

AT3G50630.1	ICK2 (KIP-RELATED PROTEIN 2)	0,8813	1,89E-02
AT3G50890.1	ATHB28 (*ARABIDOPSIS THALIANA* HOMEOBOX PROTEIN 28); DNA binding / transcription factor	1,2029	1,95E-02
AT4G01510.1	ARV2	0,5850	1,98E-02
AT2G28050.1	pentatricopeptide (PPR) repeat-containing protein	0,6708	1,98E-02
AT5G64320.1	pentatricopeptide (PPR) repeat-containing protein	0,4252	2,00E-02
AT2G05830	eukaryotic translation initiation factor 2B family protein / eIF-2B family protein	0,5506	2,00E-02
AT3G10660.1	CPK2 (CALMODULIN-DOMAIN PROTEIN KINASE CDPK ISOFORM 2); calcium- and calmodulin-dependent protein kinase/ kinase	-1,1076	2,02E-02
AT2G26710.1	BAS1 (PHYB ACTIVATION TAGGED SUPPRESSOR 1); oxygen binding	1,6105	2,04E-02
AT1G13910.1	leucine-rich repeat family protein	0,7508	2,07E-02
AT4G15053.1	unknown protein	-0,5832	2,07E-02
AT2G36490.1	ROS1 (repressor of silencing1)	0,6671	2,09E-02
AT1G69690.1	TCP family transcription factor, putative	0,6259	2,11E-02
AT2G21770.1	CESA9 (CELLULASE SYNTHASE 9); transferase, transferring glycosyl groups	-0,4273	2,11E-02
AT1G24625.1	ZFP7 (ZINC FINGER PROTEIN 7); nucleic acid binding / transcription factor/ zinc ion binding	0,7480	2,12E-02
AT5G66760.1	SDH1-1 (Succinate dehydrogenase 1-1)	-0,6416	2,13E-02
AT5G46470.1	disease resistance protein (TIR-NBS-LRR class), putative	-0,7053	2,15E-02
AT1G53490.1	DNA binding	0,8615	2,16E-02
AT3G05320.1	unknown protein	-1,1453	2,17E-02
AT3G27420.1	unknown protein	0,5780	2,21E-02
AT3G58560.1	endonuclease/exonuclease/phosphatase family protein	-0,5276	2,28E-02
AT2G31600	unknown protein	0,5785	2,30E-02
AT3G18080.1	glycosyl hydrolase family 1 protein	1,3754	2,30E-02
AT3G20820.1	leucine-rich repeat family protein	1,3647	2,30E-02
AT5G59550.1	zinc finger (C3HC4-type RING finger) family protein	-0,8542	2,34E-02
AT1G29720.1	protein kinase family protein	-0,9567	2,35E-02
AT2G38050.1	DET2 (DE-ETIOLATED 2)	0,8069	2,37E-02
AT3G63445	unknown protein	0,5690	2,38E-02
AT2G19570.1	CDA1 (CYTIDINE DEAMINASE 1)	-0,6122	2,40E-02
AT3G11650.1	NHL2 (NDR1/HIN1-like 2)	-0,9434	2,41E-02
AT1G14670.1	endomembrane protein 70, putative	-0,6318	2,44E-02
AT1G05310.1	pectinesterase family protein	0,9490	2,51E-02
AT4G25270.1	pentatricopeptide (PPR) repeat-containing protein	0,5643	2,54E-02
AT3G06370.1	NHX4 (sodium proton exchanger 4); sodium:hydrogen antiporter	0,6867	2,59E-02
AT2G31560	unknown protein	0,8317	2,62E-02
AT1G76370.1	protein kinase, putative	-0,5007	2,62E-02
AT5G23410.1	unknown protein	0,8505	2,63E-02
AT3G62650	binding	1,1116	2,68E-02
AT3G56040.1	unknown protein	0,6456	2,68E-02
AT4G15140.1	unknown protein	0,5453	2,74E-02
AT5G27030.1	TPR3 (TOPLESS-RELATED 3)	-0,7469	2,74E-02
AT5G48530.1	unknown protein	-0,7603	2,75E-02
AT1G70740.1	protein kinase family protein	-1,0150	2,75E-02
AT3G15540.1	*IAA*19 (indoleacetic acid-induced protein 19); transcription factor	1,2034	2,77E-02
AT4G39510.1	CYP96A12 (cytochrome P450, family 96, subfamily A, polypeptide 12); oxygen binding	1,7679	2,77E-02

AT4G16515.1	unknown protein	1,4376	2,78E-02
AT5G53080.1	kinesin light chain-related	0,7276	2,79E-02
AT4G01130.1	acetylesterase, putative	1,5852	2,81E-02
AT1G68370.1	ARG1 (ALTERED RESPONSE TO GRAVITY 1)	-0,4778	2,83E-02
AT5G08130.1	BIM1 (BES1-interacting Myc-like protein 1); DNA binding / transcription factor	1,1888	2,83E-02
AT3G17020.1	universal stress protein (USP) family protein	0,7653	2,85E-02
AT1G74670.1	gibberellin-responsive protein, putative	3,6366	2,86E-02
AT3G51950	zinc finger (CCCH-type) family protein / RNA recognition motif (RRM)-containing protein	0,8256	2,88E-02
AT1G01880.1	DNA repair protein, putative	0,8529	2,88E-02
AT4G33460.1	ATNAP13 (EMBRYO DEFECTIVE 2751)	0,4432	2,89E-02
AT3G54740	unknown protein	0,7881	2,94E-02
AT1G33670.1	leucine-rich repeat family protein	-0,4894	2,96E-02
AT1G40133.1	unknown protein	-0,6421	2,98E-02
AT1G09350.1	ATGOLS3 (*ARABIDOPSIS THALIANA* GALACTINOL SYNTHASE 3); transferase, transferring glycosyl groups / transferase, transferring hexosyl groups	2,3737	3,12E-02
AT3G19150.1	ICK4 (KIP-RELATED PROTEIN 6)	-0,6836	3,19E-02
AT3G22820.1	allergen-related	0,9412	3,22E-02
AT5G13050	5-FCL; 5-formyltetrahydrofolate cyclo-ligase	0,5490	3,28E-02
AT5G51890.1	peroxidase	0,6801	3,28E-02
AT4G16800.1	enoyl-CoA hydratase, putative	0,4706	3,30E-02
AT3G06950.1	tRNA pseudouridine synthase family protein	0,6801	3,41E-02
AT3G22930.1	calmodulin, putative	-1,5784	3,42E-02
AT1G15215	unknown protein	0,6455	3,51E-02
AT5G15640.1	mitochondrial substrate carrier family protein	-0,6557	3,52E-02
AT2G17540	unknown protein	0,7425	3,55E-02
AT5G13290	protein kinase family protein	0,5278	3,57E-02
AT5G52550.1	unknown protein	0,6125	3,60E-02
AT5G02890.1	transferase family protein	1,1110	3,60E-02
AT2G20230.1	unknown protein	0,7428	3,61E-02
AT3G11530	vacuolar protein sorting 55 family protein / VPS55 family protein	0,4243	3,65E-02
AT1G65320.1	CBS domain-containing protein	0,6329	3,72E-02
AT2G33860.1	ETT (ETTIN); transcription factor	0,8932	3,79E-02
AT1G21560.1	unknown protein	0,5482	3,87E-02
AT5G54100.1	band 7 family protein	-0,8658	3,89E-02
AT3G55240.1	unknown protein	2,6108	3,97E-02
AT3G23640.1	HGL1 (HETEROGLYCAN GLUCOSIDASE 1); hydrolase, hydrolyzing O-glycosyl compounds	0,7930	3,99E-02
AT3G02630.1	acyl-(acyl-carrier-protein) desaturase, putative / stearoyl-ACP desaturase, putative	0,5444	4,06E-02
AT1G23340	unknown protein	0,5954	4,08E-02
AT3G47030.1	F-box family protein	0,6983	4,08E-02
AT3G01490.1	protein kinase, putative	1,5408	4,09E-02
AT2G23810.1	TET8 (TETRASPANIN8)	-1,3754	4,10E-02
AT5G02080.1	DNA/panthothenate metabolism flavoprotein family protein	0,6814	4,10E-02
AT1G48330.1	unknown protein	1,5085	4,12E-02
AT3G03170.1	unknown protein	0,9552	4,12E-02
AT1G76680	OPR1 (12-oxophytodienoate reductase 1); 12-oxophytodienoate reductase	-1,2419	4,13E-02
AT2G41110.1	ATCAL4 (calmodulin-related protein 4); calcium ion binding	-1,2145	4,18E-02
AT1G28470.1	ANAC010 (*Arabidopsis* NAC domain containing protein 10);	0,5889	4,23E-02

	transcription factor		
AT1G10970.1	ZIP4 (ZINC TRANSPORTER 4 PRECURSOR); cation transporter	1,0677	4,28E-02
AT4G32270	unknown protein	0,6345	4,29E-02
AT5G38860.1	BIM3 (BES1-INTERACTING MYC-LIKE PROTEIN 3); DNA binding / transcription factor	0,7651	4,35E-02
AT2G35155.1	unknown protein	0,5436	4,39E-02
AT3G22235.1	unknown protein	-1,8138	4,41E-02
AT2G45520.1	unknown protein	0,4440	4,41E-02
AT3G45780	PHOT1 (phototropin 1); kinase	1,0237	4,42E-02
AT4G38490.1	unknown protein	0,6118	4,44E-02
AT2G23760	BLH4 (BLH4); DNA binding / transcription factor	0,5884	4,46E-02
AT1G54850.1	unknown protein	0,4928	4,50E-02
AT5G48920.1	hydroxyproline-rich glycoprotein family protein	0,9290	4,51E-02
AT3G29320.1	glucan phosphorylase, putative	1,1446	4,55E-02
AT5G65310	ATHB5 (*ARABIDOPSIS THALIANA* HOMEOBOX PROTEIN 5); transcription factor	1,3658	4,58E-02
AT1G66260.1	RNA and export factor-binding protein, putative	-0,6063	4,60E-02
AT3G15980	coatomer protein complex, subunit beta 2 (beta prime), putative	-0,5419	4,63E-02
AT4G03080.1	kelch repeat-containing serine/threonine phosphoesterase family protein	-0,6005	4,65E-02
AT1G71697.1	ATCK1 (CHOLINE KINASE)	-0,8598	4,73E-02
AT3G48290.1	CYP71A24 (cytochrome P450, family 71, subfamily A, polypeptide 24); oxygen binding	1,2382	4,76E-02
AT3G43850.1	unknown protein	0,4917	4,79E-02
AT5G38070.1	zinc finger (C3HC4-type RING finger) family protein	0,5328	4,81E-02
AT4G14560.1	*IAA*1 (INDOLE-3-ACETIC ACID INDUCIBLE); transcription factor	0,9829	4,82E-02
AT1G19340.1	methyltransferase MT-A70 family protein	0,6490	4,82E-02
AT4G08500.1	MEKK1 (MYTOGEN ACTIVATED PROTEIN KINASE KINASE); kinase	-0,8071	4,84E-02
AT4G25480.1	DREB1A (DEHYDRATION RESPONSE ELEMENT B1A); DNA binding / transcription factor/ transcriptional activator	1,9775	4,88E-02
AT4G22756.1	SMO1-2 (STEROL C4-METHYL OXIDASE); catalytic	0,6722	4,90E-02
AT4G30650.1	hydrophobic protein, putative / low temperature and salt responsive protein, putative	1,5126	4,93E-02
AT3G54260.1	unknown protein	0,6760	4,94E-02
AT1G30160	unknown protein	0,8062	4,94E-02

App. Table 10. Differentially expressed genes in leaves - diploid Col-0 vs.tetraploid Col-0 (four biological replicates: four arrays including tetraploid Col-0 lines 1326-12, -19, -26 and -28)

Gene ID	Description	log_2 FC	P-value
AT3G30775.1	ERD5 (EARLY RESPONSIVE TO DEHYDRATION 5, PROLINE OXIDASE); proline dehydrogenase	-4,85E+04	1,40E-04
AT2G40330.1	Bet v I allergen family protein	-5,78E+04	1,94E-04
AT2G02990.1	RNS1 (RIBONUCLEASE 1); endoribonuclease	4,71E+04	3,00E-04
AT3G29575	unknown protein	4,66E+04	7,28E-04
AT1G53480.1	unknown protein	4,66E+04	1,21E-03
AT5G41080	glycerophosphoryl diester phosphodiesterase family protein	-4,60E+04	1,52E-03
AT1G11210.1	unknown protein	3,40E+04	3,29E-03
AT3G21330.1	basic helix-loop-helix (bHLH) family protein	2,88E+04	4,62E-03
AT1G75450.1	CKX5 (CYTOKININ OXIDASE 5); cytokinin	4,06E+04	4,93E-03

	dehydrogenase		
AT5G54585.1	unknown protein	3,43E+04	5,32E-03
AT1G52890.1	ANAC019 (*Arabidopsis* NAC domain containing protein 19); transcription factor	3,59E+04	5,56E-03
AT5G53710.1	unknown protein	3,82E+04	5,59E-03
AT2G46970.1	PIL1 (PHYTOCHROME INTERACTING FACTOR 3-LIKE 1); transcription factor	4,31E+04	6,55E-03
AT3G55500.1	ATEXPA16 (*ARABIDOPSIS THALIANA* EXPANSIN A16)	3,21E+04	6,60E-03
AT3G02410.1	serine esterase	-4,40E+04	8,58E-03
AT4G09500	glycosyltransferase family protein	2,46E+04	9,83E-03
AT3G05640	protein phosphatase 2C, putative / PP2C, putative	3,38E+04	1,17E-02
AT5G44210.1	ATERF-9/ATERF9/ERF9 (ERF domain protein 9); DNA binding / transcription factor/ transcriptional repressor	2,34E+04	1,24E-02
AT4G21870.1	26.5 kDa class P-related heat shock protein (HSP26.5-P)	-4,63E+04	1,27E-02
AT1G53700.1	WAG1; kinase	2,32E+04	1,32E-02
AT1G79520.1	cation efflux family protein	2,50E+04	1,46E-02
AT1G30720.1	FAD-binding domain-containing protein	-2,74E+04	1,58E-02
AT2G34510.1	unknown protein	-2,52E+04	1,60E-02
AT1G20780.1	armadillo/beta-catenin repeat protein-related / U-box domain-containing protein	-2,03E+04	1,66E-02
AT3G26450.1	major latex protein-related / MLP-related	2,21E+04	1,74E-02
AT1G30730.1	FAD-binding domain-containing protein	-2,66E+04	1,78E-02
AT1G56600.1	ATGOLS2 (*ARABIDOPSIS THALIANA* GALACTINOL SYNTHASE 2); transferase, transferring glycosyl groups / transferase, transferring hexosyl groups	5,92E+04	1,81E-02
AT3G48360.1	BT2 (BTB and TAZ domain protein 2); protein binding / transcription regulator	-5,14E+04	2,08E-02
AT1G52855.1	unknown protein	2,77E+04	2,23E-02
AT1G51800.1	leucine-rich repeat protein kinase, putative	-3,08E+04	2,25E-02
AT4G39070.1	zinc finger (B-box type) family protein	-3,45E+04	2,28E-02
AT3G57930	unknown protein	-1,98E+04	2,39E-02
AT3G02020.1	aspartate kinase, lysine-sensitive, putative	2,22E+04	2,39E-02
AT2G37760	aldo/keto reductase family protein	2,21E+04	2,42E-02
AT4G21440.1	ATM4/ATMYB102 (*ARABIDOPSIS* MYB-LIKE 102); DNA binding / transcription factor	2,39E+04	2,81E-02
AT1G23200.1	pectinesterase family protein	1,99E+04	2,81E-02
AT1G72416.1	heat shock protein binding	-5,17E+04	2,83E-02
AT2G46660.1	CYP78A6 (cytochrome P450, family 78, subfamily A, polypeptide 6); oxygen binding	1,86E+04	3,04E-02
AT3G23730.1	xyloglucan:xyloglucosyl transferase, putative / xyloglucan endotransglycosylase, putative / endo-xyloglucan transferase, putative	-4,32E+04	3,21E-02
AT2G25735.1	unknown protein	-3,84E+04	3,35E-02
AT3G44310	NIT1 (NITRILASE 1)	2,24E+04	3,38E-02
AT1G16130.1	WAKL2 (WALL ASSOCIATED KINASE-LIKE 2); kinase	-2,96E+04	3,44E-02
AT3G59080	aspartyl protease family protein	-2,46E+04	3,45E-02
AT4G37610.1	BT5 (BTB and TAZ domain protein 5); protein binding / transcription regulator	-3,22E+04	3,50E-02
AT5G05440.1	unknown protein	-3,56E+04	3,51E-02
AT5G07010.1	sulfotransferase family protein	3,69E+04	3,67E-02
AT5G24270.1	SOS3 (SALT OVERLY SENSITIVE 3)	-2,11E+04	3,72E-02
AT1G15580.1	*IAA*5 (indoleacetic acid-induced protein 5); transcription factor	2,66E+04	3,75E-02

AT4G01250.1	WRKY22 (WRKY DNA-binding protein 22); transcription factor	-1,95E+04	3,87E-02
AT1G50040.1	unknown protein	-2,89E+04	3,90E-02
AT1G04180.1	flavin-containing monooxygenase family protein / FMO family protein	2,62E+04	4,02E-02
AT3G26460.1	major latex protein-related / MLP-related	1,87E+04	4,08E-02
AT3G16420	PBP1 (PYK10-BINDING PROTEIN 1)	1,68E+04	4,32E-02
AT4G08950.1	phosphate-responsive protein, putative (EXO)	-2,45E+04	4,33E-02
AT5G57560.1	TCH4 (TOUCH 4); hydrolase, acting on glycosyl bonds	-4,18E+04	4,44E-02
AT1G16390.1	organic cation transporter-related	-2,46E+04	4,55E-02
AT2G37170.1	PIP2B (plasma membrane intrinsic protein 2;2); water channel	2,11E+04	4,58E-02
AT5G43150.1	unknown protein	2,51E+04	4,60E-02
AT1G03870.1	FLA9 (FLA9)	-2,93E+04	4,79E-02
AT5G22690.1	disease resistance protein (TIR-NBS-LRR class), putative	-1,75E+04	4,87E-02

App. Table 11. Differentially expressed genes in leaves - diploid Col-0 vs. tetraploid Col-0 (four arrays including four individuals of line 1326-26)

Gene ID	Description	log_2 FC	P-value
AT5G59310.1	LTP4 (LIPID TRANSFER PROTEIN 4); lipid binding	-6,70E+04	2,23E-05
AT1G53480.1	unknown protein	4,76E+04	3,54E-05
AT1G05100.1	MAPKKK18 (Mitogen-activated protein kinase kinase kinase 18); kinase	-3,10E+04	4,95E-04
AT2G47770.1	benzodiazepine receptor-related	-4,03E+04	5,22E-04
AT3G26830.1	PAD3 (PHYTOALEXIN DEFICIENT 3); oxygen binding	-2,92E+04	6,35E-04
AT1G80160	lactoylglutathione lyase family protein / glyoxalase I family protein	-4,19E+04	6,95E-04
AT5G13930.1	CHS (CHALCONE SYNTHASE); naringenin-chalcone synthase	-2,88E+04	7,98E-04
AT3G51240.1	F3H (TRANSPARENT TESTA 6); naringenin 3-dioxygenase	-3,04E+04	8,43E-04
AT3G46660.1	UDP-glucoronosyl/UDP-glucosyl transferase family protein	-4,20E+04	1,31E-03
AT1G26945.1	transcription regulator	3,00E+04	1,34E-03
AT3G02480.1	ABA-responsive protein-related	-4,17E+04	1,35E-03
AT2G30770.1	CYP71A13 (cytochrome P450, family 71, subfamily A, polypeptide 13); oxygen binding	-3,47E+04	1,50E-03
AT1G29460.1	auxin-responsive protein, putative	3,17E+04	1,84E-03
AT5G49170.1	unknown protein	2,25E+04	2,43E-03
AT4G14690.1	ELIP2 (EARLY LIGHT-INDUCIBLE PROTEIN 2); chlorophyll binding	-2,13E+04	2,63E-03
AT5G17220.1	ATGSTF12 (GLUTATHIONE S-TRANSFERASE 26); glutathione transferase	-2,79E+04	2,68E-03
AT1G17170.1	ATGSTU24 (*Arabidopsis thaliana* Glutathione S-transferase (class tau) 24); glutathione transferase	-3,26E+04	2,75E-03
AT5G18020.1	auxin-responsive protein, putative	2,64E+04	3,15E-03
AT5G39860.1	PRE1 (PACLOBUTRAZOL RESISTANCE1); DNA binding / transcription factor	3,38E+04	3,47E-03
AT4G38860.1	auxin-responsive protein, putative	2,11E+04	3,53E-03
AT1G29500.1	auxin-responsive protein, putative	2,35E+04	3,79E-03
AT1G79400.1	ATCHX2 (CATION/H+ EXCHANGER 2); monovalent cation:proton antiporter	-2,75E+04	3,97E-03
AT5G50335.1	unknown protein	2,29E+04	4,32E-03
AT5G18050.1	auxin-responsive protein, putative	3,28E+04	4,69E-03

AT5G18010.1	auxin-responsive protein, putative	3,26E+04	4,73E-03
AT2G47930.1	AGP26/ATAGP26 (ARABINOGALACTAN PROTEINS 26)	2,22E+04	5,06E-03
AT5G66590.1	allergen V5/Tpx-1-related family protein	2,21E+04	5,28E-03
AT1G74670.1	gibberellin-responsive protein, putative	5,10E+04	5,47E-03
AT1G66600.1	WRKY63 (WRKY DNA-binding protein 63); transcription factor	-3,17E+04	5,51E-03
AT3G19850.1	phototropic-responsive NPH3 family protein	2,66E+04	6,28E-03
AT5G59320.1	LTP3 (LIPID TRANSFER PROTEIN 3); lipid binding	-5,00E+04	6,33E-03
AT4G15480.1	UGT84A1; UDP-glycosyltransferase/ sinapate 1-glucosyltransferase/ transferase, transferring glycosyl groups	-2,85E+04	6,43E-03
AT2G34170.1	unknown protein	2,10E+04	6,77E-03
AT1G23080	PIN7 (PIN-FORMED 7); auxin:hydrogen symporter/ transporter	1,87E+04	6,78E-03
AT1G43160.1	RAP2.6 (related to AP2 6); DNA binding / transcription factor	-3,63E+04	7,03E-03
AT1G70260.1	nodulin MtN21 family protein	-3,04E+04	7,06E-03
AT5G62350.1	invertase/pectin methylesterase inhibitor family protein / DC 1.2 homolog (FL5-2I22)	1,91E+04	7,07E-03
AT1G29450.1	auxin-responsive protein, putative	2,64E+04	7,44E-03
AT4G01460.1	basic helix-loop-helix (bHLH) family protein	2,14E+04	7,85E-03
AT3G20820.1	leucine-rich repeat family protein	1,88E+04	8,13E-03
AT1G20190.1	ATEXPA11 (*ARABIDOPSIS THALIANA* EXPANSIN A11)	3,28E+04	8,25E-03
AT5G57760.1	unknown protein	2,82E+04	8,49E-03
AT5G05270	chalcone-flavanone isomerase family protein	-2,78E+04	9,01E-03
AT4G37530	peroxidase, putative	-2,27E+04	9,22E-03
AT3G12610.1	DRT100 (DNA-DAMAGE REPAIR/TOLERATION 100); protein binding	2,00E+04	9,26E-03
AT4G13890.1	SHM5 (SERINE HYDROXYMETHYLTRANSFERASE 5); glycine hydroxymethyltransferase	-2,33E+04	9,78E-03
AT5G48900.1	pectate lyase family protein	2,34E+04	9,85E-03
AT1G75170	SEC14 cytosolic factor family protein / phosphoglyceride transfer family protein	-1,84E+04	1,00E-02
AT2G40610.1	ATEXPA8 (*ARABIDOPSIS THALIANA* EXPANSIN A8)	3,73E+04	1,01E-02
AT4G30660.1	hydrophobic protein, putative / low temperature and salt responsive protein, putative	2,39E+04	1,05E-02
AT4G17660.1	protein kinase, putative	-2,93E+04	1,05E-02
AT1G65060	4CL3 (4-coumarate:CoA ligase 3); 4-coumarate-CoA ligase	-2,42E+04	1,06E-02
AT2G28570.1	unknown protein	-1,93E+04	1,08E-02
AT5G57780.1	unknown protein	1,84E+04	1,08E-02
AT3G01490.1	protein kinase, putative	2,16E+04	1,09E-02
AT1G78570.1	RHM1/ROL1 (RHAMNOSE BIOSYNTHESIS1); UDP-glucose 4,6-dehydratase/ catalytic	-2,03E+04	1,13E-02
AT5G39240.1	unknown protein	1,75E+04	1,15E-02
AT1G70260.1	nodulin MtN21 family protein	-3,33E+04	1,16E-02
AT5G12050.1	unknown protein	3,08E+04	1,21E-02
AT2G22122.1	unknown protein	2,37E+04	1,24E-02
AT4G03190.1	GRH1 (GRR1-LIKE PROTEIN 1); ubiquitin-protein ligase	1,86E+04	1,25E-02
AT5G02490.1	heat shock cognate 70 kDa protein 2 (HSC70-2) (HSP70-2)	-2,93E+04	1,25E-02
AT1G04490	unknown protein	-2,23E+04	1,26E-02
AT3G07350.1	unknown protein	1,63E+04	1,27E-02

AT5G64860.1	DPE1 (DISPROPORTIONATING ENZYME); 4-alpha-glucanotransferase	1,84E+04	1,28E-02
AT3G51290.1	proline-rich family protein	2,07E+04	1,31E-02
AT4G34760.1	auxin-responsive family protein	1,95E+04	1,34E-02
AT1G56120.1	leucine-rich repeat family protein / protein kinase family protein	-2,13E+04	1,38E-02
AT5G44260.1	zinc finger (CCCH-type) family protein	2,89E+04	1,43E-02
AT2G18193.1	AAA-type ATPase family protein	-1,81E+04	1,43E-02
AT5G46690.1	BHLH071 (BETA HLH PROTEIN 71); DNA binding / transcription factor	2,75E+04	1,46E-02
AT2G04070.1	MATE efflux family protein	-2,27E+04	1,46E-02
AT2G42900.1	unknown protein	2,10E+04	1,50E-02
AT5G52640.1	HSP81-1 (HEAT SHOCK PROTEIN 81-1); ATP binding / unfolded protein binding	-2,90E+04	1,57E-02
AT5G60100.1	APRR3 (PSEUDO-RESPONSE REGULATOR 3); transcription regulator	1,87E+04	1,60E-02
AT1G62570.1	flavin-containing monooxygenase family protein / FMO family protein	-1,86E+04	1,67E-02
AT5G44210.1	ATERF-9/ATERF9/ERF9 (ERF domain protein 9); DNA binding / transcription factor/ transcriptional repressor	1,70E+04	1,70E-02
AT1G67800	copine-related	-1,55E+04	1,70E-02
AT5G08640.1	FLS (FLAVONOL SYNTHASE)	-3,39E+04	1,70E-02
AT2G26170	CYP711A1 (MORE AXILLARY BRANCHES 1); oxygen binding	-1,76E+04	1,72E-02
AT1G49130	zinc finger (B-box type) family protein	2,43E+04	1,72E-02
AT1G02340.1	HFR1 (LONG HYPOCOTYL IN FAR-RED); DNA binding / transcription factor	1,86E+04	1,72E-02
AT3G29370.1	unknown protein	2,31E+04	1,74E-02
AT4G26670.1	mitochondrial import inner membrane translocase subunit Tim17/Tim22/Tim23 family protein	1,76E+04	1,75E-02
AT1G07050.1	CONSTANS-like protein-related	2,47E+04	1,77E-02
AT3G55240.1	unknown protein	3,18E+04	1,80E-02
AT1G20310.1	unknown protein	-1,95E+04	1,80E-02
AT4G24000.1	ATCSLG2 (Cellulose synthase-like G2); transferase/ transferase, transferring glycosyl groups	-1,92E+04	1,88E-02
AT5G25260.1	unknown protein	-2,58E+04	1,90E-02
AT2G45210.1	auxin-responsive protein-related	-1,79E+04	1,96E-02
AT3G53800.1	armadillo/beta-catenin repeat family protein	2,53E+04	1,97E-02
AT1G10640.1	polygalacturonase, putative / pectinase, putative	2,49E+04	1,98E-02
AT1G04180.1	flavin-containing monooxygenase family protein / FMO family protein	2,15E+04	2,05E-02
AT1G69610.1	unknown protein	-1,57E+04	2,05E-02
AT5G11550.1	binding	2,07E+04	2,08E-02
AT3G07720.1	kelch repeat-containing protein	-1,74E+04	2,08E-02
AT1G17665.1	unknown protein	1,61E+04	2,14E-02
AT5G39210.1	CRR7 (CHLORORESPIRATORY REDUCTION 7)	1,59E+04	2,14E-02
AT5G24080.1	protein kinase family protein	-2,15E+04	2,15E-02
AT1G29440.1	auxin-responsive family protein	2,57E+04	2,18E-02
AT1G65610.1	endo-1,4-beta-glucanase, putative / cellulase, putative	-2,21E+04	2,18E-02
AT1G75800.1	pathogenesis-related thaumatin family protein	1,53E+04	2,20E-02
AT5G50740	metal ion binding	2,33E+04	2,21E-02
AT4G19120	ERD3 (EARLY-RESPONSIVE TO DEHYDRATION 3)	1,70E+04	2,21E-02
AT4G01700.1	chitinase, putative	-1,82E+04	2,28E-02
AT1G32940.1	subtilase family protein	-2,42E+04	2,30E-02
AT5G13140.1	unknown protein	1,79E+04	2,37E-02

AT1G74010.1	strictosidine synthase family protein	-1,59E+04	2,38E-02
AT2G43590.1	chitinase, putative	-3,02E+04	2,39E-02
AT4G23870.1	unknown protein	1,85E+04	2,42E-02
AT5G02760.1	protein phosphatase 2C family protein / PP2C family protein	3,90E+04	2,43E-02
AT5G06690	(THIOREDOXIN-LIKE 5); thiol-disulfide exchange intermediate	2,13E+04	2,43E-02
AT4G26370	antitermination NusB domain-containing protein	1,79E+04	2,50E-02
AT3G05160	sugar transporter, putative	1,48E+04	2,58E-02
AT2G06925.1	ATSPLA2-ALPHA/PLA2-ALPHA (PHOSPHOLIPASE A2-ALPHA); phospholipase A2	1,95E+04	2,59E-02
AT5G40760.1	G6PD6 (GLUCOSE-6-PHOSPHATE DEHYDROGENASE 6); glucose-6-phosphate 1-dehydrogenase	-1,84E+04	2,63E-02
AT1G06000.1	UDP-glucoronosyl/UDP-glucosyl transferase family protein	-2,06E+04	2,63E-02
AT1G72416.1	heat shock protein binding	1,66E+04	2,64E-02
AT5G44680.1	methyladenine glycosylase family protein	2,33E+04	2,69E-02
AT2G21660	ATGRP7 (COLD, CIRCADIAN RHYTHM, AND RNA BINDING 2); RNA binding / double-stranded DNA binding / single-stranded DNA binding	2,07E+04	2,69E-02
AT1G22335.1	unknown protein	1,71E+04	2,71E-02
AT1G32090.1	early-responsive to dehydration protein-related / ERD protein-related	1,57E+04	2,74E-02
AT1G25230.1	purple acid phosphatase family protein	2,38E+04	2,75E-02
AT4G33490.1	pepsin A	2,11E+04	2,78E-02
AT3G03820.1	auxin-responsive protein, putative	2,61E+04	2,79E-02
AT4G39030.1	EDS5 (ENHANCED DISEASE SUSCEPTIBILITY 5); antiporter/ transporter	-2,77E+04	2,81E-02
AT2G42560.1	late embryogenesis abundant domain-containing protein / LEA domain-containing protein	-1,88E+04	2,85E-02
AT5G22310.1	unknown protein	2,17E+04	2,87E-02
AT3G21230.1	4CL5 (4-COUMARATE:COA LIGASE 5); 4-coumarate-CoA ligase	-1,73E+04	2,90E-02
AT3G24982.1	protein binding	-1,70E+04	2,91E-02
AT3G18050.1	unknown protein	1,77E+04	2,93E-02
AT2G33830	dormancy/auxin associated family protein	2,68E+04	3,01E-02
AT5G50360.1	unknown protein	-3,26E+04	3,01E-02
AT1G17020.1	SRG1 (SENESCENCE-RELATED GENE 1); oxidoreductase, acting on paired donors, with incorporation or reduction of molecular oxygen, 2-oxoglutarate as one donor, and incorporation of one atom each of oxygen into both donors	-2,44E+04	3,04E-02
AT4G38840.1	auxin-responsive protein, putative	1,55E+04	3,04E-02
AT5G25460.1	unknown protein	2,57E+04	3,06E-02
AT1G10585.1	transcription factor	-3,43E+04	3,09E-02
AT4G04630.1	unknown protein	1,52E+04	3,12E-02
AT3G04910	WNK1 (WITH NO LYSINE (K) 1); kinase	1,52E+04	3,14E-02
AT5G45380.1	sodium:solute symporter family protein	-1,41E+04	3,19E-02
AT3G60530.1	zinc finger (GATA type) family protein	1,48E+04	3,20E-02
AT3G60390.1	HAT3 (homeobox-leucine zipper protein 3); transcription factor	1,70E+04	3,26E-02
AT4G10040.1	CYTC-2 (CYTOCHROME C-2); electron carrier	-1,46E+04	3,28E-02
AT1G15415.1	unknown protein	-1,82E+04	3,31E-02
AT1G23060	unknown protein	1,57E+04	3,31E-02
AT4G30180.1	transcription factor/ transcription regulator	2,43E+04	3,32E-02

AT4G32340.1	binding	1,87E+04	3,36E-02
AT5G50740.1	metal ion binding	2,58E+04	3,39E-02
AT5G19090	heavy-metal-associated domain-containing protein	1,40E+04	3,41E-02
AT3G02230.1	RGP1 (REVERSIBLY GLYCOSYLATED POLYPEPTIDE 1)	-1,65E+04	3,44E-02
AT3G05360.1	disease resistance family protein / LRR family protein	-1,75E+04	3,45E-02
AT1G01110.1	IQD18 (IQ-domain 18)	1,52E+04	3,46E-02
AT4G25990	CIL	1,61E+04	3,48E-02
AT1G12845.1	unknown protein	3,36E+04	3,48E-02
AT1G10370.1	ATGSTU17/ERD9/GST30/GST30B (EARLY-RESPONSIVE TO DEHYDRATION 9, GLUTATHIONE S-TRANSFERASE 30, GLUTATHIONE S-TRANSFERASE 30B); glutathione transferase	-2,33E+04	3,51E-02
AT1G66370.1	MYB113 (myb domain protein 113); DNA binding / transcription factor	-3,14E+04	3,57E-02
AT5G18030.1	auxin-responsive protein, putative	2,98E+04	3,60E-02
AT2G30550.2	lipase class 3 family protein	-1,79E+04	3,61E-02
AT3G23240.1	ATERF1/ERF1 (ETHYLENE RESPONSE FACTOR 1); DNA binding / transcription factor/ transcriptional activator	-1,88E+04	3,63E-02
AT2G42870.1	unknown protein	3,99E+04	3,63E-02
AT3G49260	IQD21 (IQ-domain 21); calmodulin binding	1,61E+04	3,71E-02
AT1G29395.1	COR414-TM1 (cold regulated 414 thylakoid membrane 1)	1,85E+04	3,71E-02
AT1G12860.1	basic helix-loop-helix (bHLH) family protein / F-box family protein	1,41E+04	3,74E-02
AT1G21550.1	calcium-binding protein, putative	-1,65E+04	3,75E-02
AT2G02100.1	LCR69/PDF2.2 (Low-molecular-weight cysteine-rich 69); protease inhibitor	1,72E+04	3,88E-02
AT3G53250.1	auxin-responsive family protein	2,17E+04	3,92E-02
AT1G14920.1	GAI (GA INSENSITIVE); transcription factor	1,68E+04	3,96E-02
AT3G23880.1	F-box family protein	2,08E+04	3,97E-02
AT3G48240.1	octicosapeptide/Phox/Bem1p (PB1) domain-containing protein	-1,44E+04	3,98E-02
AT5G61820.1	unknown protein	-1,83E+04	3,99E-02
AT5G04190.1	PKS4 (PHYTOCHROME KINASE SUBSTRATE 4)	2,76E+04	4,03E-02
AT2G36630.1	unknown protein	-1,62E+04	4,12E-02
AT3G14150	(S)-2-hydroxy-acid oxidase, peroxisomal, putative / glycolate oxidase, putative / short chain alpha-hydroxy acid oxidase, putative	1,64E+04	4,14E-02
AT5G18080.1	auxin-responsive protein, putative	2,54E+04	4,15E-02
AT5G53710.1	unknown protein	-2,37E+04	4,25E-02
AT5G11070.1	unknown protein	1,78E+04	4,30E-02
AT5G53870.1	plastocyanin-like domain-containing protein	-1,85E+04	4,33E-02
AT5G58600	PMR5 (POWDERY MILDEW RESISTANT 5)	1,77E+04	4,36E-02
AT1G77120.1	ADH1 (ALCOHOL DEHYDROGENASE 1); alcohol dehydrogenase	-1,51E+04	4,39E-02
AT3G13790.1	ATBFRUCT1/ATCWINV1 (*ARABIDOPSIS THALIANA* CELL WALL INVERTASE 1); beta-fructofuranosidase/ hydrolase, hydrolyzing O-glycosyl compounds	-1,35E+04	4,39E-02
AT4G16600.1	glycogenin glucosyltransferase (glycogenin)-related	-1,34E+04	4,42E-02
AT2G01760.1	ARR14 (*ARABIDOPSIS* RESPONSE REGULATOR 14); transcription factor/ two-component response regulator	1,54E+04	4,47E-02
AT2G02580.1	CYP71B9 (cytochrome P450, family 71, subfamily B, polypeptide 9); oxygen binding	-1,39E+04	4,51E-02
AT3G09440.1	heat shock cognate 70 kDa protein 3 (HSC70-3)	-2,36E+04	4,53E-02

	(HSP70-3)		
AT5G61590.1	AP2 domain-containing transcription factor family protein	2,11E+04	4,53E-02
AT1G68390.1	unknown protein	-1,93E+04	4,53E-02
AT3G11340.1	UDP-glucoronosyl/UDP-glucosyl transferase family protein	-5,46E+04	4,53E-02
AT3G06770	glycoside hydrolase family 28 protein / polygalacturonase (pectinase) family protein	1,64E+04	4,58E-02
AT1G50830.1	unknown protein	1,32E+04	4,66E-02
AT2G27050.1	EIL1 (ETHYLENE-INSENSITIVE3-LIKE 1); transcription factor	1,31E+04	4,71E-02
AT4G30650.1	hydrophobic protein, putative / low temperature and salt responsive protein, putative	1,95E+04	4,71E-02
AT4G34970.1	actin binding	1,37E+04	4,72E-02
AT1G68190.1	zinc finger (B-box type) family protein	1,82E+04	4,74E-02
AT5G42010.1	WD-40 repeat family protein	-1,32E+04	4,81E-02
AT2G15960.1	unknown protein	2,03E+04	4,84E-02
AT1G56720	protein kinase family protein	1,60E+04	4,85E-02
AT4G05100.1	AtMYB74 (myb domain protein 74); DNA binding / transcription factor	-2,42E+04	4,88E-02
AT4G27440	PORB (PROTOCHLOROPHYLLIDE OXIDOREDUCTASE B); oxidoreductase/ protochlorophyllide reductase	2,44E+04	4,91E-02
AT2G37950.1	zinc finger (C3HC4-type RING finger) family protein	1,66E+04	4,94E-02
AT4G21680.1	proton-dependent oligopeptide transport (POT) family protein	-1,56E+04	4,98E-02
AT3G22370.1	AOX1A (alternative oxidase 1A); alternative oxidase	-1,42E+04	4,99E-02

App. Table 12. Differentially expressed genes in leaves - diploid Col-0 vs. tetraploid Col-0 (four arrays including four individuals of line 1326-28)

Gene ID	Description	log_2 FC	P-value
AT5G59310.1	LTP4 (LIPID TRANSFER PROTEIN 4); lipid binding	-6,71E+04	1,32E-04
AT5G59320.1	LTP3 (LIPID TRANSFER PROTEIN 3); lipid binding	-5,35E+04	1,22E-03
AT3G02480.1	ABA-responsive protein-related	-4,50E+04	1,98E-03
AT2G02990.1	RNS1 (RIBONUCLEASE 1); endoribonuclease	-5,77E+04	2,34E-03
AT1G53480.1	unknown protein	4,72E+04	3,93E-03
AT1G60190.1	armadillo/beta-catenin repeat family protein / U-box domain-containing protein	-3,51E+04	4,07E-03
AT1G80160	lactoylglutathione lyase family protein / glyoxalase I family protein	-4,53E+04	5,05E-03
AT5G02760.1	protein phosphatase 2C family protein / PP2C family protein	3,65E+04	6,21E-03
AT3G46660.1	UDP-glucoronosyl/UDP-glucosyl transferase family protein	-4,35E+04	6,33E-03
AT3G11480.1	S-adenosyl-L-methionine:carboxyl methyltransferase family protein	-4,99E+04	8,31E-03
AT1G43160.1	RAP2.6 (related to AP2 6); DNA binding / transcription factor	-3,92E+04	9,55E-03
AT3G45060.1	ATNRT2.6 (*Arabidopsis thaliana* high affinity nitrate transporter 2.6); nitrate transporter	-3,95E+04	9,69E-03
AT1G66370.1	MYB113 (myb domain protein 113); DNA binding / transcription factor	-3,40E+04	1,10E-02
AT5G18010.1	auxin-responsive protein, putative	2,76E+04	2,65E-02
AT5G62280.1	unknown protein	3,04E+04	3,11E-02
AT5G44260.1	zinc finger (CCCH-type) family protein	2,57E+04	3,20E-02
AT1G07260.1	UDP-glucoronosyl/UDP-glucosyl transferase family	-3,52E+04	4,71E-02

	protein		
AT3G48520.1	CYP94B3 (cytochrome P450, family 94, subfamily B, polypeptide 3); oxygen binding	-4,57E+04	4,91E-02
AT5G24080.1	protein kinase family protein	-2,66E+04	4,93E-02

App. Table 13.enrichment/depletion of gene ontology (GO) functional groups among the differentially expressed genes in 6-8th leaves of tetraploid Col-0 plants.

GO	Name	FDR	Over/Under
GO:0044260	cellular macromolecule metabolic process	0.0	under
GO:0044267	cellular protein metabolic process	0.0	under
GO:0006412	translation	0.0	under
GO:0043284	biopolymer biosynthetic process	0.0	under
GO:0019538	protein metabolic process	0.0	under
GO:0006414	translational elongation	0.0	under
GO:0032991	macromolecular complex	0.0	under
GO:0030533	triplet codon-amino acid adaptor activity	0.0	under
GO:0060090	molecular adaptor activity	0.0	under
GO:0009059	macromolecule biosynthetic process	0.0	under
GO:0030529	ribonucleoprotein complex	0.0	under
GO:0006519	amino acid and derivative metabolic process	0.0	under
GO:0043232	intracellular non-membrane-bound organelle	0.0	under
GO:0043228	non-membrane-bound organelle	0.0	under
GO:0044249	cellular biosynthetic process	0.0	under
GO:0006807	nitrogen compound metabolic process	0.0	under
GO:0009308	amine metabolic process	0.0	under
GO:0016887	ATPase activity	0.0	under
GO:0008233	peptidase activity	0.0	under
GO:0043170	macromolecule metabolic process	0.0	under
GO:0043234	protein complex	0.0	under
GO:0016817	hydrolase activity, acting on acid anhydrides	0.0	under
GO:0005840	ribosome	0.0	under
GO:0016818	hydrolase activity, acting on acid anhydrides, in phosphorus-containing anhydrides	0.0	under
GO:0016462	pyrophosphatase activity	0.0	under
GO:0006520	amino acid metabolic process	0.0	under
GO:0009058	biosynthetic process	0.0	under
GO:0017111	nucleoside-triphosphatase activity	0.0	under
GO:0032555	purine ribonucleotide binding	0.0	under
GO:0032553	ribonucleotide binding	0.0	under
GO:0042623	ATPase activity, coupled	0.0	under
GO:0032559	adenyl ribonucleotide binding	0.0	under
GO:0005524	ATP binding	0.0	under
GO:0043283	biopolymer metabolic process	0.0	under
GO:0006508	proteolysis	0.0	under
GO:0031974	membrane-enclosed lumen	0.0	under
GO:0043233	organelle lumen	0.0	under
GO:0000166	nucleotide binding	0.0	under

GO:0017076	purine nucleotide binding	0.0	under
GO:0031225	anchored to membrane	0.0	under
GO:0003735	structural constituent of ribosome	0.0	under
GO:0044428	nuclear part	0.0	under
GO:0000003	reproduction	0.0	under
GO:0009733	response to auxin stimulus	1,08E+00	over
GO:0009725	response to hormone stimulus	1,08E+00	over
GO:0009719	response to endogenous stimulus	2,25E+01	over
GO:0050896	response to stimulus	0.00146059	over
GO:0042221	response to chemical stimulus	0.00454521	over
GO:0003700	transcription factor activity	0.00538335	over
GO:0005634	nucleus	0.0183629	over
GO:0009606	tropism	0.0183629	over
GO:0030528	transcription regulator activity	0.0183629	over
GO:0009630	gravitropism	0.0443657	over
GO:0003677	DNA binding	0.0443657	over

App. Table 14 Diploid (Cy3) vs. tetraploid Ler-0 leaves (Cy5; eight biological replicates)

Gene ID	Description	log_2 FC	P-value
AT5G20630.1	GER 3 (Germin3) oxalate oxidase	4,8300	5,55E-03
AT3G11510.1	40S ribosomal protein S14	0,6581	5,57E-03
AT1G53160	SPL4 (Squamosa promoter binding protein-like 4)	-1,6076	6,13E-03
AT2G31270.1	CDT1A (*Arabidopsis* homolog of yeast CDT 1A), cyclin dependent protein kinase	0,9577	6,39E-03
AT3G14190.1	unknown protein	0,9108	9,21E-03
AT5G53210.1	SPCH (Speechless) transcription factor	0,7227	1,09E-02
AT2G27385.1	unknown protein	1,9251	1,14E-02
AT3G02560	40S ribosomal protein S7	0,6953	1,27E-02
AT4G35730.1	unknown protein	1,2314	1,41E-02
AT2G16270.1	unknown protein	0,6309	1,61E-02
AT5G15510.1	unknown protein	0,6369	2,14E-02
AT1G80080.1	TMM (Too Many Mouths)	1,6761	2,58E-02
AT4G14320.1	60S ribosomal protein L36a/L44	0,6835	2,61E-02
AT3G15560.1	unknown protein	0,9693	2,62E-02
AT1G24600.1	unknown protein	-1,6438	2,81E-02
AT5G13840.1	FZR3 (Fizzy related) signal transducer	1,1897	3,01E-02
AT1G06360.1	fatty acid desaturase family protein	2,4842	3,27E-02
AT5G45630.1	unknown protein	-1,6532	3,45E-02
AT4G34620.1	SSR16 (small subunit ribosomal protein 16)	0,6391	3,47E-02
AT2G42840.1	PDF1 (Protodermal Factor 1)	1,8234	4,03E-02
AT5G60390	elongation fator 1-alpha/EF-1-alpha	0,5045	4,15E-02
AT2G34655.1	unknown protein	-1,0287	4,18E-02
AT2G17360.1	40S ribosomal protein S4 (RPS4A)	0,5201	4,44E-02
AT1G07770	RPS15A (ribosomal protein s15A)	0,4366	4,62E-02
AT2G25210.1	60S ribosomal protein L39 (RPL39A)	0,5750	4,75E-02
AT4G00020.1	protein-binding/single-stranded DNA binding	1,0174	4,75E-02
AT3G55280	RPL23AB (ribosomal protein L23AB)	0,5534	4,87E-02

App. Table 15. The free amino acid contents in Col-0 diploid and tetraploid plants

	In first experiment							**In Second experiment**		
	tetraploid					**diploid**				
	Col 19- 1	Col 19-2	Col 12-1	Col 12-2	Col 12-3	Col-0 2x-1	Col-0 2x-2	Col-0 2x-3	Col-0 2x-3	Col-0 2x-5
	c [µg/g]									
Alanin	106,7	74,7	77,8	88,7	139	117	139	105	99	96
Glycin	13,0	6,8	8,8	13,7	11,2	11,5	10,8	20,3	15,3	20,1
Valin	13,2	9,1	11,3	9,7	16,5	10,4	11,5	8,6	9,0	9,1
Leucin	13,9	11,9	15,2	10,6	18,3	9,3	10,6	6,5	7,0	6,7
Isoleucin	5,1	5,1	5,4	4,5	8,2	3,7	3,9	3,7	4,6	4,5
Threonin	74,5	65,3	24,7	56,4	57,6	95,6	111,7	48,2	71,6	61,5
Serin	257	244	102	110	170	68,4	111	113	135	139
Prolin	41,3	15,5	15,8	338,5	28,7	17,2	17,3	16,4	11,9	14,0
Asparagin	17,9	17,1	11,6	9,3	13,4	13,4	15,0	32,8	39,9	43,0
Asparaginacid	125	108	77	98	96	242	228	184	161	180
Methionin	2,7	1,6	3,2	3,4	3,8	0,0	1,9	0,0	0,0	0,0
Glutaminacid	116	92	76	203	181	192	208	188	193	187
Glutamin	246	226	118	121	177	384	390	260	303	406
Phenylalanin	10,8	10,7	11,3	10,5	14,8	10,3	11,1	8,1	9,0	8,3
Lysin	14,9	11,7	15,7	12,5	21,2	10,8	13,6	4,6	5,6	8,3
Histidin	9,0	4,8	6,0	7,6	10,0	4,5	5,0	3,7	4,5	6,8
Tyrosin	8,5	6,0	6,4	3,8	11,7	5,0	4,6	3,2	3,2	3,2
Tryptophan	2,0	2,0	2,8	2,0	3,2	1,3	1,5	1,4	1,3	1,3

App. Table 16. The free amino acid contents in L*er*-0 diploid and tetraploid plants

	tetraploid						**diploid**		
	L*er*10-1	L*er*10-2	L*er*10-3	L*er*41-1	L*er*41-2	L*er*41-3	L*er*-0 1	L*er*-0 2	L*er*-0 3
	c [µg/g]								
Alanin	59,4	64,3	64,3	66,3	65,9	99,6	68,0	98,2	90,3
Glycin	5,8	7,8	4,2	6,4	13,8	10,8	5,9	8,8	5,1
Valin	9,3	10,3	7,9	9,6	15,5	14,2	7,6	8,5	10,7
Leucin	8,6	12,4	9,7	14,9	14,5	15,5	9,8	12,5	11,4
Isoleucin	3,3	4,2	3,2	4,7	5,5	5,4	3,0	5,2	4,6
Threonin	79,1	76,8	41,0	54,2	116,7	160,1	45,6	95,9	86,0
Serin	51,4	79,3	48,4	102,8	129,2	281,4	59,0	107,5	96,6
Prolin	19,3	8,5	17,7	11,3	130,2	24,8	8,8	13,2	16,7
Asparagin	13,3	18,8	11,4	9,2	25,8	33,2	14,1	18,2	33,1
Asparaginacid	272,7	187,3	221,8	143,5	129,1	291,4	192,5	289,2	272,3
Methionin	2,1	2,3	2,9	3,4	3,5	4,0	3,3	3,3	3,4
Glutaminacid	259,9	264,5	223,8	294,4	213,5	370,2	183,4	259,0	407,3
Glutamin	227,3	180,2	184,5	110,2	453,2	434,1	272,3	412,2	439,7
Phenylalanin	8,7	10,9	9,6	12,6	12,7	13,3	8,0	9,9	10,7
Lysin	8,6	9,5	9,5	12,4	13,4	14,0	6,2	9,9	11,6
Histidin	3,2	2,0	4,1	3,3	4,3	7,6	2,8	3,8	4,0
Tyrosin	4,8	4,8	5,9	8,0	6,9	6,8	4,6	4,4	3,6
Tryptophan	1,5	1,7	1,8	1,5	2,4	1,9	1,4	1,6	1,5

8. Appendix Published Work

Impact of natural genetic variation on the transcriptome of autotetraploid *Arabidopsis thaliana*

Zheng Yu[a], Georg Haberer[b], Michaela Matthes[a], Thomas Rattei[c], Klaus F. X. Mayer[b], Alfons Gierl[a], and Ramon A. Torres-Ruiz[a,1]

[a]Lehrstuhl für Genetik, Technische Universität München, Wissenschaftszentrum Weihenstephan, D-85350 Freising, Germany; [b]Munich Information Center for Protein Sequences/Institute for Bioinformatics, Helmholtz Zentrum München, D-85764 Neuherberg, Germany; and [c]Department of Genome Oriented Bioinformatics, Technische Universität München, Wissenschaftszentrum Weihenstephan, 85350 Freising, Germany

Edited by Detlef Weigel, Max Planck Institute for Developmental Biology, Tübingen, Germany, and approved August 19, 2010 (received for review January 22, 2010)

Polyploidy, the presence of more than two complete sets of chromosomes in an organism, has significantly shaped the genomes of angiosperms during evolution. Two forms of polyploidy are often considered: allopolyploidy, which originates from interspecies hybrids, and autopolyploidy, which originates from intraspecies genome duplication events. Besides affecting genome organization, polyploidy generates other genetic effects. Synthetic allopolyploid plants exhibit considerable transcriptome alterations, part of which are likely caused by the reunion of previously diverged regulatory hierarchies. In contrast, autopolyploids have relatively uniform genomes, suggesting lower alteration of gene expression. To evaluate the impact of intraspecies genome duplication on the transcriptome, we generated a series of unique *Arabidopsis thaliana* autotetraploids by using different ecotypes. *A. thaliana* autotetraploids show transcriptome alterations that strongly depend on their parental genome composition and include changed expression of both new genes and gene groups previously described from allopolyploid *Arabidopsis*. Alterations in gene expression are stable, nonstochastic, developmentally specific, and associated with changes in DNA methylation. We propose that *Arabidopsis* possesses an inherent and heritable ability to sense and respond to elevated, yet balanced chromosome numbers. The impact of natural variation on alteration of autotetraploid gene expression stresses its potential importance in the evolution and breeding of plants.

allopolyploidy | autopolyploidy | evolution

Polyploidy has fundamentally influenced the speciation and evolution of plants and animals (1–6). To succeed, newly occurring polyploids must overcome notable challenges: genomic instability based on aberrant chromosome segregation during meiosis (3, 4, 6), and rapid adaption to selective environmental pressures that includes competition, for instance, with their diploid progenitors (4, 5, 7). Among known polyploid plants, allopolyploids show a taxonomic predominance (2, 3, 5). However, increasing evidence indicates that the actual appearance of autotetraploid plants in nature might be significantly underestimated (3, 5, 8, 9). The basis for their evolutionary success remains unclear.

Polyploidy has not only significantly shaped the genomes of plants throughout their evolutionary history (2, 9, 10) but has also impacted other genetic and epigenetic aspects including gene expression (4, 7). Studies on differential gene expression and transcriptomics have mainly focused on (neo-) allotetraploids such as wheat, cotton, maize (a segmental allotetraploid; ref. 11), and prominently, resynthesized *Arabidopsis suecica* from (neo-) tetraploid *A. thaliana* and *A. arenosa* (12–18). Transcriptional profiling of two *A. suecica* lines revealed that the expression of >1,400 genes diverged from the midparent value (16). This profiling demonstrated that allopolyploid plants exhibit considerable transcriptome alterations as compared with their diploid progenitors. As allopolyploids arise from interspecies hybrids, part of these changes are likely caused by reunion of previously diverged regulatory hierarchies. In contrast, autopolyploid plants, which result from intraspecies genome duplication, have uniform genomes whereby significant transcriptome alterations would be unexpected. Supporting this notion, an accompanying control experiment of the *A. suecica* analysis detected only negligible differences in gene expression between diploid and a tetraploid *A. thaliana* ecotype L*er* line (16). Similarly, the analysis of 9,000 genes in potato auto(poly)ploids revealed few very weak differences in comparison with diploids (19). Together with the uniformity of autopolyploid genomes, these albeit limited analyses suggested an absence of significant transcriptome alterations in autopolyploid plants, reminiscent to findings in tetraploid yeast (20).

We were interested to test whether significant gene expression alterations can be found among newly synthesized autopolyploids. A series of *A. thaliana* autotetraploids from nine different ecotypes was subjected to gene expression/transcriptome analysis. Our study uncovers an ecotype-dependent, heritable capacity to significantly change gene expression in autotetraploid *A. thaliana*.

Results

***A. thaliana* Col-0 but Not L*er*-0 Ecotype Shows Significant Transcriptome Alteration in Response to Tetraploidy.** To evaluate the impact of intraspecies genome duplication, we conducted a series of transcriptome analyses with numerous *A. thaliana* neo-autotetraploids (see Table S1 for overview of experimental layouts). First, we compared the seedling transcriptome of tetraploid Col-0 lines with their diploid Col-0 progenitor (Fig. 1*A*). Four recently generated independent tetraploid lines of the third generation, after induction, were used (21). Although tetraploid plants typically exhibited enlarged cells and tissues in comparison with diploids, overall structural morphology remained unchanged (Fig. 1*B*). These lines were repeatedly assessed by flow cytometry and chromosome counts for their ploidy (ref. 21; Fig. 1*C*). Transcriptome analysis defined 476 genes (286 up- and 190 down-regulated) that exhibited significant changes in gene expression (cutoff threshold, 1.5-fold; additional 112 genes displayed more subtle fold changes, FCs) (Dataset S1). We performed the same analysis with a series of independently generated tetraploid L*er*-0 lines (Fig. 1 *D* and *E*). In contrast to Col-0, comparison of tetraploid L*er*-0 vs. diploid L*er*-0 seedlings detected only nine genes of disparate functions (all >1.5-fold suppressed; Fig. 1*F*) (Dataset S2). Notably, these nine genes in L*er*-0 tetraploids were not altered between di- and tetraploid Col-0.

Using the same lines as described, the transcriptome of the sixth to eighth rosette leaves of tetraploid Col-0, versus diploid Col-0, was then analyzed to represent a second developmental stage and tissue. Correspondingly, 247 genes were differentially

PLANT BIOLOGY

Author contributions: A.G. and R.A.T.-R. designed research; Z.Y., G.H., and M.M. performed research; Z.Y., G.H., M.M., T.R., K.F.X.M., and R.A.T.-R. analyzed data; and R.A.T.-R. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE18482).

[1]To whom correspondence should be addressed. E-mail: ramon.torres@wzw.tum.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1000852107/-/DCSupplemental.

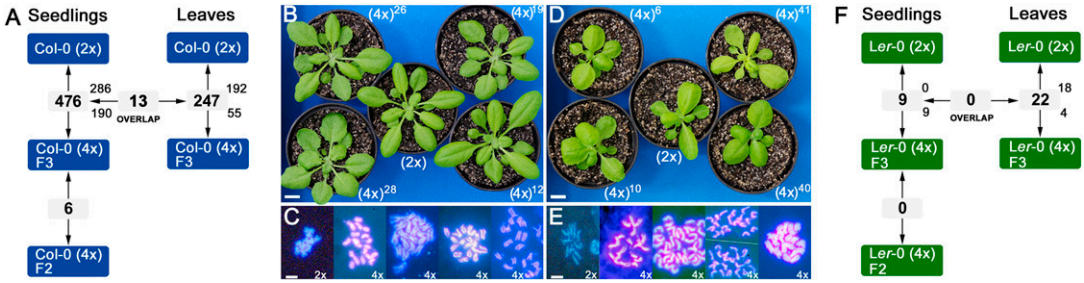


Fig. 1. Transcriptome alterations and morphology of autotetraploid *A. thaliana* Col-0 and L*er*-0. (*A*) Transcriptome alterations in Col-0 between diploids and tetraploids (tissues and generations as indicated). Up- or down-regulated genes at the top and bottom of the shaded boxes, respectively. (*B*) Morphology of diploid (2×) and tetraploid Col-0 (4×) lines (indicated by numbers) at the rosette stage. (*C*) Mitotic chromosome figures of Col-0 root tip cells (diploid line, tetraploid lines 12, 19, 26, and 28 from left to right). (*D*) Morphology of diploid (2×) and tetraploid L*er*-0 lines (indicated by numbers) at the rosette stage. (*E*) Mitotic chromosome figures of L*er*-0 root tip cells, at different mitotic stages (diploid line, tetraploid lines 6, 10, 40, and 41 from left to right). (*F*) Transcriptome alterations in L*er*-0 between diploids and tetraploids (tissues and generations as indicated; up- and down-regulated genes as in *A*). Only alterations with >1.5-fold changes are shown ($P < 0.05$). Note the morphological similarity of di- and tetraploid plants at the rosette stage. (Scale bars: *B* and *D*, 1 cm; *C* and *E*, 10 μm.)

expressed between the sixth and eighth tetraploid and diploid rosette leaves, of which 192 were more than 1.5× up- and 55 were more than 0.67× down-regulated, respectively (42 additional genes exhibited more subtle changes) (Dataset S1). Again, as observed in the seedlings, the sixth to eighth tetraploid L*er*-0 leaves exhibited few transcriptome changes; in total 22, with 18 up- and 4 down-regulated (five additional genes exhibited more subtle changes) (Dataset S2). Although the microarrays used are based on the Col-0 sequence (22), we calculated that this explains little of the response difference of tetraploid L*er*-0 vs. Col-0 (*SI Materials and Methods*). Thus, upon shift from di- to tetraploidy, Col-0 responds with the alteration of gene expression of several hundred genes, whereas L*er*-0 shows minimal altered gene expression. We classified these ecotypes as responder (Col-0) and nonresponder (L*er*-0), respectively.

Alteration of Gene Expression Response to Tetraploidy Depends on Developmental Stage. The Gene Ontology (GO) groups represented by the detected genes, as described in The Arabidopsis Information Resource (TAIR) representations, covered almost all important functional groups of biological processes and molecular functions (Fig. S1). Further, an analysis for significant enrichments of GO groups refined this overview and uncovered under- and over-representation related to various functions/processes (Dataset S3). This analysis was extended by a deeper term-supported comparative in silico analysis based on term-supported matching (*Materials and Methods*), which delivered a striking enrichment of genes related to specific functional categories (Fig. 2*A*). In seedlings, we found gene groups related to photosynthesis and chlorophyll, sugar and cell wall biosynthesis, metal ions, calcium, ATPases, and transcriptional control including six NAC transcription factors (Fig. 2*A*). Several of the most highly up- or down-regulated genes covered ethylene-, stress-, senescence- and defense-related processes, respectively, many with adjusted *P* values far below 0.05 (Table S2). Subsequent RT-PCR tests on diploid and tetraploid tissue directly compared amplification products on gels. Only those genes that showed clear differences were further followed. According to this preselection, ≈55% of the selected genes (Figs. 2*B* and 3; together for seedling and leaf material) displayed significant differences between di- and tetraploid gene expression. These cases enabled us to control the representation of alterations in gene expression in the different functional categories by quantitative RT-PCR (qRT-PCR) (Fig. 2*B*). In comparison with the seedlings, the in silico scan of the detected Col-0 leaf genes indicated a significantly changed pattern of altered activity (Fig. 2*A*). The "cell wall/sugar program" had been extensively reduced, 87 vs. 26 genes, with only two overlaps (*At1g22400* and *At4g30270*). Only one NAC transcription factor was found. The seedling ethylene/stress program had been considerably reduced in "favour" of an auxin synthesis/signaling program with many IAA-antagonists of auxin responsive factors (ARFs), and Short Auxin Upregulated RNAs (SAUR)-like genes (20 genes; FCs > 2.0; Table S3). Microarray data were confirmed by qRT-PCR analyses of genes representing diverse GO functional groups (Fig. 2*B*). One particularly interesting case included an overexpressed SAUR gene-cluster (designated *At5g180-c*) comprising six highly homologous copies (Fig. S2). The genes are dispersed in a region of 20 kb with some copies <2 kb apart. Overexpression of *At5g180-c* was predominantly caused by *At5g18010*. Its overrepresentation within cDNA clones was 12/58 in tetraploids vs. 2/29 in diploids. Only 13 of the genes with >1.5-fold up- or down-regulation, respectively, overlapped between seedlings and leaves (Fig. 1).

Transcriptome of Tetraploid Col-0 and L*er*-0 Is Highly Stable in Consecutive Generations. The tetraploid L*er*-0 and Col-0 lines analyzed in this study exhibited high chromosome number stability during consecutive generations (21). We investigated the stability of the tetraploid transcriptome by analyzing microarray expression profiles of seedlings, two and three generations after induction. This analysis revealed an almost complete identity at a genome-wide level. We did not find any differences in the second vs. third tetraploid L*er*-0 comparison. The comparison of second vs. third tetraploid Col-0 revealed only six differences (Fig. 1 and Dataset S1). Thus, both the unaltered and the altered tetraploid transcriptome of the nonresponding L*er*-0 and the responding Col-0, respectively, remain genetically stable.

Microarray Analysis Detects a Species-Specific Locus That Is Strongly Overexpressed in Both Seedlings and Leaves. Among the transcripts more abundant in tetraploids than in diploids was *At1g53480*. The corresponding gene, named *MRD1,* had been shown to be transcriptionally suppressed in a former microarray analysis of the *Arabidopsis mto1-1* mutant (23). However, its function remained unclear. *MRD1* is (weakly) expressed throughout the adult plant development i.e., in seedlings, young rosette leaves, old rosettes, and siliques (ref. 23; Fig. 3 *A* and *B*). Analysis of T-DNA insertion lines (Fig. 3*C* and *SI Materials and Methods*) did not reveal a conspicuous phenotype with respect to seedling viability, overall morphology, and fertility. *MRD1* and its homolog *At5g03090* appear to be species-specific loci of unknown function, because truncated copies were only found in *A. lyrata* among all plant sequence compilations (Fig. S3). *MRD1* displays a weak basic expression in diploid Col-0, diploid L*er*-0, and tetraploid L*er*-0. However, as verified by qRT-PCR, this locus displayed >20- to 110-fold (leaves vs. seedlings) overexpression in tetraploid Col-0 (Fig. 3*A*; compare also with Fig. 4). Northern blot analysis of *MRD1* confirmed this observation (Fig. 3*B*). Interestingly, *MRD1* overlaps with a second gene coded on the opposite strand, *At1g53490*. This gene is also altered in its expression in tetraploids but at much lower level. Although overlapping, this gene seems

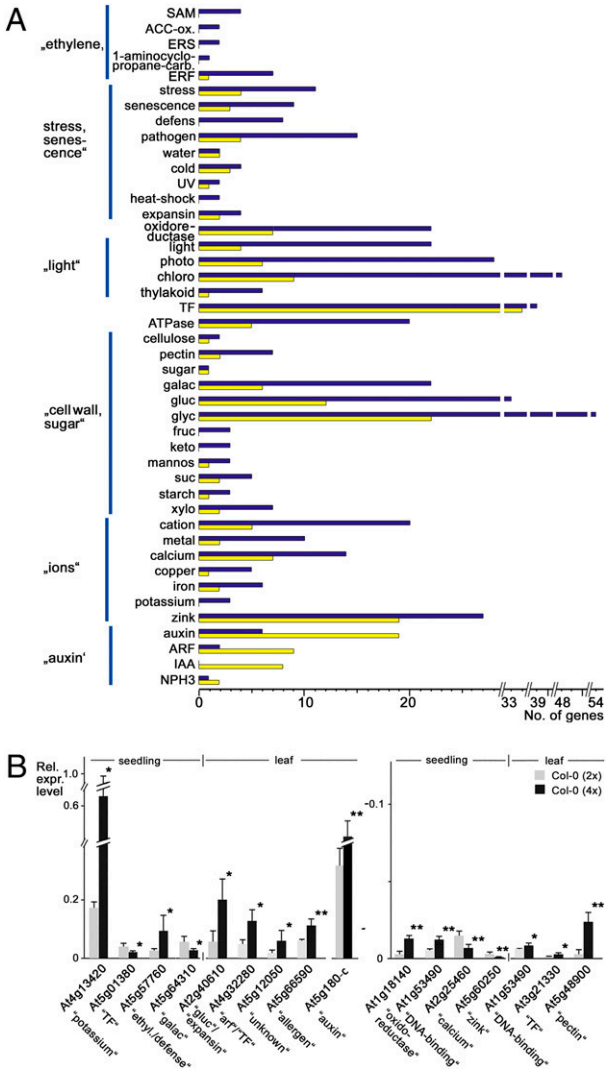


Fig. 2. Development and transcriptome alteration in tetraploid Col-0. (*A*) Comparison of conspicuous functional GO terms with altered expression in Col-0 tetraploids: seedling (blue bars) vs. leaf (yellow bars). Terms in quotation marks indicate key processes covered by the selected functional terms (for details see text and *SI Materials and Methods*). (*B*) Altered expression of selected genes in *A. thaliana* Col-0 autotetraploids as shown by qRT-PCR of genes representing disparate functional categories. At5g180-c indicates qRT-PCR of a complete "SAUR-like" gene cluster comprising six members: *At5g18010-30*, *At5g18050-60* and *At5g18080*. The reference gene in these analyses was *ACT2* (as in ref. 16). This analysis verified 55% of the selected genes from the microarray analysis to be altered between di- and tetraploids. Significance values of one-tailed *t* test: $*P \leq 0.05$; $**P \leq 0.01$; $***P \leq 0.0005$; bars with SD.

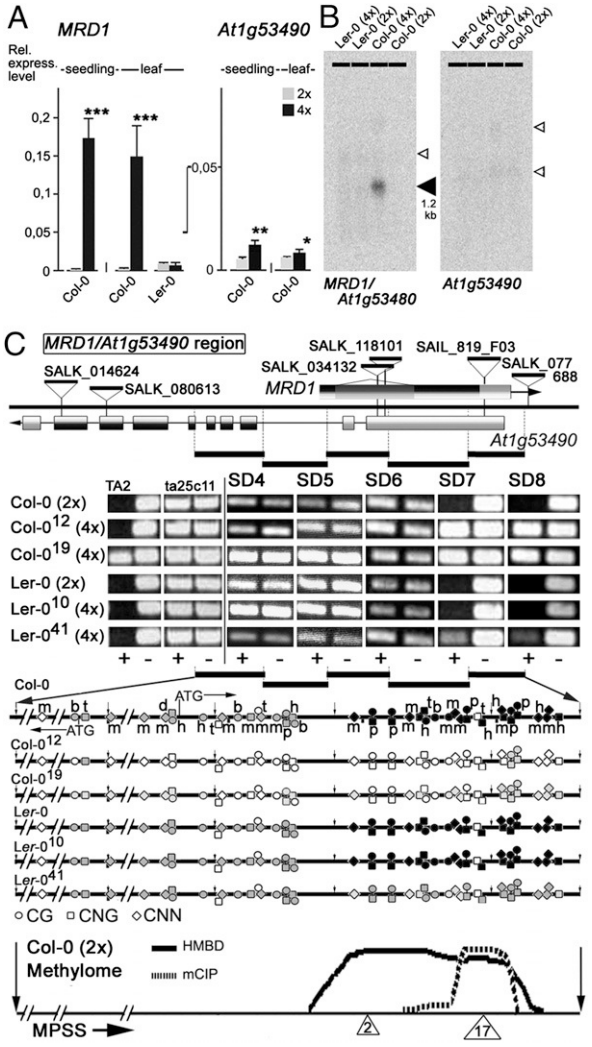


Fig. 3. Expression and methylation of *MRD1*(*At1g53480*) and *At1g53490*. (*A*) qRT-PCR of *MRD1* and *At1g53490* in diploid vs. tetraploid Col-0 and Ler-0. (*B*) Northern blot with *MRD1* (*Left*) and *At1g53490* (*Right*). Size of approximate *MRD1* transcript length is given (filled arrowhead); open arrowheads indicate weak bands probably including homologous gene copies. (*C*) Integrates analysis with methylation requiring (McrBC) and methylation sensitive enzymes. (*Top*) Structure of the *MRD1*/*At1g53490* region including tested T-DNA insertions. A TAIR annotated intron (triangle) was not found in this study. (*Middle*) McrBC analysis of subregions SD4-SD8, transposon TA2, and an nonmethylated ta25c11 repeat DNA sequence tile (taken from ref. 33). Complete methylation is indicated by the absence of a band ("+" and "–" indicate that McrBC was included or excluded, respectively). Note the demethylation of TA2 in Col-0^{19} (4×). (*Bottom*) Methylation and demethylation at sites for BstUI (b), MboI (m), DrdI (d), HpaII/MspI (p), Hpy188III (h), and TseI (t) resolved as CG, CNG, and CNN methylation sites is indicated by circles, squares, and diamonds, respectively. Shading indicates strong (black), weak (dark and light gray) and no methylation (blank). The methylome in this region (24) shows strong methylation at the 3′-end of *MRD1* for Col-0 (2×) detected with monoclonal methylcytosine antibodies (mCIP) and affinity purification with the methylcytosine binding domain of human MeCP2 (HMBD) (24). The MPSS project (26) revealed several short RNA signatures in particular for the 3′-region of *MRD1* (numbers in triangles). Lines and ploidies are indicated. For details, see text, *SI Materials and Methods*, and Figs. S3 and S4.

not to detectably affect the expression of *MRD1* (Fig. 3 *A* and *B*). This finding is in line with other pairs of overlapping genes (24). Alteration of (trans) gene expression has been shown to correlate with epigenetic phenomena in tetraploids, including modulation of DNA methylation (12, 15, 25). We therefore performed DNA methylation analyses, which scanned the methylation status of consecutive segments of this region by comparing the effects of methylation-sensitive enzymes with the enzyme McrBC, which cuts only when DNA contains methylated cytosines (Fig. 3*C* and Figs. S3 and S4). This analysis showed that low transcriptional activity of *MRD1* in tetraploid L*er*-0 is accompanied by partial or complete methylation in the 3′-region, whereas in tetraploid Col-0, its strong expression is correlated with strong demethylation in the same region (Fig. 3*C*). The promoter region of *MRD1* is generally, although not completely, demethylated in di- and tetraploid lines. In fact, the methylome project of diploid Col-0 has shown that this gene is "body-methylated" not "promoter-methylated" (ref. 24; TAIR9 GBRowse: http://gbrowse.arabidopsis.org/cgi-bin/gbrowse/arabidopsis/). In addition, data from the Massively Parallel Signature Sequencing (MPSS) project indicate an accumulation of small

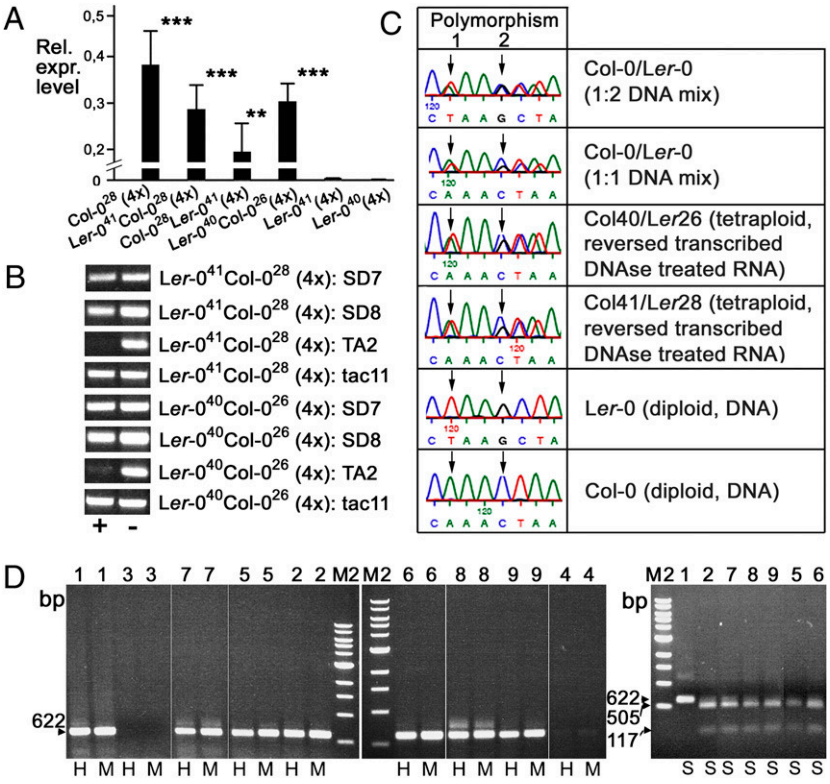


Fig. 4. Inheritance of *MRD1* overexpression and methylation in tetraploid *A. thaliana* F_1 hybrids. (*A*) Real-time qRT-PCR of leaf material of hybrid tetraploid Col-0/L*er*-0 combinations. (*B*) McrBC-Methylation analysis of *MRD1* (regions SD7 and SD8; Fig. S3). Comparisons with TA2 and the ta25c11 (=tac11) are as in Fig. 3. (*C*) Sequence reactions of RT-PCR-amplified transcripts identifying $MRD1^{Col}$ and $MRD1^{Ler}$ alleles in tetraploid Col-0/L*er*-0-hybrids. Arrows point to positions of sequence polymorphisms in *MRD1* (Fig. S3). (*D*) Genomic DNA of diploid lines Col-0 [1] and L*er*-0 [2], tetraploid lines Col-0^{12} [3], Col-0^{19} [4], L*er*-0^{10} [5], L*er*-0^{41} [6], tetraploid hybrids L*er*-0^{40}/Col-0^{26} [7], L*er*-0^{41}/Col-0^{28} [8], and Col-0^{28}/L*er*-0^{41} [9], respectively, were digested with HpaII (H) and MspI (M) (left gels). PCR was performed by using primers flanking a Col-0/L*er*-0 StuI restriction enzyme polymorphism (Fig. S3). Blocking of HpaII/MspI digestion enabled the generation of a band (at 622 bp). The resulting bands amplified from the HpaII and MspI-digested genomic DNA were isolated, purified, combined for each line, redigested with StuI (*S*), and separated again (right gel). Significance values of one-tailed *t* test: $*P \leq 0.05$; $**P \leq 0.01$; $***P \leq 0.0005$; bars with SD (comparison with tetraploid L*er*-0^{40} and L*er*-0^{41}). Lines are as in Fig.1, with ecotypes and ploidies indicated.

RNAs in particular for the 3′-region of the gene (Fig. 3*C* and ref. 26; TAIR9 GBrowse).

Altered Transcription of *MRD1* in Tetraploids Is Heritable. The striking difference between Col-0 and L*er*-0 neo-tetraploids allowed us to test for *MDR1* expression in tetraploid hybrids and, consequently, for inheritance and outcrossing of the tetraploid *MDR1* Col-0 response. We therefore generated reciprocal tetraploid Col-0/L*er*-0 hybrids by using the established lines. In fact, qRT-PCR showed that the capability of sensing and responding to tetraploidy by Col-0 is transmitted to the hybrid (Fig. 4*A*). In addition, methylation analyses showed that this overexpression was accompanied by maintained demethylation at this locus (Fig. 4*B*). Because $MRD1^{Col\text{-}0}$ and $MRD1^{Ler\text{-}0}$ display several polymorphisms (Fig. S3) we directly sequenced reverse transcribed mRNA from different tetraploid Col-0/L*er*-0 hybrids to assess ecotype specific polymorphisms in the transcripts. Interestingly, the sequence signal peaks indicated that both Col-0 and L*er*-0 *MDR1* alleles were expressed with the same intensity (Fig. 4*C*). Thus, the transcription of $MRD1^{Ler\text{-}0}$ appeared to be higher in F_1 Col-0/L*er*-0-hybrid tetraploids than in diploid and tetraploid L*er*-0. We then analyzed the methylation status of *MRD1* in the hybrids by taking advantage of a polymorphic StuI-restriction enzyme recognition site (present in L*er*-0 and absent in Col-0 SD7 region; Fig. S3). Purified SD7-DNA, which blocked methylation-sensitive HpaII and MspI enzymes turned out to originate almost exclusively from L*er*-0 (Fig. 4*D*).

***A. thaliana* Tetraploid Transcriptome Response Is Ecotype Specific.** We were interested to test whether some of the genes detected, in particular *MRD1*, would show expression alteration response in other ecotypes. We generated tetraploids of seven additional ecotypes. Because of its strong overexpression, we reasoned that *MRD1* might be a valuable tool for monitoring ploidy-affected gene expression in *A. thaliana*. In fact, expression analyses of leaf material of the new neo-tetraploid ecotypes revealed considerable variability with respect to absolute and relative expression differences of *MRD1* (Fig. 5*A*). Not surprising, the absolute expression levels differed between ecotypes. For four of seven ecotypes significantly altered *MRD1* expression was observed. This test was complemented with experiments by using two additional genes, *IAA29* (*At4g32280*) and the SAUR-gene cluster (*At5g180-c*). These experiments uncovered almost the same variability (Fig. 5 *B* and *C*). CT-1 and L*er*-1 turned out to be nonresponders in all cases, whereas the other five showed significant expression differences in two or all three genes analyzed.

Discussion

Alteration of Transcriptome in *A. thaliana* Autotetraploids Depends on Ecotype, i.e., Genome Composition. It was generally expected that the uniform genomes of autopolyploids, in contrast to those of allopolyploids, should not exhibit significant gene expression alterations. This observation is supported by limited analysis (16, 19). The presented data on Col-0 vs. L*er*-0 transcriptome comparison demonstrate significant ecotype specific differences in gene expression alterations when the diploid is compared with the tetraploid. Col-0 alters several hundred genes in two tissues, suggesting that more might be uncovered in other tissues. Although this amount is significantly less than found in allote-

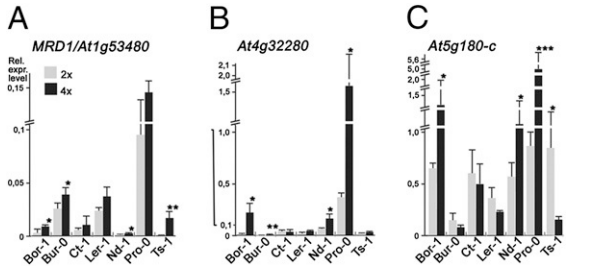


Fig. 5. Di- vs. tetraploid expression profile of selected genes in various *A. thaliana* ecotypes. (*A*) Expression of *MRD1* in seven additional diploid vs. tetraploid ecotypes (qRT-PCR). (*B*) Same analysis as in *A* for *IAA29/At4g32280*. (*C*) Same analysis as in *A* for SAUR gene cluster *At5g180-c*. Leaf material, ecotypes, and ploidies are indicated. Significance values of one-tailed *t* test: $*P \leq 0.05$; $**P \leq 0.01$; $***P \leq 0.0005$; bars with SD. For details, see text.

traploids (16), it is in sharp contrast to tetraploid L*er*-0, which displays an almost diploid expression profile. Limited analysis of other ecotypes with selected probes supports the notion that the response to tetraploidy is variable and depends on the genomic composition. In the Bor-1 and Nd-1 ecotypes, all three genes were up-regulated, whereas in other ecotypes, only two of the genes were altered in their expression. Ct-1 and L*er*-1 did not show any response to all three genes. Whether this observation indicates variable degrees of response capability has to be further investigated. Thus, in answer to our question, some autopolyploids react in a similar, but more subtle way than allopolyploids. It should be mentioned, that the ecotype specific gene expression alterations shown in this study are also clearly distinct from aneuploid syndromes (27–29), because they occurred in *A. thaliana* autotetraploids, i.e., balanced euploids. In contrast, aneuploidy is an out-of-balance situation leading to extensive gene expression alterations in *Arabidopsis* (28) and segregation distortion of loci such as *Arabidopsis SENSITIVE TO DOSAGE IMBALANCE* (*SDI*) (29).

Transcriptome Alterations in Autotetraploid *Arabidopsis* Are Developmentally Specific. The data show that gene expression alterations in autotetraploids are developmental stage specific. This finding is reflected by the low overlap (13 genes) between altered seedling and leaf transcriptomes and by different representation of GO groups (Fig. 1 and Datasets S1 and S3). Apparently, the Col-0 response is a general alteration or relaxation of gene expression control covering genes of different stages. The functional gene groups displayed by seedlings and leaf are well known from these tissues. Seedlings display a biphasic mode of ethylene-related gene activity (30), whereas any form of leaf organogenesis is tightly linked to localized auxin accumulation and auxin-driven gene activities (31, 32). Interestingly, neo-allopolyploid *A. suecica* also revealed a conspicuous alteration of ethylene/stress-related genes (16) showing partly similar reactions in both forms of polyploidy. However, they also revealed different gene expression alterations not observed in autopolyploids such as those considering heat shock genes. It is likely that some of these genes are active or inactive during stages, which do not correspond to the developmental program of their parents.

***A. thaliana* Transcriptome Alteration Response to Tetraploidy Has a Genetic Basis and Displays Epigenetic Phenomena.** The comparison of Col-0 vs. L*er*-0 tetraploids clearly showed that the transcriptome alteration response does not depend on the chromosome number per se, but on the origin of the chromosomes. Furthermore, the alteration was completely transmitted through selfing to the next generation. Selecting a strongly overexpressed gene (*MRD1/At1g53480*) to study the transmission in reciprocal crosses of Col-0 × L*er*-0 tetraploids demonstrated that the response in Col-0 is transmitted to the hybrids as well. Notably, in these cases, only two chromosome sets originate from the "responsive" Col-0 ecotype. Taken together, this result suggests that Col-0 but not L*er*-0 possesses one or more genetic factors that are capable of sensing the alteration of genome dosage and inducing gene expression alterations. Also, the analysis of other ecotypes shows that this ability depends at least partly on the genotype. Possibly, the absence of *MRD1* overexpression in some tetraploids is due to mutation. It is known that L*er*-0 originates from X-irradiated parents (NW20; TAIR). However, the reasons for the observed expression alterations might be more complex. For instance, diploid Col-0 and L*er*-0 genomes possess variable DNA methylation patterns (33). Although this natural epigenetic variability seems not to cause significant gene expression differences in diploids (33), we do not know whether this variability could contribute as such at the tetraploid level.

At this point of discussion, it seems necessary to differentiate between sensing vs. induction vs. transmission/preservation. Although we do not know the sensing factors, we can speculate what they could sense. Altered nuclear surface to volume ratios in tetraploids have been discussed as causative for gene expression/regulatory changes (4, 34). Polyploids generally show increased nuclei, which implies an altered nuclear surface to volume ratio. The gene expression alteration of *MRD1* in various Col-0 vs. L*er*-0 tetraploids and Col-0/L*er*-0 tetraploid hybrids are strongly correlated with DNA (de)methylation. Several analyses of selected (trans) genes have demonstrated changes in gene expression between plants with altered ploidy grade (12–15, 25, 35, 36), some of these have also uncovered a link to epigenetic phenomena, in particular DNA (de)methylation. Upon sensing a higher chromosome number in a nucleus with an altered surface to volume ratio, the induction of DNA (de)methylation of selected genes could be caused by targeted re- and demethylation mechanisms, which have been recently discovered in *Arabidopsis* (37, 38). These and similar mechanisms are also responsible for the preservation of the DNA methylation. Basically, the study of *MRD1*, which belongs to the ≈33% "body-methylated" *A. thaliana* genes (24), indicates one epigenetic option for maintaining the observed transcriptome alterations. However, the observed alterations should not be assigned to DNA methylation alone. Epigenetic effects can be based on other DNA modifications. Furthermore, alteration of the DNA methylation pattern of one transcription factor/repressor could be sufficient to alter the expression of other genes without any further change of their methylation.

Based on the sequence data of reversed transcribed *MRD1*-RNA, it is tempting to speculate that $MRD1^{Ler\text{-}0}$ displays higher transcriptional activation in the hybrids. This activation could happen post fertilization unlike transcriptional reactivation of transposons in pollen (39). Alternatively, this gene could be activated during gametogenesis and then silenced upon fertilization. Then this silencing would be suppressed in tetraploid Col-0 and F_1 Col-0/L*er*-0 hybrids because of the presence of chromosomes originating from tetraploid Col-0. The final effect resembles the opposite of paramutation of loci such as maize B-I (40). However, it is also possible that the dosage of a suppressor not present in Col-0 is diluted in the hybrids. This observation is also complicated by the fact that a considerable part of $MRD1^{Ler\text{-}0}$ is strongly methylated in the hybrids. In addition, there is always a basal level of *MRD1* transcription in the tissues tested regardless of the ploidy level. The data of the MPSS project (26) suggest that methylation and, in turn, activity of *MDR1*, could be influenced by small iRNA-linked mechanisms. In this context, it is worth it to mention that *MRD1* was found to be suppressed in the *mto1-1* mutant, which overaccumulates soluble methionine (23). Taken together, our observations indicate a complex control of *MRD1* transcription and it remains to be determined whether paramutation-like phenomena are involved.

Implications for Evolution and Plant Breeding. Significant changes in cellular morphology and physiology are known in allo- and autotetraploids (1–5, 15–18, 21). In the former, some trait changes have clearly been associated with gene expression alterations (17, 18). Similar effects are expected to occur in *Arabidopsis* autotetraploids. Here, we consider solely the potential of such

alterations in context of the evolution of these two forms of ploidy. The data on allotetraploids together with our observations open up alternative evolutionary scenarios for allo- vs. autopolyploids, which both exhibit equally stable chromosome segregation (3, 6). Allopolyploids and their homoploid progenitors could resort to numerous alterations in gene expression, allowing for rapid adaptations to extreme habitats. On the other hand, they might be prone to developmental accidents due to the interference of ploidy, heterosis, and effects that result from the reunion of divergent genomes (2, 4, 6, 7, 9, 41, 42). Neo-autopolyploids could resort to a lower and stably heritable number of ploidy-induced alterations allowing selective adaptations. In the long term, these processes might entail mutations that would act to fix such alterations (7), which would otherwise be lost. If so, this mechanism could appreciably impact the evolution of autopolyploids, together with other known mechanisms such as point mutations or genetic drift. Additional aspects complicate these considerations. First, autopolyploidy can occur recurrently (5, 10, 42). Second, autopolyploids could "feed" allopolyploid evolution. For instance, the generation of synthetic *A. suecica* allopolyploids was only possible through crosses of synthetic autotetraploid *A. thaliana* with *A. arenosa* because of the lethality of homoploid hybrids (15). Allopolyploids are taxonomically predominate, but a reliable estimate for the frequency of autopolyploid species is yet to be found. In fact, autopolyploids might be much more prevalent in nature than presently known (2, 3, 5, 6, 8, 9, 41), because they are sometimes difficult to recognize based on morphology. Our results support this notion and indicate that the success of autotetraploids might critically depend on the magnitude of a species' natural genetic variability. This observation could impact plant breeding because autopolyploidy might be much better exploited if the natural variability of a species is considered.

Materials and Methods

Plant Material. European Arabidopsis Stock Centre (Loughborough, UK) and *Arabidopsis* Biological Resource Center (Columbus, OH) provided *A. thaliana* ecotypes. We used established (21) or converted new ecotypes to tetraploids as described (21).

Gene Expression and Microarray Analysis. Protocols for isolation, purification, and storage of (c)RNA, (q)RT-PCR analysis, and (q)RT-PCR-primers can be found in *SI Materials and Methods* and Table S4. The *Arabidopsis* 60-mer OligoMicroarray Agilent 4 × 44K platform was used. Cy3/Cy5-two-color experiments comprised at least four biological replicates (Table S1). Seedling transcriptome analyses between diploid L*er*-0 vs. Col-0 and between tetraploid Col-0 vs. L*er*-0 lines revealed 860 and 348 ecotype specific differences, respectively. These and the other microarray data in this work are deposited at NCBI/GEO; accession no.: GSE18482.

Additional Experimental Procedures. A detailed description of experimental procedures including microarray, methylation, and qRT-PCR analysis as well as bioinformatics and statistics can be found in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank Lynette Fulton (and Farhah Assaad) for language editing (earlier version); Elena Torres Ruiz for additional corrections; Otti Peis for cultivating plants; Dagmar Engl, Stephan Kotschote and Kerstin Stegmüller (both GPC-Biotech), Monika Frey, Michael Pfaffl, Dirk Haller (Technische Universität München), Gert Daniel (Bayer Landesanstalt für Landwirtschaft, Freising, Germany), Jörg Durner, Uta von Rad, Tony Schaeffner (Helmholtz Zentrum), and Götz Frommer (Agilent) for their helpful support; our Department of Energy-Joint Genome Institute colleagues for the *A. lyrata* genome assembly; and an anonymous reviewer for valuable hints. We thank Deutsche Forschungsgemeinschaft for financial support (Grant GI140/12-1 to A.G. and R.A.T-R.).

1. Grant V (1971) *Plant Speciation* (Columbia Univ Press, New York).
2. Otto SP, Whitton J (2000) Polyploid incidence and evolution. *Annu Rev Genet* 34: 401–437.
3. Ramsey J, Schemske DW (2002) Neopolyploidy in flowering plants. *Annu Rev Ecol Syst* 33:589–639.
4. Comai L (2005) The advantages and disadvantages of being polyploid. *Nat Rev Genet* 6:836–846.
5. Soltis DE, Soltis PS, Tate JA (2003) Advances in the study of polyploidy since Plant speciation. *New Phytologist* 161:173–191.
6. Mallet J (2007) Hybrid speciation. *Nature* 446:279–283.
7. Osborn TC, et al. (2003) Understanding mechanisms of novel gene expression in polyploids. *Trends Genet* 19:141–147.
8. Darlington CD (1963) *Chromosome Botany and the Origins of Cultivated Plants.* (Hafner, New York), 2nd Ed.
9. Soltis PS, Soltis DE (2009) The role of hybridization in plant speciation. *Annu Rev Plant Biol* 60:561–588.
10. De Bodt S, Maere S, Van de Peer Y (2005) Genome duplication and the origin of angiosperms. *Trends Ecol Evol* 20:591–597.
11. Gaut BS, Doebley JF (1997) DNA sequence evidence for the segmental allotetraploid origin of maize. *Proc Natl Acad Sci USA* 94:6809–6814.
12. Kashkush K, Feldman M, Levy AA (2002) Gene loss, silencing and activation in a newly synthesized wheat allotetraploid. *Genetics* 160:1651–1659.
13. Adams KL, Percifield R, Wendel JF (2004) Organ-specific silencing of duplicated genes in a newly synthesized cotton allotetraploid. *Genetics* 168:2217–2226.
14. Riddle NC, Jiang H, An L, Doerge RW, Birchler JA (2010) Gene expression analysis at the intersection of ploidy and hybridity in maize. *Theor Appl Genet* 120:341–353.
15. Comai L, et al. (2000) Phenotypic instability and rapid gene silencing in newly formed *arabidopsis* allotetraploids. *Plant Cell* 12:1551–1568.
16. Wang J, et al. (2006) Genomewide nonadditive gene regulation in Arabidopsis allotetraploids. *Genetics* 172:507–517.
17. Wang J, Tian L, Lee HS, Chen ZJ (2006) Nonadditive regulation of FRI and FLC loci mediates flowering-time variation in Arabidopsis allopolyploids. *Genetics* 173: 965–974.
18. Ni Z, et al. (2009) Altered circadian rhythms regulate growth vigour in hybrids and allopolyploids. *Nature* 457:327–331.
19. Stupar RM, et al. (2007) Phenotypic and transcriptomic changes associated with potato autopolyploidization. *Genetics* 176:2055–2067.
20. Galitski T, Saldanha AJ, Styles CA, Lander ES, Fink GR (1999) Ploidy regulation of gene expression. *Science* 285:251–254.
21. Yu Z, Haage K, Streit VE, Gierl A, Ruiz RA (2009) A large number of tetraploid Arabidopsis thaliana lines, generated by a rapid strategy, reveal high stability of neo-tetraploids during consecutive generations. *Theor Appl Genet* 118:1107–1119.
22. Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature* 408:796–815.
23. Goto DB, Naito S (2002) AtMRD1 and AtMRU1, two novel genes with altered mRNA levels in the methionine over-accumulating mto1-1 mutant of Arabidopsis thaliana. *Plant Cell Physiol* 43:923–931.
24. Zhang X, et al. (2006) Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. *Cell* 126:1189–1201.
25. Mittelsten Scheid O, Afsar K, Paszkowski J (2003) Formation of stable epialleles and their paramutation-like interaction in tetraploid Arabidopsis *thaliana*. *Nat Genet* 34: 450–454.
26. Lu C, et al. (2005) Elucidation of the small RNA component of the transcriptome. *Science* 309:1567–1569.
27. Birchler JA, Riddle NC, Auger DL, Veitia RA (2005) Dosage balance in gene regulation: Biological implications. *Trends Genet* 21:219–226.
28. Huettel B, Kreil DP, Matzke M, Matzke AJM (2008) Effects of aneuploidy on genome structure, expression, and interphase organization in *Arabidopsis thaliana*. *PLoS Genet* 4:e1000226.
29. Henry IM, Dilkes BP, Comai L (2007) Genetic basis for dosage sensitivity in Arabidopsis thaliana. *PLOS Genet* 3 e70:0593–0602.
30. Etheridge N, Chen YF, Schaller GE (2005) Dissecting the ethylene pathway of Arabidopsis. *Brief Funct Genomics Proteomics* 3:372–381.
31. Benková E, et al. (2003) Local, efflux-dependent auxin gradients as a common module for plant organ formation. *Cell* 115:591–602.
32. Treml BS, et al. (2005) The gene ENHANCER OF PINOID controls cotyledon development in the Arabidopsis embryo. *Development* 132:4063–4074.
33. Vaughn MW, et al. (2007) Epigenetic natural variation in Arabidopsis thaliana. *PLoS Biol* 5:e174.
34. Misteli T (2007) Beyond the sequence: Cellular organization of genome function. *Cell* 128:787–800.
35. Guo M, Davis D, Birchler JA (1996) Dosage effects on gene expression in a maize ploidy series. *Genetics* 142:1349–1355.
36. Wang J, et al. (2004) Stochastic and epigenetic changes of gene expression in Arabidopsis polyploids. *Genetics* 167:1961–1973.
37. Teixeira FK, et al. (2009) A role for RNAi in the selective correction of DNA methylation defects. *Science* 323:1600–1604.
38. Zheng X, et al. (2008) ROS3 is an RNA-binding protein required for DNA demethylation in Arabidopsis. *Nature* 455:1259–1262.
39. Slotkin RK, et al. (2009) Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. *Cell* 136:461–472.
40. Chandler V, Alleman M (2008) Paramutation: Epigenetic instructions passed across generations. *Genetics* 178:1839–1844.
41. Rieseberg LH, Willis JH (2007) Plant speciation. *Science* 317:910–914.
42. Leitch AR, Leitch IJ (2008) Genomic plasticity and the diversity of polyploid plants. *Science* 320:481–483.

Theor Appl Genet (2009) 118:1107–1119
DOI 10.1007/s00122-009-0966-9

ORIGINAL PAPER

A large number of tetraploid *Arabidopsis thaliana* lines, generated by a rapid strategy, reveal high stability of neo-tetraploids during consecutive generations

Zheng Yu · Kristina Haage · Verena E. Streit · Alfons Gierl · Ramón A. Torres Ruiz

Received: 26 February 2008 / Accepted: 6 January 2009 / Published online: 11 February 2009

Abstract *Arabidopsis thaliana* has, in conjunction with *A. arenosa*, developed into a system for the molecular analysis of alloplolyploidy. However, there are very few *Arabidopsis* lines available to study autopolyploidy. In order to investigate polyploidy on a reliable basis, we have optimised conventional methodologies and developed a novel strategy for the rapid generation and identification of polyploids based on trichome branching patterns. The analysis of more than two dozen independently induced *Arabidopsis* lines has led to interesting observations concerning the relationship between cell size and ploidy levels and on the relative stability of tetraploidy in *Arabidopsis* over at least three consecutive generations. The most important finding of this work is that neo-tetraploid lines exhibit considerable stability through all the generations tested. The systematic generation of tetraploid collections through this strategy as well as the lines generated in this work will help to unravel the consequences of polyploidy, particularly tetraploidy, on the genome, on gene expression and on natural diversity in *Arabidopsis*.

Communicated by M. Kearsey.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-009-0966-9) contains supplementary material, which is available to authorized users.

Z. Yu · K. Haage · V. E. Streit · A. Gierl (✉) · R. A. Torres Ruiz (✉)
Lehrstuhl für Genetik, Wissenschaftszentrum Weihenstephan, Technische Universität München, Am Hochanger 8, 85350 Freising, Germany
e-mail: Ramon.Torres@wzw.tum.de

Present Address:
K. Haage
Lehrstuhl für Genetik, Department Biologie I, Ludwig-Maximilians-Universität München, Großhadener Str. 4, 82152 Planegg-Martinsried, Germany

Introduction

Polyploidy is a widespread phenomenon in animals and plants (Comai 2005). Estimates of the polyploid fraction amongst angiosperm species range from 30 to 80%. In fact, possibly all angiosperm species might be so called palaeopolyploids (Otto and Whitton 2000; Bennett 2004 and references therein). Polyploidy, particularly tetraploidy, has several potential advantages because the organisms can resort to a higher number of genes and higher maximum number of allelic variants. This is believed to be advantageous for plant metabolism in terms of elevated rates of synthesis or a higher variability of metabolically relevant compounds (Wolters and Visser 2000; Osborn et al. 2003), which may be one reason why certain tissues of diploid plants increase their genome content endogeneously by endopolyploidisation (Larkins et al. 2001). A further advantage lies in the subfunctionalization of gene copies that take over the task in different tissues (Adams and Wendel 2005a). Plant breeders have taken advantage of polyploidy in order to improve agronomic traits of economically important plants, some of which have been generated by allopolyploidisation of interspecies hybrids or by autopolyploidisation within species (Otto and Whitton 2000; Bennett 2004). Polyploid plants are generally known to exhibit a decrease in fertility and viability due to perturbations in meiotic segregation, mitotic division and/or perturbations in gene expression caused by altered gene dosage and gene

silencing (Comai 2005; Mittelsten-Scheid et al. 2003). It is therefore not clear whether a newly arisen polyploid plant will be evolutionarily successful or not.

An advanced molecular understanding of plant polyploidy will help us to ameliorate or circumvent detrimental effects and contribute to the improvement of plant breeding and agriculture. To this effect, *Arabidopsis thaliana* is increasingly becoming a model system for the molecular analysis of allopolyploidy (Madlung et al. 2002; Wang et al. 2004; Wang et al. 2006 and references therein). Although allopolyploidy is considered much more prevalent in nature, autopolyploids might be much more common than assumed (Soltis and Soltis 2000). Several studies using tetraploid lines obtained with diverse methods have been reported over the years (Redei 1964; Bouharmont 1965; Melaragno et al. 1993; Heslop-Harrison and Maluszynska 1994; Koornneef 1994; Altmann et al. 1994; Weiss and Maluszynska 2000; Mittelsten-Scheid et al. 2003; Santos et al. 2003). However, some of these lines no longer exist, some are unstable and others have undesirable traits due to EMS and X-ray treatment. At present, the stock centers provide six tetraploid, seven (telo)trisomic lines, one aneuploid and two natural tetraploid ecotypes (NASC catalogue http://arabidopsis.info/). Taken together, there are only a few reliable lines available for analysis. Consequently this limits the generalisation of observations when studying autopolyploidy. The identification of the molecular principles of tetraploidy will thus require the generation and analysis of several independent lines (Comai 2005). These will provide a reliable scientific basis when analysing epigenetic changes or exploring the potential of polyploidy in the context of natural diversity. Investigators of allopolyploidy in Brassicas have recognised this point and produced multiple colchicine-induced lines for analysis (Albertin et al. 2006; Lukens et al. 2006; Gaeta et al. 2007).

Although methods for the induction of tetraploids in *Arabidopsis* and other plants are known, the scarcity of available lines highlights the tedious nature of generating polyploids. The bottleneck lies not so much in the induction of polyploidy, as in the rapid identification and assessment of new lines. We have scanned a number of cellular traits that can act as markers for ploidy levels. This study presents a rapid and reliable procedure, based on the evaluation of trichome morphology, for the systematic establishment of collections of polyploid *Arabidopsis* lines. Our analysis is limited to the assessment of correct total chromosome numbers. However, in this respect it demonstrates a surprisingly high stability of tetraploidy in the lines generated and provides new data on the relationship between ploidy and cell size.

Materials and methods

Plant material and growth conditions

We obtained the following *Arabidopsis* ecotypes from the *Arabidopsis* stock centers *Arabidopsis* Biological Resource Center (ABRC) and Nottingham *Arabidopsis* Stock Center (NASC): Bor-1 (CS22590), Br-0 (CS22628), Bur-0 (CS22656), C24 (CS22620), CIBC-5 (CS22602), Cvi-0 (CS22614), Col-0 (N1092), Ct-1 (CS22639), Eden-1(CS22572), Kas-2 (CS22638), Kas-2 (N1264), L*er*-0 (NW20), L*er*-1 (CS22618), Mt-0 (CS22642), Nd-0 (N1390), Nd-1 (CS22619), Oy-0 (CS22658), Ra-0 (CS22632), Tamm2 (CS22604), Wa-1 (N1587, a natural tetraploid), Ws-2 (CS22659), Yo-0 (CS22624), selected RI-lines (see Table 1 and Supplementary Table 2) from the Col-0xL*er*-0 Recombinant Inbred Population set (N1899, Lister and Dean 1993) and the tetraploid lines N141, N3900, N3247, N3151, N3238, N3432.

Further lines were kindly provided by the following colleagues: ecotypes Mt-0 and Oy-0 by Thomas Debener (University of Hannover; these lines are distinguished from the same ecotypes from NASC/ABRC by the suffix TD and were used in our initial work), the diploid and tetraploid Zürich lines (Zü, Zü4x) and the repeatedly assessed tetraploid Col-line (N3432) by Ortrun Mittelsten-Scheid (GMI Wien), the tetraploid Wilna ecotype by Jolanta Maluszynska (University Silesia, Poland), the transgenic lines *CYCAt1*:CDB:GUS by D. Celenza (via M.-T. Hauser, University BOKU Wien) and $DR5_{rev}$::GFP by J. Friml (University Ghent).

Seeds were sown on soil (mixture of one third quartz sand and two thirds peaty mould) and grown under constant light (80–100 µmol photons/m^2 s), 40% relative humidity and 18°C in a Heareus (HEMZ 20/240/S) walk-in growth chamber.

Conversion of diploid into tetraploid plants

The first step in the generation and identification of polyploid plants is induction with colchicine. Colchicine concentrations are critical and had to be tested separately for every ecotype (Fig. 1). We also tested several procedures. A number of polyploid plants were generated by sterile submersion for 3–4 h in 0.5% colchicine solution, followed by an interim sterile culture on ½ Murashige-Skoog medium (ca. 1 week) and subsequently planted on soil. In a second treatment (modified after Santos et al. 2003), a drop of colchicine solution (~15 µl) was placed on the apex of young seedlings with less than five primary leaves (one-drop method). We tested several concentrations in combination with different numbers of treatments and found that a single treatment was sufficient for the generation

Table 1 Efficiency of colchicine treatments in L*er*-0 (NW20) and Col-0 (N1092) as estimated from "polyploid" sectors

Ecotype/stock-no., line designation	Colchicine conc. (%)/ no. of treatments	No. of seedlings	Survived seedlings (% of treated seedlings)	Seedlings with polyploid sectors (% of survived seedlings)
L*er*-0/NW20	0.5/4 h submerged	142	1 (0.7)	0 (0)
Col-0/N1092	0.5/4 h submerged	140	17 (12)	7 (41)
L*er*-0/NW20	0.5/9[a]	18	1 (5.5)	0 (0)
Col-0/DR5[c]	0.5/3[b]	11	9 (82)	6 (67)
Col/CYCAT1[d]	0.5/3[b]	11	7 (64)	5 (71)
Col-0/DR5[c]	0.1/3[b]	10	9 (90)	5 (56)
Col/CYCAT1[d]	0.1/3[b]	11	11 (100)	4 (36)
RI-lines group I[e]	0.5/3[b]	128	29 (22.6)	21 (72)
RI-lines group II[f]	0.1/1	242	151 (62.4)	128 (84.8)
Col-0/N1092	0.5/1	30	25 (83.3)	10 (40)
L*er*-0/NW20	0.5/1	150	36 (24)	7 (19.4)
L*er*-0/NW20	0.1/1	200	53 (26.5)	11 (20.8)
L*er*-0/NW20	0.05/1	40	12 (30)	3 (25)

[a] Application during 2 weeks on 9 days Tuesday–Friday and Monday–Friday

[b] Application during 1 week on 3 days Monday, Wednesday and Friday

[c] Full designation of the transgene: $DR5_{rev}$::GFP (Friml et al. 2003)

[d] Full designation of the transgene CYCAT1:CDB:GUS (Hauser and Bauer 2000)

[e] RI-lines from the Lister and Dean ColxL*er*-RI-population (NASC, Lister and Dean 1993): RI-13, -35, -191, -238, -263, -295, -302, -303, -332, -358, -367, -370 and -377 (see also supplementary Table 2)

[f] RI-lines from the Lister and Dean ColxL*er*-RI-population (NASC, Lister and Dean 1993): RI-13, -37, -115, -190, -191, -194, -217, -231, -232, -238, -242, -245, -263, -267, -283, -288, -303 and -367 (see also supplementary Table 2)

of tetra-/polyploid sectors. For "sensitive" ecotypes we reduced solutions to 0.1 or 0.05% (Table 1, Supplementary Table 1). The evaluation of cellular size, in particular trichomes, as explained in the text, enabled us to identify polyploid versus non-polyploid sectors on a treated plant. This allowed us to harvest plants producing polyploid seeds selectively, which in turn alleviated the identification of the desired tetraploids in the second round of analysis. In our hands, this treatment was the most efficient procedure and led predominantly to tetraploid plants. We did not harvest specific inflorescences and aimed to harvest polyploid plants (or sectors), which often also included diploid sectors. This selection led to progeny of mixed ploidies in the following generation. We therefore selected single plants from the first generation onwards.

Estimation of cell size and statistical analyses

We measured trichome cell size indirectly by counting the number of trichome branches on rosette leaves. For each ecotype and ploidy level, we classified the trichomes on the third leaf of five to ten plants (between 20 and 110 trichomes per plant). In some cases, we evaluated additional data sets from counts from the fifth leaf. For the analysis of stomatal cells, we did not perform epidermal lifts but took images instead. This enabled us to measure the stomatal size and to evaluate chloroplasts numbers. Thus, rosette leaves were fixed (30% EtOH, 5% acetic acid, 3.7% formaldehyde, 0.01% Triton X-100) overnight at 4°C and passed through an ethanol series (40, 50, 60, 70%; each step for ca. 45 min) at room temperature and stored in 70% at 4°C until use. Prior to analysis, the ethanol series was reversed down to 10% ethanol and the epidermis was analysed microscopically. Images of 30 randomly selected adaxial stomata (three plants, ten stomata each) were taken and cut out from the printed images (leaving out the stomatal opening). Their weights were used to estimate the actual surface area in μm^2 with the aid of a calibration curve.

For statistical analyses we used the GraphPad-PRISM® statistical software (version 4.0). All data sets were tested for normal (= Gaussian) distribution and where necessary transformed ($\exp x$, $\sin x$) to obtain normality. This allowed us to perform parametric tests of trichome and stomatal size data, which included *t* tests (and *F* tests) for comparisons of two data sets or ANOVA (one-way, two-way) and appropriate post-tests (Bonferroni, Tukey) for multiple comparisons, respectively (two-tailed analyses). The data for chloroplast number could not be adequately transformed to yield a Gaussian distribution. In this case we applied non-parametric tests (Mann–Whitney for comparison of two data sets, Kruskal–Wallis followed by Dunn's post-tests for multiple comparisons). The confidence interval applied to all tests was 95%.

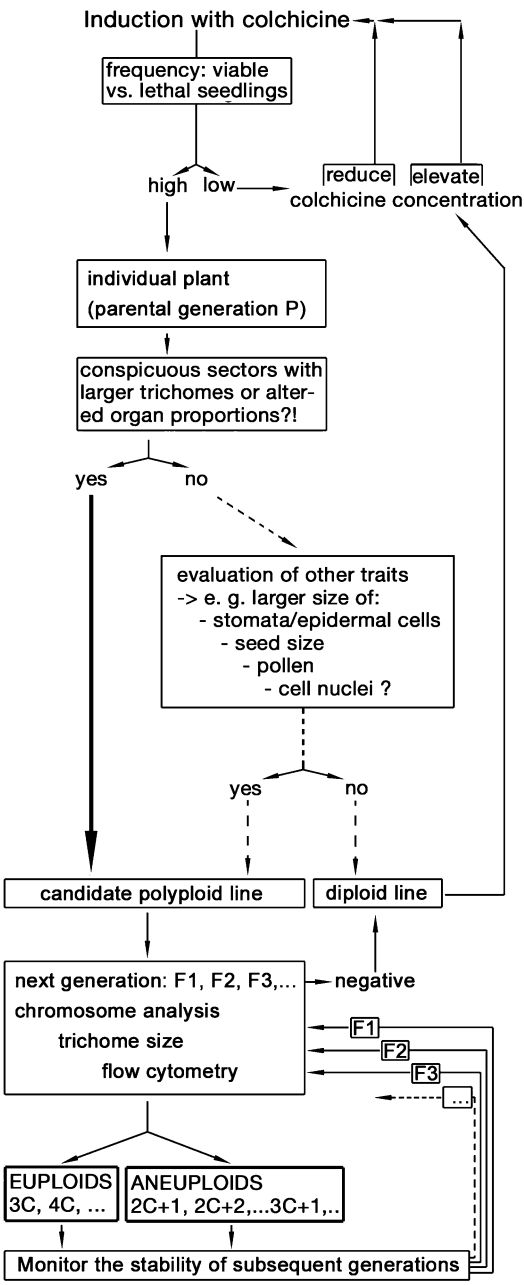


Fig. 1 Flow chart of the procedure for induction, identification and assessment of polyploids in *Arabidopsis thaliana* (for details see text). The big arrow directly leading from "conspicuous sectors…" to "candidate polyploid line" indicates the shortest and most convenient way to isolate the highest number of poly-/tetraploids with the least effort

Analysis of the polyploidy grade

We routinely analysed nuclei from rosette leaves from treated plants by flow cytometry. Where appropriate we analysed tissue from sectors of the colchicine treated plants beforehand in addition to the evaluation of trichome morphology. Flow cytometry was essentially performed as described (e.g. Henry et al. 2005) using the high-resolution kit from Partec. Briefly, leaves were chopped with a sharp razor blade in 0.3 ml of nuclei extraction buffer (solution A of the Partec kit) and filtered through 20 or 30 µm Cell Trics filters (Partec). The flow through was combined with 1.2 ml of DAPI solution (solution B of the Partec kit) and analysed in a PAS II flow cytometer (Partec) equipped with a HBO lamp for UV excitation. The PAS II FlowCytometer distributes the measured particles according to their fluorescence intensity into 1024 different channels. Routinely, several thousand particles were measured per leaf and measurements were often repeated. We compared the peak positions of the 2C, 4C, 8C, 16C and 32C nuclei between diploid and tetraploid plants whose ploidy was already known or had been assessed by chromosome counts of metaphases. Due to endopolyploidy, flow cytometry measurements from a diploid *Arabidopsis* plant not only exhibit a 2C peak but also further peaks up to 32C. Consequently, plants with a higher basic ploidy level e.g. tetraploids lack the 2C peak. However, flow cytometry is not sensitive enough to discriminate between certain euploids and aneuploids, e.g. 4C vs. 4C + 1. Bearing this in mind, the chromosome number of plants was assessed by counting metaphase chromosomes of root tips. These chromosome counts uncovered a number of Col-0 plants originating from the most intensive colchicine induction procedure, which displayed irregular mitotic chromosome figures (see "Results").

We followed standard protocols for metaphase chromosome preparation (e. g. Maluszynska and Heslop-Harrison 1991; Zhong et al. 1996) with slight modifications. Briefly, roots from seedlings grown on agar plates were dissected and incubated for 1 h at 4°C in 2 mM hydroxychinolin until the tissue was transferred to Carnoy's solution (ethanol:acetic acid; 3:1; v:v). Alternatively, roots were dissected and immediately submerged in 1 ml Carnoy's solution to fix the root tips for at least 12 h, washed three times for ca. 5 min with H_2O_{bidest} and incubated for at most 15 min at 37°C in enzyme solution (10% Macerozyme R-10 from Duchefa; 2% Cellulase "Onozuka R-10" from Duchefa in 100 mM Citric-buffer pH4.8). Roots were washed three times with H_2O_{bidest} and placed on small Petri dishes to separate the root tips (they often separate automatically upon digestion). About 10 root tips were transferred onto a slide (SuperFrost®Plus, Menzel), squashed with the tip of a needle and covered with a drop of 60% acetic acid for ca. 1 min. After adding 800 µl Carnoy's solution for 2 min the slide was submerged in 70% ethanol, air-dried and 60 µl DAPI solution (4′,6-diamidino-2-phenylindol-dihydrochloride, 1 µg/ml) was added. The slide was covered with a cover slip and stored for 10 min in the dark. Afterwards the cover slip and DAPI were washed off with H_2O_{bidest}, the slide was air-dried, covered with 10% glycerol and a new cover slip and analysed under the epifluorescence microscope. For each line and generation we evaluated between 5 and 20 metaphases.

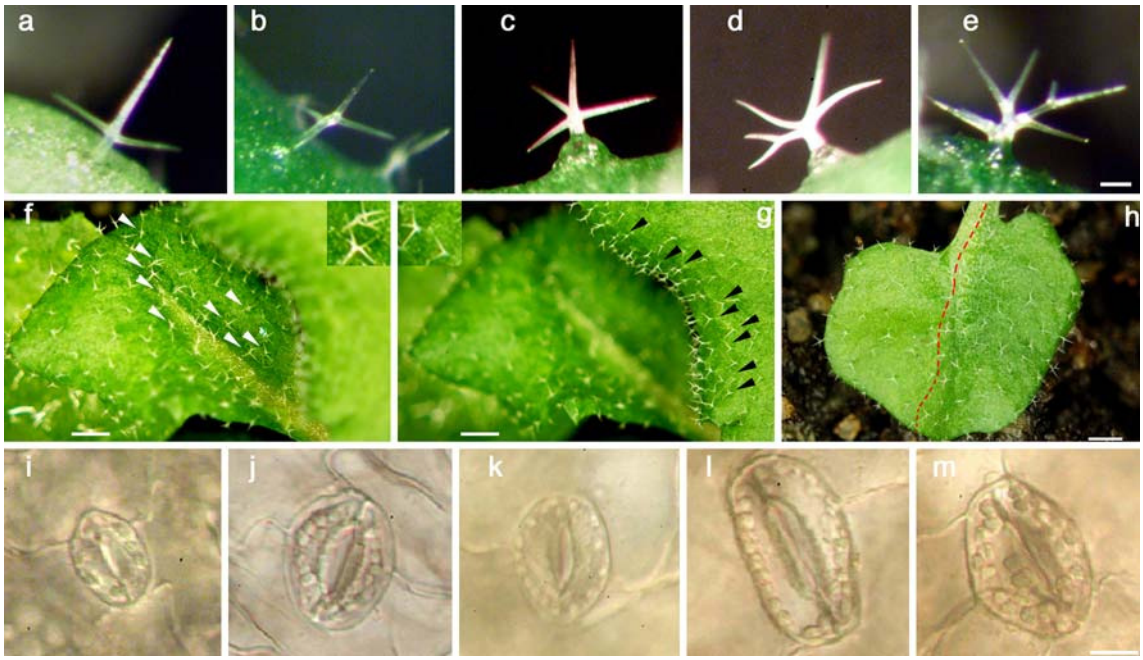


Fig. 2 Cell size and tissue size effects of polyploidy. Shown is the trichome morphology of diploid Col-0 (**a**), diploid L*er*-0 (**b**), tetraploid Col-0 (**c**), octoploid Col-0 (**d**) and aneuploid Col-0 (**e**) plants. Sectors of colchicine treated plants as visualized by trichome morphology and leaf morphology (**f**–**h**); (**f**) polyploid trichomes (white arrowheads) on a leaf of a sectored plant ecotype Yo-0 (inset shows a magnification of such trichomes), (**g**) a neighboured leaf with diploid trichomes (black arrowheads; inset shows a magnification); (**h**) sectored leaf of a treated CIBC-5 ecotype, diploid sector (left) separated by a stippled line from the polyploid sector (right). Sizes of stomata from diploid Col-0 (i), tetraploid Col-0 (j), octoploid Col-0 (k) and two different (4x + 1 and 6x + 2) aneuploid Col-0 plants, respectively (l + m). Note that there are elevated numbers of chloroplasts in the stomata of plants with higher ploidies (see also Fig. 7). Scale bars: 0.1 mm in **e** is same for **a**–**e**; 1 mm in **f**–**h**; 10 μm in **m** is same for **i**–**m**

Results

The generation of polyploids involves three major steps: induction with colchicine, identification of candidate lines (sectors) and assessment of polyploid lines by analysing the progeny of candidate lines (Fig. 1). The rationale of our procedure is based on the straightforward verification of the architecture of trichome cells as a morphological surface "marker", which discriminates between diploid and polyploid *Arabidopsis* sectors and lines, respectively (Fig. 2). The treated parental generation was examined with respect to the presence of polyploid sectors. Plants with polyploid sectors were always counted as potential parents for polyploid lines, because we observed that, if a polyploid sector developed reproductive organs, it regularly produced polyploid progeny. The identification of sectored plants therefore works as a short cut in our strategy (Fig. 1).

Induction via colchicine treatment

The initial tests started with high concentrations of colchicine and either extreme durations of exposition to this agent or high numbers of treatments. In these cases, most seedlings became necrotic and died. However, the data indicate that the tolerance to colchicine varies between ecotypes. For instance, 12% of Col-0 seedlings survived 4 h of submersion in 0.5% (w/v) colchicine giving one tetraploid line (named P9A) and six aneuploid lines, three of which are presented in this work solely for comparison (4x + 1, 6x + 1 and 7x[+ 3], respectively; see below). Less than 1% of L*er* survived (Table 1). We also tested this procedure with the ecotypes Kas-2 (N1264), Nd-0, Mt-0 (TD), and Oy-0 (TD), with comparable numbers (60–120) of seedlings (not shown). These showed survival frequencies between 0% (Mt-0 TD) and 7.5% (Nd-0). However, except with Col-0, we did not obtain any plants with polyploid sectors. The nine-fold treatment did not lead to polyploid sectors in L*er*-0 (Table 1) but also not in Mt-0 (TD) and Nd-0 (survival frequency of 45.8 and 20.7%, respectively; not shown). A three-fold or one-fold treatment (with different concentrations) always produced polyploids. Plants with Col-0 background are less sensitive and outrun L*er*-0 plants when one aims to obtain many polyploids. We have continued to treat diverse ecotypes with the one-drop method and concentrations of 0.5 and 0.1% colchicine, respectively, with variable results (Supplementary Table 1). For some ecotypes, it was not possible to obtain any plants with polyploid sectors so far. Interestingly, high survival frequencies do not guarantee a high frequency of polyploid

sectors (compare Col-0/DR5 and Col/CYCAT1 at 3x 0.5% with the same at 3x 0.1%). The result with RI-lines is conspicuous. They show low or moderate survival frequencies but very high frequencies of survival with polyploid sectors.

Sector analysis in colchicine treated plants

If a plant survives the colchicine treatment, it might remain unaffected and diploid. Alternatively, if colchicine has suppressed mitosis in a cell of the shoot apex a mosaic plant will grow. In this plant the root, hypocotyl, cotyledons and first primary leaves will be diploid.

Polyploidy often leads to an increase in cell size. We reasoned that an increase in the size of epidermal cells would be readily detectable and that this could act as a reliable indicator for polyploid sectors in colchicine treated plants. In fact, the morphology of trichomes turned out to be perfectly suited for this purpose. We found that the way trichomes react to a change of polyploidy in their sector is not just by an overall increase in size but also an increase in the number of branches (Fig. 2). This roughly correlated with the degree of polyploidy in progeny originating from sectors with altered versus non-altered trichomes. Sectors that were sufficiently large were also analysed by flow cytometry analysis. We assume that trichome cells that are already endopolyploid in diploid plants might have an even more elevated degree of endopolyploidy in polyploid plants. They at least had larger nuclei in the cases examined (SFig. 1 in supplementary material). The identification of polyploid sectors by surface markers alone was not a sufficient parameter for our strategy. Most important was whether epidermal (L1) sectors reflected polyploid tissues in the corresponding sub-epidermal layers (L2, L3) because these are the origins for (polyploid) gametes and progeny, respectively. This was regularly the case (see below).

Stability of newly induced tetraploid *Arabidopsis* plants

We analysed the relative stability of the newly induced tetraploids over three consecutive generations in about two dozen lines from the Col-0 and L*er*-0 ecotypes as schematically outlined in Fig. 3. We took advantage of the fact that an accumulation of multiple branched trichomes is a reliable marker for a sector/plant, which gives polyploid progeny in all (even the parental) generations. Single colchicine treated seedlings were given a code number, which henceforth identifies the corresponding line in our lab (Table 2; first column: ecotype; second column: line/code-no.). Plants, that survived and displayed polyploid sectors (as judged by trichome and organ morphology) were grown to maturity and harvested. Seeds of each line (F1 generation) were germinated and the root tips of ten young seedlings were taken for chromosome analysis (Fig. 4). The F1 generation might consist of different ploidies dependent on the size of the two competing sectors, diploid vs. polyploid, in the parent plant (Fig. 3). We evaluated the metaphase figures and calculated the average chromosome number for each line (Table 2, third column: F1chr.). The chromosome analysis of the F1 progeny indicated more or less polyploid progeny for virtually all plants, as indicated by the chromosome index (F1chr.-column in Table 2). Note that, due to the variable quality of the metaphase figures taken, the chromosome count does not precisely sum up to even numbers. This is also true for the control diploid and tetraploid plants. This is because it was sometimes difficult to distinguish all the chromosomes in the mitotic figures. We considered a value of ca. 19 together with a corresponding 4C flow cytometry value to indicate a tetraploid. Five seedlings of each line were further grown and inspected for trichome morphology (Fig. 3). In all except three cases, 5/5 seedlings showed polyploid trichomes and/or abnormally structured/sized organs. One of the five mature plants was selected for flow cytometry analysis (Table 2; fourth column: F1flc.). This step already separated lines with respect to their ploidy (F1flc-column in Table 2): these included three diploids, twelve tetraploids, one pentaploid, two hexaploids and two octoploids. The penta- and hexaploids were not followed through all generations. The hexa- and octoploids, respectively, exhibited instability through all generations and produced progeny of mixed ploidy (the former more than the latter). Plants with a basic ploidy level higher than diploid lacked the 2C peak. Plants with a basic ploidy level higher than tetraploid in addition lacked the 4C peak (Fig. 5). All the plants that turned out to be tetraploid or octoploid were harvested and further analysed (F2 generation). We proceeded as in the first generation, i.e. ca. 10 seedlings were taken for chromosome analysis (Table 2; fifth column: F2chr.), five seedlings were grown, evaluated for their trichomes and one of the five adult plants was harvested for flow cytometry (Table 2; sixth column: F2flc.) to give the progeny of the next generation and so on. In all but one instance, the tetraploid lines remained tetraploid. The octoploids exhibited chromosome number instability as deduced from their average number of metaphase chromosomes. All tetraploids of the F2 generation remained tetraploid in the next generation (Table 2, seventh and eighth column: F3chr. and F3flc., respectively).

By analysing metaphase chromosome figures, we also detected six aneuploid lines amongst the seven Col lines, which resulted from the most intensive colchicine induction procedure (see "Materials and methods"). The variability of chromosome numbers observed for these lines by far

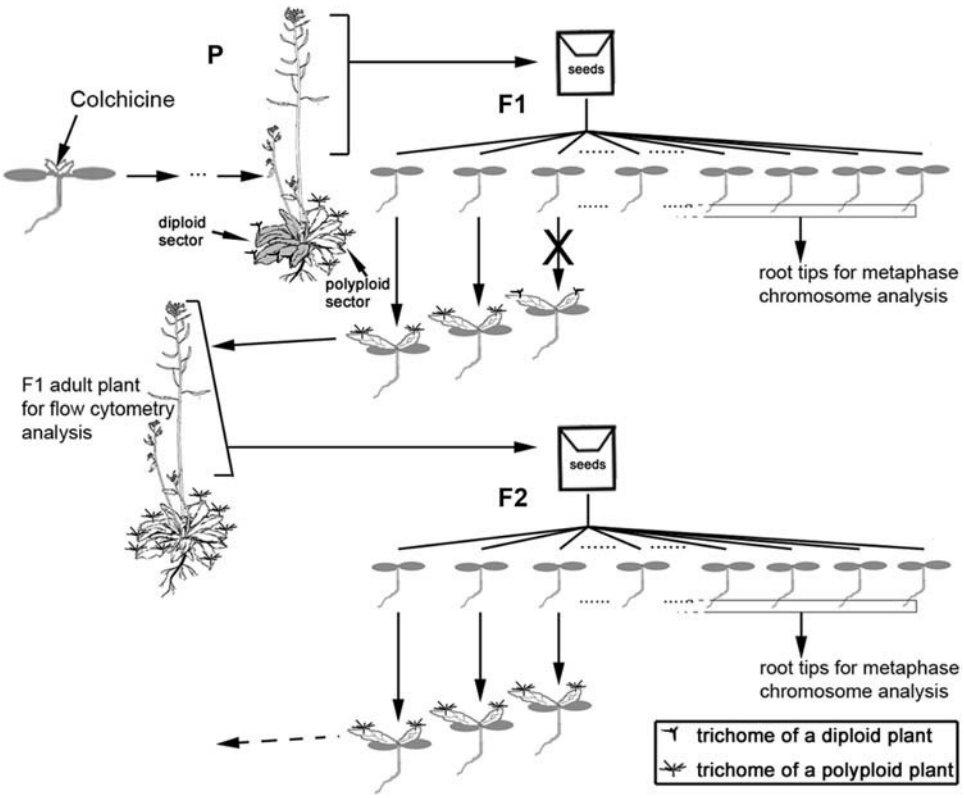


Fig. 3 Selection and assessment scheme for polyploid lines in consecutive generations (see text)

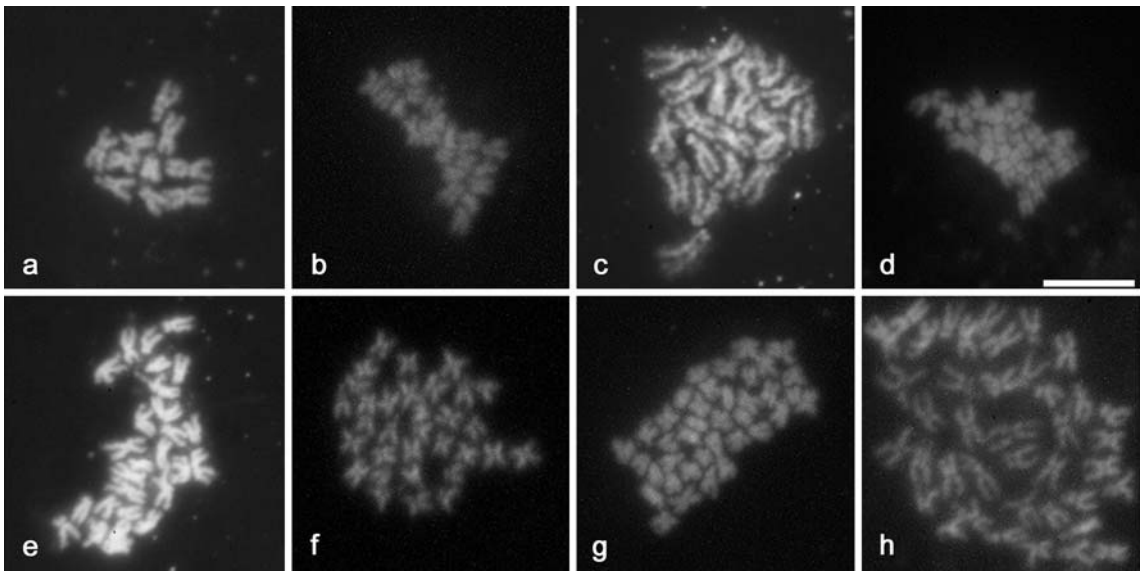


Fig. 4 Metaphase chromosomes of dividing root tip cells. Shown are metaphase chromosomes of different ploidies from the Col-0 ecotype (except the hexaploid, which is of mixed ColxLer background): **a** 2x, **b** and **c** 4x, **d** 4x + 1, **e** 6x, **f** 6x + 2, **g** 7x[+ 3] and **h** 8x. The chromosomes in **a**, **c**, **e** and **h** have not reached full condensation. Scale bar in **d** is same for **a**–**h**: 10 μm

exceeded that which results from differences in the quality of particular metaphases as mentioned above. We estimated the most likely chromosome numbers of these plants by evaluating metaphase figures of siblings of generation F2. Since this study focuses on tetraploids, we did not follow further generations of these plants. For a comparison of cell size features (see below), we selected siblings of the same generation from three of these six lines with an estimated chromosome number of 4x + 1, 6x + 1 and 7x[+ 3]. We suspect that the aneuploid constitution of these lines resulted from the strength of the colchicine induction procedure.

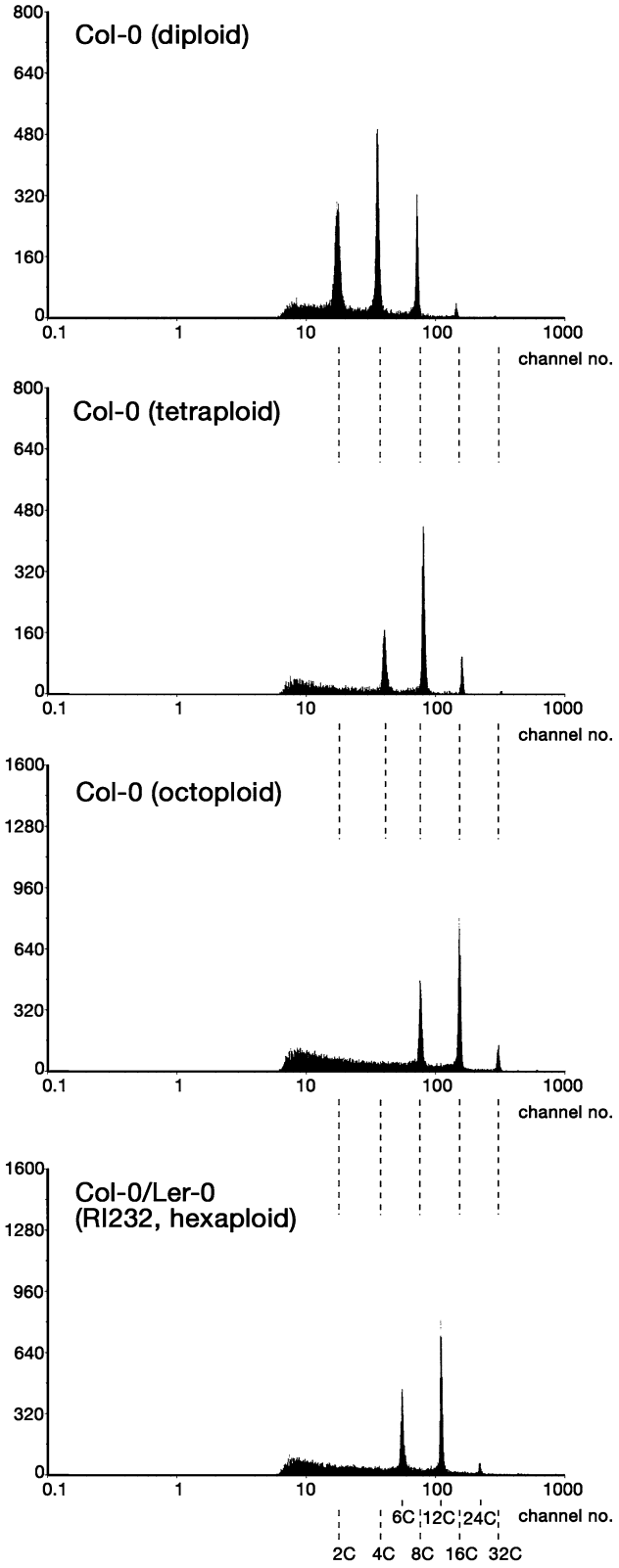


Fig. 5 Flow cytometry of *Arabidopsis* polyploids. The number of measured particle counts (nuclei) versus the channel of the Partec II analyser is given. Note the absence of the 2C peak in the tetraploid and the absence of the 2C + 4C peaks in the octoploid plant respectively. Note also the position of the peaks of the hexaploid plant

Cell size features of polyploid lines

The numerous lines assessed, spanning a large range of ploidies, gave us the opportunity to test a number of characters, in particular cell size, in more detail. This analysis revealed some known but also new interesting features. The effect of polyploidy in the epidermis is summarised by the average number of branches per trichome. The trichome index gives the proportion of a particular trichome branch class in the corresponding population (Fig. 6). The average branch number of a trichome starts at 3.00, 3.00 and 2.34 in diploid L*er*-0, Col-0 and Zü, respectively, continues with 4.04, 3.33 and 2.94 in tetraploid L*er*-0, Col-0 and Zü, respectively, and reaches 4.27 in hexaploid Col-0xL*er*-0 (RI232), 5.19 in octoploid Col-0 and 3.83, 4.51 and 4.52, respectively, in the aneuploids (4x + 1, 6x + 1, 7x[+ 3]). Plants with numerous six- and seven-branched trichomes are more likely hexa- or octoploid than tetra- or diploid, respectively, (Fig. 6). Interestingly, the higher the degree of ploidy, the more trichomes disperse along a higher number of trichome branch classes. Diploids have two- to four-branched trichomes. Tetraploids have predominantly two- to five-branched trichomes. Trichomes of hexa- and octoploids scatter along three to eight branched variants. Aneuploids resemble plants with higher ploidies in this respect, irrespective of their actual chromosome number (Fig. 6). Thus, an elevated number of trichome branches in comparison to the corresponding progenitor ecotype generally tends to indicate higher ploidy. The new lines we obtained allowed us to compare the impact of ploidy (2x vs. 4x) and the impact of ecotype (Col-0, Ler-0, Zü) simultaneously. We obtained significant values for the general impact of both factors on trichome branch number ($P < 0.001$, two-way ANOVA). Therefore, we separated the ecotypes subsequently and performed the analyses of ploidy effect within a given ecotype. This revealed significant differences between Ler 2x vs. Ler 4x and Zü 2x vs. Zü 4x, respectively, (*t* tests, $P < 0.0001$ for both). Within Col-0 multiple comparisons (one-way ANOVA) were performed. Euploids and aneuploids were also separately compared because analyses indicated differences in variance between these groups (e.g. Col-0 2x vs. Col-0 7x[+ 3] with $P < 0.0001$, *F* test). Comparison of Col-0 euploids displayed significant differences ($P < 0.001$) except that one comparison, Col-0 2x vs. Col-0 4x, was critical and gave non-significant ($P > 0.05$) and significant ($P < 0.01$) values depending on the transformation of data to obtain a Gaussian distribution (sin*x* vs. exp*x*). One-way ANOVA of aneuploids did not reveal significant differences. The hexaploid line was excluded from these comparisons since it is of mixed Ler-0/Col-0 origin.

Leaves having supernumerary trichome branches also had larger stomata (Fig. 7). This difference is significant

Table 2 Assessment of polyploidy grade of selected lines during three consecutive generations

Ecotype	Line code-no.	F1chr.	F1flc.	F2chr.	F2flc.	F3chr.	F3flc.
Col-0	N1092(diploid)[a]	9.7	2C	–	–	–	–
Col-0	N3432(tetrapl.)[b]	19.3	4C	–	–	–	–
Col-0	1326-12	18.2	4C	19.8	4C	19.8	4C
Col-0	1326-15	15.7	4C	20.1	4C	19.8	4C
Col-0	1326-18	14.9	2C	–	–	–	–
Col-0	1326-19	19.1	4C	19.6	4C	19.7	4C
Col-0	1326-26	19.7	4C	19.8	4C	19.6	4C
Col-0	1326-28	19.2	4C	19.3	4C	19.5	4C
Col-0	P9A	19	–	18.6	4C	–	4C
Col-0	3115-1	27.9	8C	19.2	4C/6C	23.8	4C
Col-0	3115-2	11.7	4C	9.2	2C	–	–
Col-0	3115-3	40.0	8C	33.9	8C	24.0	–
L*er*-0	1026-5	15.8	4C	20.1	4C	–	–
L*er*-0	1026-10	18.4	4C	19.1	4C	19.8	4C
L*er*-0	1026-19	11.8	2C	–	–	–	–
L*er*-0	1026-40	17.6	4C	19.5	4C	19.6	4C
L*er*-0	1026-41	16.8	4C	19.3	4C	20.0	4C
L*er*-0	1026-27	12.6	2C	–	–	–	–
L*er*-0	3116-1	17.8	6C	–	–	–	–
L*er*-0	3116-2	15.3	5C	–	–	–	–
L*er*-0	3116-6	20.0	4C	20.0	4C	19.6	4C
RI232-2		–	–	–	6C	28.6	6C

[a] Reference diploid line Col-0 N1092 used for comparison in all flow cytometry analyses

[b] Established and repeatedly assessed tetraploid Col-0 control line N3432 from Dr. O. Mittelsten-Scheid. For details see text

when di- and tetraploids are compared ($P < 0.0001$). The three analysed aneuploid lines did not reveal significant differences in multiple comparisons, which also indicated that a higher chromosome number does not always necessarily lead to a bigger cell size. Similarly, diploid Col-0 displayed a significant difference in the number of chloroplasts per stomata in comparison to tetraploid Col-0 ($P < 0.0001$). Aneuploids revealed a difference only with respect to Col-0 4x + 1 vs. Col-0 6x + 1 ($P < 0.05$). In random samples, we monitored other features such as size of cell nuclei, roots and seeds (SFig. 1). In all cases, these structures had larger sizes in polyploids in comparison to diploids. With respect to roots, the increased size of these organs is due to increase of cell size instead of cell number (not shown).

Discussion

Colchicine: effects and induction of polyploidy

The most important aspect of our colchicine treatments pertains to colchicine effects on the epidermal (L1) and sub-epidermal (L2, L3) layers. Following treatment, the apex can develop in three different ways. First, it might produce detectable diploid and polyploid sectors (Fig. 2). Second, diploid cells overgrow polyploid cells and most of the plant turns diploid again and produces diploid progeny. Conversely, polyploid cells overgrow diploid cells; i.e. polyploid progenies are mainly produced. In any case, it is a crucial prerequisite of our strategy that the altered polyploidy of the epidermis (L1 layer) on the treated plant reflects an altered polyploidy of inner cell layers (L2 and L3 layer, respectively) because these are the cells that give rise to generative cells. This is not self-evident because L1 and the inner layers (L2, L3) are developmentally separated early on (Duckett et al. 1994; Takada and Jürgens 2007). In fact, we found that once a polyploid sector had been identified on a plant, the same plant regularly delivered polyploid progeny (Table 2). This can be explained by two possible effects, which have been observed previously (Dawe and Freeling 1991; Tilney-Bassett 1986 and references therein). The first is that colchicine intruded into epidermal and subepidermal layers in most of the treatments. The second is that cells from one tissue "invaded" the adjacent lineage.

Colchicine is a highly poisonous secondary metabolite from *Colchicum autumnale* (and other Colchiceae), which inhibits microtubule polymerization by binding to tubulin. This study reveals ecotype and line differences in sensitivity to colchicine (some RI and transgenic lines; Table 1, Supplementary Tables 1 and 2). Whether the observed differences in sensitivity have a genetic basis, controlling

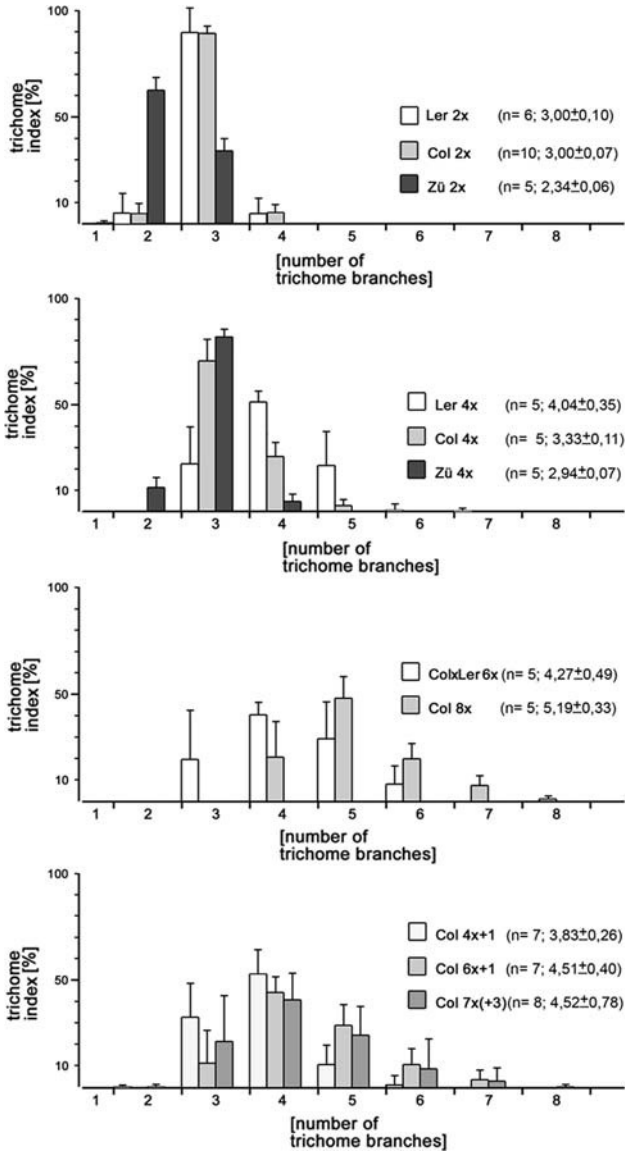


Fig. 6 The relation of basic polyploidy and branch numbers in *Arabidopsis* trichomes. The plots separate the data for different ploidy classes, i.e. diploid, teraploid, hexa plus octoploid and aneuploid respectively. Ecotypes and (an-)euploidies are indicated with different grey scales. The number of analysed plants, the mean of branches per trichome of all plants and their standard deviations are given on the right

processes such as uptake and detoxification of colchicine, remains to be determined.

Whether colchicine has mutagenic effects in addition to its toxic effects in *Arabidopsis* or other plants is not clear. In our experiments, we did not detect any of the mutants that most frequently occur in mutagenic screens (i.e. the albino, crème and fusca phenotypes; Jürgens et al. 1991; Mayer et al. 1991). Similarly, an independent study of over 40 colchicine induced *Brassica* allopolyploids did not provide any indication of a mutagenic effect of colchicine (Lukens et al. 2006). There is also no evidence for long term genomic instabilities caused by colchicine. In fact, these could be caused by other factors, such as scaling incompatibilities of cell volume versus intracellular structures (e.g. spindle geometry; Storchova et al. 2006).

Size effects in generated polyploids

When we examined the plants after colchicine induction we visually separated wild-type (diploid) plants from plants with putative polyploid sectors. These were defined as sectors carrying a conspicuously high number of trichomes with more than three branches (more than two in case of ecotype Zü). Actually, the subsequent analyses of the pedigree revealed that the trichome branch number tends to increase with polyploidy until it reaches a certain limit (Fig. 6). From an analysis of triploids, it is known that trichomes of such plants have values between diploids and tetraploids (Perazza et al. 1999). The statistical analyses of the di- and tetraploids demonstrate a general impact of both the genetic background (i.e. ecotype) and the grade of polyploidy. This does not exclude the possibility that some comparisons do not exhibit a difference (e. g. Col-0 vs. L*er*-0 diploid plants). A tendency to increase the number of trichome branches with increasing polyploidy grade is also given (despite particular exceptions) in the larger Col-0 polyploid series. Another interesting observation concerns the scattering of trichomes in different trichome branch classes. Thus, octo-, hexa- and aneuploid trichomes have four to seven trichome branch classes, as compared to three to four classes in diploids and tetraploids. This probably indicates that a plant with a high ploidy, e.g. an octoploid plant, can realize a broader range of ploidies in its cells. It has been shown in *Arabidopsis* and other plants that many cells have elevated ploidy levels due to endoreduplication (Galbraith et al. 1991; Schmuths et al. 2004). In addition, in aneuploids, cells might adopt variable ploidies due to enhanced genome instability. Very similar effects might influence the size of stomata (and in turn the number of chloroplasts they harbour). In diploids and tetraploids the difference is significant. Aneuploids display some variability in this respect but often the differences are not significant although they differ in chromosome number. Taken together, we could verify all known size effects (Melaragno et al. 1993; Perazza et al. 1999 and references therein) in polyploids, i.e. size increase of (epidermal) cells, organs and seeds as well as cell nuclei (Figs. 2, 6, 7 and SFig. 1).

Relative stability of neo-tetraploid lines

In this study, we focussed on the analysis of absolute chromosome numbers in tetraploid lines during three consecutive

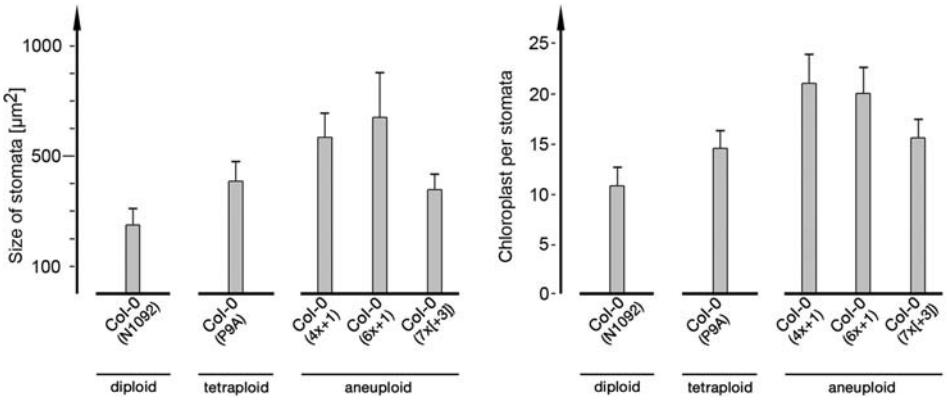


Fig. 7 Stomata in polyploids. The sizes of stomata (both cells) for corresponding lines are given in μm^2 (left). The number of chloroplasts per stomata from the same lines as in the left part is given in the right part. The corresponding lines and standard deviations are indicated

generations. We have addressed neither aspects concerning the correct number of particular chromosomes nor epigenetic variation.

The detected tetraploid lines displayed a high degree of stability in chromosome number because, once they had passed the selection in F1 they remained stable. There was only one exception, Col-0 line 3115-2. The selected F1 plant only produced diploid progeny (note the near-to-diploid F1 chromosome index; Table 2). Such instabilities are also known for established lines. We observed this in the tetraploid lines: Wilna ecotype, N3247 and Wa-1 (N1587), which is a natural tetraploid (Henry et al. 2005; Schmuths et al. 2004). Interestingly, L*er*-0 seems to be more colchicine-sensitive than Col-0. However, once established L*er*-0 tetraploids seem to be as stable as Col-0 tetraploids. The hexa- and octoploids produced progeny of mixed ploidy (the former more than the latter). They could only be maintained through successive generations by selection of true hexaploids and octoploids in the pedigree by means of flow cytometry analysis. Though we have not analysed these in this detail, we found the aneuploids isolated from the experiment with the strongest colchicine induction treatment (see "Materials and methods") to be very variable with respect to genome composition. The chromosome analysis revealed an approximate estimation of their chromosome number. It is likely that this composition will vary in the next generations and lead to divergent lines in this respect.

The relative stability of tetraploids has been previously observed, indicating that they can be easily maintained through selfing (Bouharmont 1965; Heslop-Harrison and Maluszynska 1994). The ability to produce progeny of the same ploidy requires some degree of diploidisation or chromosomal stabilisation in meiosis. In a tetraploid plant, this is acquired by a high frequency of bivalents versus a low frequency of uni-, tri- and tetravalents. Two studies report different but significant degrees of multivalent formation in *Arabidopsis* tetraploids. Both used different ecotypes and lines established for quite a different number of generations (Weiss and Maluszynska 2000; Santos et al. 2003). In addition, the Wilna line had undergone some chromosomal rearrangements in the rRNA gene clusters, probably promoting bivalent formation i.e. chromosomal stability (Weiss and Maluszynska 2000). Santos et al. (2003) demonstrated the importance of the history of a line because established lines had acquired a higher degree of diploidisation than newly induced lines (note that one of the lines analysed in this work N3427, identical to CS3427, was found to be unstable in our analyses, see above). The observed stability of our tetraploid lines seems to be in contrast with the aberrant meiotic products reported in previous work (Weiss and Maluszynska 2000; Santos et al. 2003). According to those studies, our new tetraploid lines should have produced mixed progeny with chromosome indexes significantly deviating from 20. However, the chromosome indexes of all tetraploids evaluated in F3 lie between 19.5 and 20.0, which indicates a homogeneous pool of tetraploid seedlings. There are at least two possible explanations for these discrepancies. The first is that unbalanced chromosomes had a greater impact on the viability of gametes in our lines. This "filtering" effect excluded aberrant gametes and was not detected since we did not analyse meiotic stages. A second possibility is that similar (or more extensive) rearrangements such as those observed by Weiss and Maluszynska (2000) promoted bivalent formation and chromosome stability. These could also not be detected by our analysis. Interestingly, rapid and early rearrangements seem to be common in re-synthesized *Brassica* allopolyploids (Song et al. 1995). Besides *Arabidopsis* and *Brassica* auto- and allopolyploids, respectively, other (allo- or amphidiploid) species such as *Triticum* and *Aegilops* are known to undergo rapid rearrangements including gene loss (Kashkush et al. 2002; Adams and Wendel 2005b and references therein). However, this is not a general rule as shown by cotton (Gossypium) species (Liu et al. 2001). Considering the current data in *Arabidopsis*, additional aspects are important in this context. For instance, our procedure selected for tetraploidy in every new generation.

Furthermore, Santos et al. (2003), and Weis and Maluszynska (2000) analysed meiotic stages but did not analyse the progeny during consecutive generations of a larger number of independently induced lines.

In conclusion, our data show that the strategy we have developed to augment the resources of tetra-/polyploid *Arabidopsis* lines works efficiently. The data presented indicate that the generated tetraploids exhibit a considerable stability in successive generations.

Acknowledgements We thank Dr. Daniel and his group (Bayerische Landesanstalt für Landwirtschaft) for generous provision of the Partec II Flow Cytometer and help, Dr. Stephan Haug (Lehrstuhl für Mathematische Statistik, TU München) for helpful advice in statistics and an anonymous reviewer for useful hints. We owe very special thanks to Farhah Assaad for critically reading the manuscript and useful discussions. We thank the Nottingham *Arabidopsis* Stock Centre (NASC) for lines. We are particularly indebted to the following colleagues for kindly providing lines: Ortrun Mittelsten-Scheid (Wien), Jolanta Maluszynska (Katowice), Marie-Theres Hauser (Wien), Thomas Debener (Hannover) and Jiri Friml (Ghent). The financial support by the Deutsche Forschungsgemeinschaft (Grant GI 140/12-1 to A.G. and R.A.T.R.) is gratefully acknowledged.

References

Adams KL, Wendel JF (2005a) Novel patterns of gene expression in polyploidy plants. Tends Genet 21:539–543

Adams KL, Wendel JF (2005b) Polyploidy and genome evolution in plants. Curr Opin Plant Biol 8:135–141

Albertin W, Balliau T, Brabant P, Chevre AM, Eber F, Malosse C, Thiellement H (2006) Numerous and rapid nonstochastic modifications of gene products in newly synthesized *Brassica napus* allotetraploids. Genetics 73:1101–1113

Altmann T, Damm B, Frommer WB, Martin T, Morris PC, Schweizer D, Willmitzer L, Schmidt R (1994) Easy determination of ploidy level in *Arabidopsis thaliana* plants by means of pollen size measurement. Plant Cell Rep 13:652–656

Bennett MD (2004) Perspectives on polyploidy in plants—ancient and neo. Biol J Lin Soc 82:411–423

Bouharmont J (1965) Fertility studies in polyploid *Arabidopsis thaliana*. In: Röbbelen G (ed) *Arabidopsis* research. Wasmund, Gelsenkirchen, pp 31–36

Comai L (2005) The advantages and disadvantages of being polyploid. Nat Rev Genet 6:836–846

Dawe RK, Freeling M (1991) Cell lineage and its consequences in higher plants. Plant J 1:3–8

Duckett CM, Oparka KJ, Prior DAM, Dolan L, Roberts K (1994) Dye-coupling in the root epidermis of *Arabidopsis* is progressively reduced during development. Development 120:3247–3255

Friml J, Vieten A, Sauer M, Weijers D, Schwarz H, Hamann T, Offringa R, Jurgens G (2003) Efflux-dependent auxin gradients establish the apical-basal axis of *Arabidopsis*. Nature 426:147–153

Gaeta RT, Pires JC, Iniguez-Luy F, Leon E, Osborn TC (2007) Genomic changes in resynthesized *Brassica napus* and their effect on gene expression and phenotype. Plant Cell 19:3403–3417

Galbraith DW, Harkins KR, Knapp S (1991) Systemic endopolyploidy in *Arabidopsis thaliana*. Plant Physiol 96:985–989

Hauser MT, Bauer E (2000) Histochemical analysis of root meristem activity in *Arabidopsis thaliana* using a cyclin: GUS (ß-glucuronidase) marker line. Plant Soil 226:1–10

Henry IM, Dilkes BP, Young K, Watson B, Wu H, Comai L (2005) Aneuploidy and genetic variation in the *Arabidopsis thaliana* triploid response. Genetics 170:1979–1988

Heslop-Harrison JS, Maluszynska J (1994) Molecular cytogenetics of *Arabidopsis*. In: Meyerowitz EM, Somerville CR (eds) *Arabidopsis*. Cold Spring Harbor Laboratory Press, New York, pp 63–88

Jürgens G, Mayer U, Torres Ruiz RA, Berleth T, Misera S (1991) Genetic analysis of pattern formation in *Arabidopsis* embryo. Dev Suppl 1:27–38

Kashkush K, Feldman M, Levy AA (2002) Gene loss, silencing and activation in a newly synthesized wheat allotetraploid. Genetics 160:1651–1659

Koornneef M (1994) *Arabidopsis* genetics. In: Meyerowitz EM, Somerville CR (eds) *Arabidopsis*. Cold Spring Harbor Laboratory Press, New York, pp 89–120

Larkins BA, Dilkes BP, Dante RA, Coelho CM, Woo YM, Liu Y (2001) Investigating the hows and whys of DNA endoreduplication. J Exp Bot 52:183–192

Lister C, Dean C (1993) Recombinant inbred lines for mapping RFLP and phenotypic markers in *Arabidopsis thaliana*. Plant J 4:745–750

Liu B, Brubaker CL, Mergeai G, Cronn RC, Wendel JF (2001) Polyploid formation in cotton is not accompanied by rapid genomic changes. Genome 44:321–330

Lukens LN, Pires JC, Leon E, Vogelzang R, Oslach L, Osborn T (2006) Patterns of sequence loss and cytosine methylation within a population of newly resynthesized *Brassica napus* allopolyploids. Plant Physiol 140:336–348

Madlung A, Masuelli RW, Watson B, Reynolds SH, Davison J, Comai L (2002) Remodeling of DNA methylation and phenotypic and transcriptional changes in synthetic *Arabidopsis* allotetraploids. Plant Physiol 129:733–746

Maluszynska J, Heslop-Harrison JS (1991) Localization of tandemly repeated DNA sequences in *Arabidopsis thaliana*. Plant J 1:159–166

Mayer U, Torres Ruiz RA, Berleth T, Misera S, Jürgens G (1991) Mutations affecting body organization in the *Arabidopsis* embryo. Nature 353:402–407

Melaragno JE, Mehrotra B, Coleman AW (1993) Relationship between endopolyploidy and cell size in epidermal tissue of *Arabidopsis*. Plant Cell 5:1661–1668

Mittelsten-Scheid O, Afsar K, Paszkowski J (2003) Formation of stable epialleles and their paramutation-like interaction in tetraploid *Arabidopsis thaliana*. Nat Genet 34:450–454

Osborn TC, Pires JC, Birchler JA, Auger DL, Chen ZJ, Lee HS, Comai L, Madlung A, Doerge RW, Colot V, Martienssen RA (2003) Understanding mechanisms of novel gene expression in polyploids. Trends Genet 19:141–147

Otto SP, Whitton J (2000) Polyploid incidence and evolution. Annu Rev Genet 34:401–437

Perazza D, Herzog M, Hülskamp M, Brown S, Dorne AM, Bonneville JM (1999) Trichome cell growth in *Arabidopsis thaliana* can be derepressed by mutations in at least five genes. Genetics 152:461–476

Redei GP (1964) Crossing experiences with polyploids. Arabidopsis Inf Serv 1:13

Santos JL, Alfaro D, Sanchez-Moran E, Armstrong SJ, Franklin FC, Jones GH (2003) Partial diploidization of meiosis in autotetraploid *Arabidopsis thaliana*. Genetics 165:1533–1540

Schmuths H, Meister A, Horres R, Bachmann K (2004) Genome size variation among accessions of *Arabidopsis thaliana*. Ann Bot 93:317–321

Soltis PS, Soltis DE (2000) The role of genetic and genomic attributes in the success of polyploids. Proc Natl Acad Sci USA 97:7051–7057

Song K, Lu P, Tang K, Osborn TC (1995) Rapid genome change in synthetic polyploids of Brassica and its implications for polyploid evolution. Proc Natl Acad Sci USA 92:7719–7723

Storchova Z, Breneman A, Cande J, Dunn J, Burbank K, O'Toole E, Pellman D (2006) Genome-wide genetic analysis of polyploidy in yeast. Nature 443:541–547

Takada S, Jürgens G (2007) Transcriptional regulation of epidermal cell fate in the *Arabidopsis* embryo. Development 134:1141–1150

Tilney-Bassett RA (1986) Plant chimeras. Edward Arnold Ltd, Baltimore

Wang J, Tian L, Madlung A, Lee HS, Chen M, Lee JJ, Watson B, Kagochi T, Comai L, Chen ZJ (2004) Stochastic and epigenetic changes of gene expression in *Arabidopsis* polyploids. Genetics 167:1961–1973

Wang J, Tian L, Lee HS, Wie NE, Jiang H, Watson B, Madlung A, Osborn TC, Doerge RW, Comai L, Chen ZJ (2006) Genomewide nonadditive gene regulation in *Arabidopsis* tetraploids. Genetics 172:507–517

Weiss H, Maluszynska J (2000) Chromosomal rearrangement in autotetraploid plants of *Arabidopsis thaliana*. Hereditas 133:255–261

Wolters AM, Visser RG (2000) Gene silencing in potato: allelic differences and effect of ploidy. Plant Mol Biol 43:377–386

Zhong XB, de Jong JH, Zabel P (1996) Preparation of tomato meiotic-pachytene and mitotic metaphase chromosomes suitable for fluorescence in situ hybridization (FISH). Chromosom Res 4:24–28