

Dynamics of chromatin modifications and other nuclear features in response to intraspecific hybridization in *Arabidopsis thaliana*

Dissertation

*zur Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat.)*

vorgelegt der
Mathematisch-Naturwissenschaftlich-Technischen Fakultät
(mathematisch-naturwissenschaftlicher Bereich)
der Martin-Luther-Universität Halle-Wittenberg

von Herr Ali Mohammad Banaei Moghaddam
geb. am 24.08.1979 in Yazd, Iran

verteidigt am
19.01.2010
begutachtet von:

Herr Prof. Dr. Klaus Humbeck [Halle]
Herr Prof. Dr. Ulrich Wobus [Gatersleben]
Herr Prof. Dr. Neil R. Jones [Aberystwyth, UK]

Halle (Saale)

ACKNOWLEDGEMENT

This work was started in June 2005 at the Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben. It was funded by Deutsche Forschungsgemeinschaft (DFG).

I am very grateful to my supervisor Dr. Andreas Houben for giving me the opportunity to work in his group Chromosome Structure and Function (CSF), continuous support and guidance, permanent encouragement and invaluable effort in correcting this thesis.

I thank my co-supervisor Dr. Michael Florian Mette, for the great help, support, fruitful discussions and comments during the work and invaluable efforts in correcting this thesis.

I would like to give many thanks to Prof. Vincent Colot for giving me the opportunity to perform ChIP on chip experiments in his laboratory and Dr. François Roudier for his patience and great support during my staying. I thank Dr. Tim Sharbel, for supporting me with his group facilities to perform MSAP analysis.

I wish to express my gratitude to Dr. Marc Strickert and especially his PhD. student Michael Seifert for support in the analyzing of *Arabidopsis* whole genome NimbleGen array data.

Many thanks to Dr. Richard Pickering for helpful suggestions and valuable help in correcting my thesis.

I thank Dr. Jörg Fuchs, for his help by doing flowcytometry analysis.

I thank Prof. Thomas Altmann for making available unpublished gene-expression data.

I would like to appreciate technicians Oda Weiss, Katrin Kumke and Margit Hantschmann for their perfect technical assistance.

Thanks to all members of CSF group for their helpful discussions and creating a nice scientific atmosphere in the lab.

Special thanks also to Raheleh for continuous support, encouragement and helpful suggestions.

بادود و سپاس بیکران این پایان نامه را به مادر عزیز و مهربانم، پدر بزرگوارم، برادر و خواهران گرامیم، همراه و
همسرم راحله به خاطر از خودگذشتگی و مهربانی بیدریغش، به روان اهورایی تحسین آموزگارم شادروان رستم که به
من خواندن و نوشتن آموخت و به جناب آقای طراز تحسین آموزگار زبان انگلیسیم همراه با آرزوی تندرستی
تقدیم می‌کنم.

Content

1.	Introduction.....	16
1.1.	Effects associated with the processes of inter- and intraspecific hybridization.....	16
1.1.1.	Genetic (DNA sequence-based) changes in allopolyploid formation.....	18
1.1.2.	Allopolyploidization can be accompanied by reactivation of transposable elements.....	19
1.1.3.	Transposon activation can happen in interspecific hybrids.....	20
1.1.4.	Changes in gene expression patterns due to allopolyploidization.....	20
1.1.5.	Hybridization-induced DNA and histone modifications.....	21
1.1.5.1.	Epigenetics.....	21
1.1.5.1.1.	DNA methylation.....	22
1.1.5.1.2.	Post-translational histone modifications.....	24
1.1.5.1.3.	Histone methylation.....	25
1.1.5.2.	DNA methylation changes in allopolyploids.....	26
1.1.6.	Activation and inactivation of rRNA genes in allopolyploid species.....	27
1.2.	Heterosis.....	28
2.	Aims of the thesis.....	32
2.1.	Is endopolyploidization altered by intraspecific hybridization?.....	32
2.2.	Is rRNA gene activity altered in intraspecific hybrids?.....	32
2.3.	Does microscopic analysis detect altered DNA methylation, histone H3K4me2, H3K27me3, or H3K9me2 in intraspecific hybrids?.....	32
2.4.	Do DNA methylation patterns change in response to intraspecific hybridization?..	33
2.5.	Is DNA methylation in 5' regulatory regions of selected genes altered in intraspecific hybrids?.....	33
2.6.	Are allele-specific transcript levels of selected genes altered in intraspecific hybrids?.....	33
2.7.	Is intraspecific hybridization accompanied by establishment of new histone modification patterns?.....	34
2.7.1.	Comparative genomic hybridization (CGH) analysis of different <i>A. thaliana</i> accessions.....	34

2.7.2.	ChIP on chip comparison of histone modifications in parental inbred lines and their hybrid offspring.....	34
3.	Material and methods.....	35
3.1.	Plant material.....	35
3.1.1.	Production of hybrid seeds for <i>A. thaliana</i> accessions Col-0, C24 and Cvi.....	35
3.1.2.	Cultivation of <i>A. thaliana</i> seedlings in soil.....	35
3.1.3.	Soil-free cultivation of <i>A. thaliana</i> seedlings for chromatin extraction.....	36
3.2.	Determination of the endopolyploidization and flow sorting of nuclei.....	36
3.3.	Silver staining of size-fractionated nuclei for nucleolus measurement.....	36
3.4.	Indirect immunostaining of histone modifications in sorted nuclei.....	37
3.4.1.	Immunostaining for histone H3K4, H3K9 and H3K27 methylation.....	37
3.4.2.	Immunostaining for 5-methylcytosine.....	37
3.5.	Extraction of genomic DNA and restriction cleavage.....	38
3.6.	Methylation sensitive amplified polymorphism (MSAP) assay.....	38
3.7.	Primers.....	41
3.7.1.	Design of primers.....	41
3.7.2.	Calculation of PCR efficiency.....	41
3.8.	Cloning and sequencing of C24 alleles of selected genes.....	41
3.9.	Allele-specific DNA methylation assay.....	42
3.10.	Extraction of plant RNA and cDNA synthesis.....	42
3.11.	Transcript determination.....	45
3.11.1	Quantitative RT-PCR.....	45
3.11.2	Allele-specific semiquantitative RT-PCR.....	46
3.12.	Quantitative PCR analysis of DNA from immunoprecipitated chromatin.....	48
3.13.	Comparative genomic hybridization (CGH) assay.....	49
3.13.1.	CGH experiment using chromosome 4 tiling microarrays.....	49
3.13.2.	CGH experiment using <i>Arabidopsis</i> whole genome NimbleGen arrays.....	50
3.14.	ChIP on chip experiments.....	50
3.14.1	ChIP on chip experiments using chromosome 4 tiling microarrays.....	53
3.14.2.	ChIP on chip experiments using <i>Arabidopsis</i> whole genome NimbleGen arrays....	53
3.15.	Validation of ChIP on chip results via quantitative ChIP-PCR.....	53

4.	Results.....	55
4.1.	The level of endoreduplication and nucleolus size does not differ between inbred parents and their intraspecific hybrids.....	55
4.2.	The nuclear distribution of DNA methylation and histone H3-methylation at positions K4, K9 and K27 in inbred parents and intraspecific hybrids is similar.....	57
4.3.	A limited gain of DNA methylation appeared in hybrid offspring compared to inbred parents.....	57
4.4.	No changes in allele-specific methylation of 5' regulatory regions of selected genes were detected in intraspecific hybrid offspring.....	61
4.5.	Expression of exemplary genes is additive in intraspecific hybrid offspring.....	64
4.6.	Allele-specific transcript levels of exemplary genes are unaltered in intraspecific hybrid offspring.....	65
4.7.	Comparative genomic hybridization (CGH).....	68
4.7.1.	CGH revealed sequence polymorphisms between <i>A. thaliana</i> accessions Col-0 and C24.....	70
4.7.1.1.	Chromosome 4 tiling microarrays.....	70
4.7.1.2.	<i>Arabidopsis</i> whole genome NimbleGen arrays.....	74
4.7.2.	CGH revealed sequence polymorphisms between <i>A. thaliana</i> accessions Col-0 and Cvi.....	79
4.8.	Analysis of histone H3K4me2 and H3K27me3 distribution between Col-0, C24, Cvi, and their intraspecific hybrids.....	81
4.8.1.	Natural differences in the distribution of histone H3K4me2 along chromosome 4 of Col-0 and C24.....	82
4.8.2.	Histone H3K4me2 is a stable modification which is additively inherited in intraspecific hybrids between <i>A. thaliana</i> accessions Col-0 and C24.....	86
4.8.3.	Natural differences in the distribution of histone H3K4me2 between <i>A. thaliana</i> accessions Col-0 and Cvi.....	89
4.8.4.	Histone H3K4me2 shows stability in response to intraspecific hybridization between <i>A. thaliana</i> accessions Col-0 and Cvi.....	91
4.8.5.	Natural differences in the distribution of histone H3K27me3 between <i>A. thaliana</i> accessions Col-0 and C24.....	93

4.8.6.	Histone H3K27me3 shows dynamics in response to intraspecific hybridization between Col-0 and C24.....	96
4.8.7.	Natural differences in the distribution of histone H3K27me3 between <i>A. thaliana</i> accessions Col-0 and Cvi.....	99
4.8.8.	Histone H3K27me3 shows dynamics in response to intraspecific hybridization between Col-0 and Cvi.....	101
4.8.9.	H3K27me3 shows higher dynamics in response to intraspecific hybridization than H3K4me2.....	103
5.	Discussion.....	106
5.1.	Intraspecific hybridization in <i>A. thaliana</i> does not change the endopolyploidy level.....	106
5.2.	Nucleolus size is not altered by intraspecific hybridization.....	106
5.3.	Global distribution of chromatin modifications is not altered by intraspecific hybridization.....	107
5.4.	Intraspecific hybrid formation is accompanied by a slight increase of DNA methylation.....	108
5.5.	DNA methylation and transcript levels of single genes were largely unaffected by intraspecific hybrid formation.....	110
5.6.	Sequence polymorphisms exist among different accessions of <i>A. thaliana</i>	111
5.6.1.	Comparative genomic hybridization is a powerful technique for detecting sequence polymorphisms.....	111
5.6.2.	Genes related to disease resistance and transposable elements show a high level of sequence polymorphisms between different accessions of <i>A. thaliana</i>	112
5.6.3.	<i>A. thaliana</i> accession Cvi shows slightly more sequence polymorphism than C24 compared with Col-0.....	114
5.7.	Variations in the levels of chromatin modifications marks exist among different accessions of <i>A. thaliana</i>	114
5.8.	Histone H3K4me2 is a stable histone modification mark in response to intraspecific hybridization.....	117
5.9.	Histone H3K27me3 is more dynamic than H3K4me2 in response to intraspecific hybridization.....	118

5.10.	Significance of findings with regard to models explaining the mechanism of heterosis.....	119
6.	Outlook.....	120
6.1.	Confirmation of histone H3K27me3 level changes detected by ChIP on chip analysis in hybrid offspring by ChIP-quantitative PCR.....	120
6.2.	Expression analysis of genes colocalizing with H3K27me3 hybridization responsive tiles.....	120
6.3.	Analysis of the heritability of H3K27me3 level changes induced by intraspecific hybridization.....	120
6.4.	Analysis of the activity of transposable elements in inbred parental lines and their hybrid offspring.....	121
6.5.	Systematically analyze selected regions of the <i>A. thaliana</i> genome to determine the changes in the level of different histone modifications and their effects on gene expression levels in response to intraspecific hybridization.....	121
7.	Summary.....	122
8.	Zusammenfassung.....	125
9.	Literature.....	128
	Appendix.....	138
	Appendix Figure 1 Fragmentation of genomic DNA by sonication.....	138
	Appendix Figure 2 Fragmentation of chromatin by sonication.....	138
	Appendix Figure 3 Efficiency of IP experiments for H3K4me2 and H3K27me3 on chromatin from Col-0, Col-0xC24, C24xCol-0, C24, Col-0xCvi and Cvi.....	139
	Appendix Figure 4 Ploidy levels at different developmental stages in inbred parental lines and their interspecific hybrid offsprings.....	140
	Appendix Table 1 Number of nuclei of defined ploidy for each genotype at different developmental stages.....	141
	Appendix Table 2 Number of nuclei for which nucleolus areas were Determined.....	142
	Appendix Table 3 Candidate tails covering sequences with significantly higher or lower level of H3K4me2 in hybrid offspring	

compared to their inbred parents based on chromosome 4 tiling DNA microarray.....	142
Publications related to the submitted thesis.....	143
Eidesstattliche Erklärung.....	144
CURRICULUM VITAE.....	145

Data-CD

- Excel File 1.** Genes coinciding with CGH polymorphic tiles Col-0_v_C24 chromosome 4 array
- Excel File 2.** Genes coinciding with CGH polymorphic tiles Col-0_v_C24 NimbleGen array
- Excel File 3.** Genes coinciding with CGH polymorphic tiles Col-0_v_Cvi NimbleGen array
- Excel File 4.** Genes coinciding with H3K4me2 polymorphic tiles Col-0_v_C24 chromosome 4 array
- Excel File 5.** Genes coinciding with H3K4me2 polymorphic tiles Col-0_v_Cvi Rep 1 NimbleGen array
- Excel File 6.** Genes coinciding with H3K4me2 polymorphic tiles Col-0_v_Cvi Rep 2 NimbleGen array
- Excel File 7.** Genes coinciding with H3K27me3 polymorphic tiles Col-0_v_C24 Rep 1 NimbleGen array
- Excel File 8.** Genes coinciding with H3K27me3 polymorphic tiles Col-0_v_C24 Rep 2 NimbleGen array
- Excel File 9.** Genes coinciding with H3K27me3 hybridization responsive tiles in Col-0xC24 combination of Re1 and rep2
- Excel File 10.** Genes coinciding with H3K27me3 polymorphic tiles Col-0_v_Cvi Rep 1 NimbleGen array
- Excel File 11.** Genes coinciding with H3K27me3 polymorphic tiles Col-0_v_Cvi Rep 2 NimbleGen array
- Excel File 12.** Genes coinciding with H3K27me3 hybridization responsive tiles in Col-0xCvi combination of Rep1 and Rep2

Abbreviations

ATP	<u>a</u> denosine <u>t</u> ri- <u>p</u> hosphate
AgNOR	argyrophilic nuclear organizer region
BLAST	<u>b</u> asic <u>l</u> ocal <u>a</u> lignment <u>s</u> earch <u>t</u> ool
bp	<u>b</u> ase <u>p</u> air
BSA	<u>b</u> ovine <u>s</u> erum <u>a</u> lbumin
cDNA	<u>c</u> omplementary <u>d</u> eoxyribo <u>n</u> ucleic <u>a</u> cid
CGH	<u>c</u> omparative <u>g</u> enomic <u>h</u> ybridization
Col-0	<u>C</u> olumbia- <u>0</u>
Cvi	<u>C</u> ape <u>V</u> erde <u>I</u> slands
DAPI	4',6- <u>d</u> iamidino-2-phenyl <u>i</u> ndole
DNA	<u>d</u> eoxyribo <u>n</u> ucleic <u>a</u> cid
DTT	<u>d</u> ithio <u>t</u> hreitol
EDTA	<u>e</u> thylenediaminetetraacetic <u>a</u> cid
FITC	<u>f</u> luorescein <u>i</u> sothiocy <u>a</u> nate
g	gravity (centrifugal force, $g = 9.8 \text{ m/s}^2$)
GTP	guanosine 5'- <u>t</u> riphosphate
h	hour
H3	histone H3
H4	histone H4
HCl	hydrochloric acid
InDels	<u>i</u> nsersions and <u>d</u> eletions
kb	<u>k</u> ilo <u>b</u> ase
Ler	<u>L</u> andsberg <u>e</u> recta
min	<u>m</u> inute
Mg	magnesium
mL	milliliter
mM	<u>m</u> illimolar
mRNA	messenger ribonucleic acid
MS medium	<u>M</u> urashige and <u>S</u> koog medium

MSAP	<u>m</u> ethylation <u>s</u> ensitive <u>a</u> mplified <u>p</u> olymorphism
NaCl	sodium chloride
NaOH	sodium hydroxide
ng	nanogram
NOR	<u>n</u> ucleolar <u>o</u> rganizer <u>r</u> egions
PBS	phosphate buffered saline
PCR	polymerase <u>c</u> hain <u>r</u> eaction
PIs	protease inhibitor
PMSF	phenyl <u>m</u> ethanesulfonyl <u>f</u> luoride
QTL	quantitative <u>t</u> rait <u>l</u> oci analysis
RFLP	<u>r</u> estriction <u>f</u> ragment <u>l</u> ength polymorphism
RNA	<u>r</u> ibon <u>u</u> cleic <u>a</u> cid
RNAi	RNA <u>i</u> nterference
rRNA	<u>r</u> ibosomal RNA
rpm	<u>r</u> evolutions per <u>m</u> inute
RT-PCR	<u>r</u> everse <u>t</u> ranscription - <u>p</u> olymerase <u>c</u> hain <u>r</u> eaction
s	<u>s</u> econd
SDS	<u>s</u> odium <u>d</u> odecyl <u>s</u> ulfate
siRNA	<u>s</u> mall <u>i</u> nterfering RNA
SNP	<u>s</u> ingle <u>n</u> ucleotide polymorphism
SSC	<u>s</u> aline- <u>s</u> odium <u>c</u> itrate
TAE	Tris-acetate-EDTA buffer
TBE	Tris-borate-EDTA buffer
TE	<u>t</u> ransposable <u>e</u> lement
Tris	tris(hydroxymethyl)aminomethane
UTR	<u>u</u> n <u>t</u> ranslated <u>r</u> egion
WT	<u>w</u> ild type
°C	centigrade
μE	microeinstein
μg	microgram
μL	microlitre
μm	micrometer

μm^2	square micrometer
1BR	first biological replicate
2BR	second biological replicate
5mC	5 methylcytosine

List of figures

Figure 1	Models for the formation of autopolyploids.....	17
Figure 2	Models for the formation of allopolyploids.....	17
Figure 3	Genetic models of heterosis.....	30
Figure 4	Endopolyploidization levels in inbred parents and their hybrid offspring.....	56
Figure 5	Nucleolus size in inbred parents and their hybrid offspring.....	56
Figure 6	Nuclear distribution of histone H3-methylation at positions K4, K9 and K27 and DNA methylation.....	58
Figure 7	Methylation sensitive amplified polymorphism (MSAP) analysis of inbred parents and their hybrid offspring.....	60
Figure 8	Semiquantitative allele-specific analysis of methylation in 5' regulatory regions of exemplary genes in inbred parents and their hybrid offspring.....	63
Figure 9	Relative transcript levels of exemplary genes obtained by quantitative RT-PCR.....	65
Figure 10	Semiquantitative allele-specific analysis of transcript levels of selected genes in inbred parents and their hybrid offspring.....	67
Figure 11	Distribution of CGH detected sequence polymorphisms between Col-0 and C24 on chromosome 4.....	72
Figure 12	Length distribution of domains on chromosome 4 containing contiguous CGH polymorphic tiles between accessions Col-0 and C24.....	73
Figure 13	Sequence type classification of CGH polymorphic tiles on chromosome 4 between Col-0 and C24.....	73
Figure 14	Size distributions of CGH polymorphic domains between Col-0, and C24 or Cvi.....	75
Figure 15	Sequence type classifications of CGH polymorphic tiles of Col-0 versus C24 and Cvi. using <i>Arabidopsis</i> whole genome NimbleGen arrays.....	76
Figure 16	Ontology classification of genes localized in CGH polymorphic regions.....	77
Figure 17	CGH polymorphic regions of chromosome 4 detected with different hybridization array platforms.....	78
Figure 18	Identification of CGH polymorphic regions (e.g. Col-0 versus Cvi) in dye-swap experiments is highly reproducible.....	79
Figure 19	Tiles with significant signal ratios in CGH and H3K4me2 ChIP on chip experiments.....	83
Figure 20	H3K4me2 polymorphisms between Col-0 and C24 along <i>A. thaliana</i> chromosome 4.....	85
Figure 21	Size distribution of domains of contiguous tiles polymorphic for H3K4me2 levels in Col-0 compared with C24.....	85
Figure 22	Sequence type classification of H3K4me2 polymorphic tiles in C24 compared with Col-0 along <i>A. thaliana</i> chromosome 4.....	86
Figure 23	Intersection analysis of tiles showing differences in H3K4me2 in direct	

	comparison between Col-0 / C24, Col-0 / Col-0xC24, and Col-0xC24 / C24.....	88
Figure 24	ChIP-quantitative PCR results for candidate regions showing responsiveness of H3K4me2 in intraspecific hybrids in Chip on chip analysis.....	89
Figure 25	Sizes of genomic regions with higher (in red) or lower level (in green) of histone H3K4me2 in Cvi versus Col-0.....	90
Figure 26	Class distribution of sequences covered by tiles with significant H3K4me2 differences between Col-0 and Cvi.....	91
Figure 27	Intersection analysis of tiles showing differences in H3K4me2 in direct comparison between Cvi / Col-0, Col-0xCvi / Col-0, and Col-0xCvi / Cvi.....	93
Figure 28	Sizes of regions with higher (in red) or lower levels (in green) of histone H3K27me3 in C24 versus Col-0.....	94
Figure 29	Classifications of tiles covering sequences with differences in H3K27me3 levels between Col-0 and C24.....	95
Figure 30	Intersection analysis of tiles showing differences in H3K27me3 in direct comparison between C24 / Col-0, Col-0xC24 / Col-0, and Col-0xC24 / C24.....	97
Figure 31	Ontology classification of genes localized in H3K27me3 hybridization responsive regions common to both biological replicates for Col-0 versus C24.....	98
Figure 32	Sizes of regions with higher (in red) or lower levels (in green) of histone H3K27me3 in Cvi versus Col-0.....	99
Figure 33	Classifications of tiles covering sequences with differences in H3K27me3 levels between Col-0 and Cvi.....	100
Figure 34	Intersection analysis of tiles showing differences in H3K27me3 in direct comparison between Cvi / Col-0, Col-0xCvi / Col-0, and Col-0xCvi / Cvi.....	102
Figure 35	Ontology classification of genes localized in H3K27me3 hybridization responsive regions common to both biological replicates for Col-0 versus Cvi.....	104

List of tables

Table 1	Adapters used in MSAP assays.....	39
Table 2	Pre-selective primers used in MSAP assays.....	40
Table 3	Selective primers in MSAP assays.....	40
Table 4	Primer combinations used in MSAP assays.....	40
Table 5	Primers used for amplification of candidate genes from accession C24.....	43
Table 6	Restriction sites for allele-specific DNA methylation analysis.....	44
Table 7	Primers for allele-specific DNA methylation assays.....	45
Table 8	Primers used for transcript analysis of candidate genes by quantitative and semiquantitative RT-PCR.....	46
Table 9	Primers used for controls in transcript analysis by quantitative and semiquantitative RT-PCR.....	47
Table 10	Candidate genes and restriction sites for allele-specific gene expression analysis....	47
Table 11	Primers for IP efficiency determination and quantitative PCR of reference regions in ChIP-PCR.....	48
Table 12	Primers for amplification of immunoprecipitated DNA for microarray hybridization.....	52
Table 13	Primers for quantitative PCR examination of ChIP on chip results from chromosome 4 tiling microarrays.....	54
Table 14	Summary of 5' regulatory region methylation analysis.....	64
Table 15	Summary of allele-specific expression assays.....	68
Table 16	Terms for the interpreting CGH and ChIP on chip results.....	69
Table 17	Design of CGH experiments to compare inbred parental lines.....	70
Table 18	Number of domains with at least two contiguous CGH polymorphic tiles in CGH experiment using chromosome 4 tiling microarrays.....	72
Table 19	Design of ChIP on chip experiments.....	84
Table 20	Domains of contiguous H3K4me2 polymorphic tiles between C24 and Col-0.....	85

1. Introduction

1.1. Effects associated with the processes of inter- and intraspecific hybridization

Interbreeding among different species (interspecific hybridization) creates hybrid individuals, which bring together different genetic material from isolated gene pools. Recent data document that interspecific hybridization widely occurs among plants and animals (Schwenk et al., 2008). The dissimilarity among parental chromosomes in a diploid hybrid issues a serious challenge to meiotic chromosome pairing and can cause hybrid sterility or dramatic reduction in fertility (Hegarty and Hiscock, 2008). One solution to overcome this chromosome incompatibility and restore fertility in hybrids is polyploidization.

“Polyploidy” refers to the heritable condition in which a diploid organism has more than two complete chromosome sets (Comai, 2005). Different types of polyploidy exist. In autopolyploids, multiple sets of chromosomes originated from the same species (Fig. 1). In contrast, allopolyploids contain combinations of two or more divergent sets of chromosomes that came together through the process of interspecific hybridization (Fig. 2). Duplication of each parental genome before (Fig. 2C) or after (Fig. 2A) interspecific hybridization (allopolyploidy) ensures that each chromosome pairs correctly with a homologous partner during meiosis. In addition, many organisms possess endopolyploid somatic cells that do not contribute to sexual reproduction. Endopolyploidy refers to polyploid cells whose chromosome number has been increased by endomitosis and for which the degree of ploidy is proportional to the number of endomitoses which have taken place. It seems that endopolyploidization is a response to physiological and developmental changes of an organism (Chen and Ni, 2006).

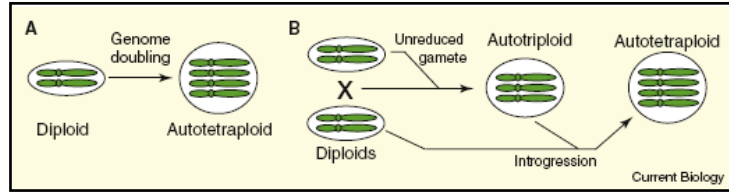


Figure 1 Models for the formation of autopolyploids.

A: Formation of autopolyploids by spontaneous chromosome duplication or fusion of two unreduced ($2n$) gametes within a single species. B: Fusion of an unreduced gamete ($2n$) with a normally reduced gamete within a single species produces an autotriploid. If this triploid can produce viable triploid gametes, hybridization with a normally reduced gamete will result in an autotetraploid (copied from Hegarty and Hiscock, 2008).

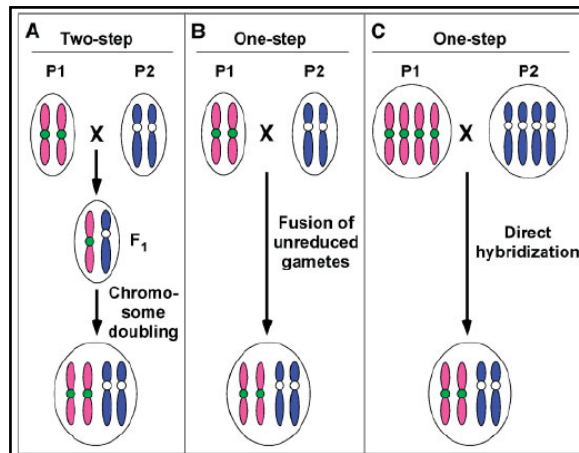


Figure 2 Models for the formation of allopolyploids.

A: According to this two-step model hybridization of two divergent genomes is followed by chromosome doubling. For simplicity each diploid genome is depicted to contain just one pair of chromosomes. P_1 and P_2 are representative of two progenitors. F_1 refers to transient diploid hybrid. B: In this one-step model allopolyploids are produced by the fusion of an unreduced male and female gamete. C: Another one-step model indicating how allopolyploids could arise immediately after hybridization between two autopolyploid species (copied from Chen and Ni, 2006).

The total number of chromosome sets in polyploid individuals is indicated by a prefix; for instance, tri-(3), tetra-(4), penta-(5), hexa-(6) and octa-(8). Polyploidy (whether allopolyploidy, autopolyploidy or endopolyploidy) has been found in all eukaryotes with particular abundance in flowering plants including many important agricultural crops. Even species with a small

genome like *Arabidopsis thaliana* appear to have undergone polyploidization in their evolutionary history (The *Arabidopsis* Genome Initiative, 2000).

In addition to the balancing of chromosomes in meiosis, further theories have been proposed to explain the ubiquity and success of polyploids. For instance, increased heterozygosity and hybrid vigour, decreased inbreeding depression due to masking of deleterious alleles via redundant favourable alleles and increased genetic diversity have been discussed. Redundancy in allopolyploids would release duplicate genes from natural selection pressure and thereby permit them to achieve new functions, which can result in better adaptability (Hegarty and Hiscock, 2008).

Artificial “synthetic” allopolyploids are usually phenotypically and genotypically less stable than their natural counterparts. As, in contrast to many natural allopolyploids, their progenitors are known and available, they provide excellent model systems for the study of genomic events associated with the early stages of genome hybridization and polyploidization. Genetic (that is, DNA sequence-based) changes would include translocations and transpositions, sequence deletions and insertions. Epigenetic (that is, not DNA sequence-based) changes would comprise, silencing, sub-functional expression of homoeologous genes, and transcription of previously silent transposons. A polyploid organism may experience some, but not all of these changes (Osborn et al., 2003; Chen and Ni, 2006).

1.1.1. Genetic (DNA sequence-based) changes in allopolyploid formation

Genomic rearrangement associated with allopolyploidization is a general phenomenon in the plant kingdom (Ma and Gustafson, 2005). Rapid genome rearrangements have been reported for a number of allopolyploid plants. Restriction fragment length polymorphism (RFLP) analysis revealed rapid genomic reorganization in synthetic *Brassica* allotetraploids. These changes included the absence of parental fragments as well as the presence of novel ones in synthetic allotetraploids. It was shown that the frequency of genomic changes in allotetraploids of *Brassica* is associated with the degree of divergence between the diploid parental genomes (Song et al., 1995). By assaying the occurrence of (sub) genome- and chromosome-specific low-copy sequences in tetraploid and hexaploid wheat, newly generated synthetic hexaploid

wheat, and its respective diploid progenitors, a non-random elimination of sequences in the process of allopolyploid formation was demonstrated (Feldman et al., 1997). It has also been shown that some specific genomic regions such as rDNA loci underwent rapid rearrangements in synthetic *Arabidopsis* allotetraploids (Pontes et al., 2004). Possible roles of DNA recombination and repair in the maintenance of genomic stability during the early stages of polyploidy formation were proposed in a study on synthetic *Arabidopsis* auto- and allotetraploids (Wang et al., 2004). Contrary to these findings, another study on *Brassica juncea* showed that its parental genomes remain unchanged during allopolyploid formation (Axelsson et al., 2000). Stability of parental genomes was also found in cotton allopolyploids (Liu et al., 2001). Similarly, a study of *Spartina anglica*, which is a very young natural allopolyploid, did not indicate extensive genomic changes arising from allopolyploid formation (Baumel et al., 2002). These contrasting observations among different species suggest that DNA sequence-based responses to allopolyploidization are species dependent and can vary from quiescence to massive and rapid genome reorganization (Ma and Gustafson, 2005).

1.1.2. Allopolyploidization can be accompanied by reactivation of transposable elements

Transposable elements (TEs) were discovered in maize by Barbara McClintock more than a half century ago (McClintock, 1951). They comprise sequences that can insert into new chromosomal location and thereby often create duplicate copies of themselves. TE-related sequences constitute ~40% of the human genome and 50-80% of some grass genomes (Feschotte et al., 2002). Based on the mechanism of their transposition, TEs are divided to two classes. Class 1 (retro) transposons use RNA as a transposition intermediate, while class 2 (DNA) transposons use DNA. Based on their structure and transposition mechanism, class 1 TEs are subdivided into A) LTR retrotransposons possessing long terminal repeats (LTRs) in direct orientation, and B) non-LTR retrotransposons (Feschotte et al., 2002). Most of the transposable elements are transcriptionally quiescent in the genome in which they have inserted. However, Barbara McClintock already proposed that interspecific hybridization could activate transposons due to a “genomic shock” (McClintock, 1984). In fact it has been shown that allopolyploidization events cause reactivation of retrotransposons and DNA transposons in newly formed wheat and *Arabidopsis* allopolyploids (Comai et al., 2000; Kashkush et al., 2003; Madlung et al., 2005). Reactivation of a retrotransposon in wheat was shown to be responsible

for silencing a neighbouring gene (Kashkush et al., 2003). In a synthetic *Arabidopsis* allotetraploid, microarray-based gene expression analysis revealed an increase of transposon-derived transcripts (Madlung et al., 2005).

1.1.3. Transposon activation can happen in interspecific hybrids

There are also well-documented examples of TE mobilization in animal hybrids. For instance, an interspecific hybrid between two species of Australian wallaby (*Macropus eugenii* and *Wallabia bicolor*) showed an amplification of an endogenous retroviral elements located in centromeric regions (O'Neill et al., 1998). An outstanding example of retrotransposon proliferation due to interspecific hybridization was shown for three “hybrid” sunflower species *Helianthus anomalous*, *H. deserticola* and *H. paradoxus*. These hybrids are the products of ancient hybridizations between the parental species *H. annuus* and *H. petiolaris*. The hybrid species are diploid, like their parents and have the same chromosome number (n=17). Nevertheless, at the same time they have at least a 50% higher genomic DNA content. This difference in genome size was suggested to be due to a higher relative abundance of a LTR retrotransposon sequences in the species with hybrid history (Ungerer et al., 2006). Therefore the combination of two diverged genomes itself could initiate activation of transposable elements.

1.1.4. Changes in gene expression patterns due to allopolyploidization

Inter-genomic silencing of several protein coding genes was observed in allopolyploid formation in wheat (Galili and Feldman, 1984). Analysis of gene expression in natural and synthetic *Arabidopsis* allotetraploid lines revealed differential gene expression patterns (Wang et al., 2004). Gene expression studies in synthetic allopolyploid *Arabidopsis* (Comai et al., 2000) and cotton (Brubaker et al., 1999) revealed that gene silencing occurs during the first or second generation after species hybridization. Comparative gene expression studies in synthetic allopolyploid lines of *Arabidopsis* demonstrated that around 5% of genes had a non-additive expression level (Wang et al., 2006). Floral gene expression was compared between allohexaploid *Senecio cambrensis*, its parental species, *S. vulgaris* (tetraploid) and *S. squalidus* (diploid), as well as their synthetic triploid F₁ hybrid *S. x baxteri* using microarray-based

transcript analysis. Compared with the parental species, the changes in gene expression levels in triploid *S. x baxteri* were greater than the ones in allohexaploid *S. cambrensis*. This suggested that the genome doubling leading to formation of the allohexaploid had an “ameliorating” effect on the “transcriptome shock” caused by interspecies hybridization that was visible in the triploids (Hegarty et al., 2006).

1.1.5. Hybridization-induced DNA and histone modifications

1.1.5.1. Epigenetics

Historically, the term “epigenetic” was coined by Conrad Waddington to form a new conceptual framework for the study of development (Van Speybroeck, 2002). It was derived from “epigenesis” that refers to a theory stating that the development of an organism involves a gradual increase in complexity by successive differentiation of an unstructured egg after fertilization. Epigenesis in this context is to be considered as an opposing concept to preformation theory, which postulates that the whole complexity of a differentiated mature organism is already present in the fertilized egg, or even in the gametes before fertilization, and that development is simply an enlargement of that preformed entity (Zilberman and Henikoff, 2005). The goal of C. Waddington was to synthesize a new theory from preformation and epigenesis by uniting genetics and developmental biology. After several redefinitions, the term “epigenetics” is currently used to summarize phenomena that involve changes in the expression potential of genes that are mitotically and/or meiotically heritable, but are not based on changes in the underlying DNA sequences (Wu and Morris, 2001). It has been postulated that epigenetic mechanisms provide an "extra" layer of transcriptional control that regulates gene expression (Workman and Abmayr, 2004).

Distinct mechanisms appear to be closely related in the initiation, sustaining and inheritance of epigenetic information at particular DNA sequences. They comprise association with DNA and histone modifications, binding of non-histone chromatin proteins with specific affinity to these DNA or histone modifications, loading with particular histone variants, association with nuclear RNAs, and integration into a higher-order organization of chromatin in the nucleus (Probst et al., 2009). Two hypotheses have been proposed to explain mechanistically the function of DNA and histone modifications in this context. In the first scenario, accessibility of proteins

interacting with DNA such as transcription factors is controlled directly via changed compactness of the chromatin. The alteration of compactness could arise from either changes in the electrostatic charge or interactions of nucleosomes as a result of chromatin modifications. The open chromatin, configuration would increase the accessibility of genomic DNA to the transcription machinery while the closed chromatin configuration would reduce transcription by limiting the accessibility of genomes to these transcription factors (Pfluger and Wagner, 2007). The other scenario suggests that attachment of these modifications might alter the structure of the nucleosome and consequently cause recruitment of several different chromatin associating factors (Berger, 2007).

1.1.5.1.1. DNA methylation

The most intensively studied epigenetic modification is DNA methylation. DNA methylation involves the reversible addition of a methyl group to the fifth carbon of the cytosine's pyrimidine ring by a DNA methyltransferase. DNA methylation plays an important role in a wide range of essential biological processes such as embryogenesis, genomic imprinting, tumorigenesis in mammals and transposon silencing and gene regulation in plants (DeAngelis et al., 2008; Lister et al., 2008).

In contrast to mammals in which DNA methylation predominates at CpG dinucleotides, DNA methylation in plants can occur at any cytosine regardless of its sequence context, for instance CG [cytosine (C); guanine (G)], CHG [H is adenine (A), thymine (T), or cytosine (C)], and CHH (Chan et al., 2005). In *A. thaliana*, three different methylation systems are thought to be responsible for establishment and maintaining cytosine methylation in these three different sequence contexts. The *A. thaliana* genome contains methylation at 24% of its CG, 6.7% of its CHG and 1.7% of its CHH sites (Cokus et al., 2008).

Indirect immunostaining using an antibody against 5-methylcytosine in *A. thaliana* revealed that methylated DNA colocalizes with heterochromatic chromocenters which contain repetitive DNA sequences flanked by pericentromeric DNA transposons and retroelements (Soppe et al., 2002). Most of the transposons localized in euchromatin are also highly methylated. In contrast, expressed genes inserted within the heterochromatic chromocenters, are free of DNA

methylation although they are surrounded by methylated transposons (Lippman et al., 2004; Zhang et al., 2006). The bias of DNA methylation towards repetitive DNA sequences may indicate that transcriptional silencing of transposable elements is a primary function of DNA methylation (Gehring and Henikoff, 2008). Genome-wide analysis of DNA methylation in *A. thaliana* by coupling chromatin immunoprecipitation using a specific antibody against methylated cytosine and high-density whole-genome tiling microarrays at 35 base pair resolution reconfirmed significant enrichment of DNA methylation in pericentromeric repetitive sequences and in regions producing small interfering RNAs (Zhang et al., 2006).

It is surprising that more than one-third of *A. thaliana* genes have substantial amounts of DNA methylation within their transcribed regions. However, only 5% of genes show methylation within their promoter regions. In addition, genes with methylation in their promoter regions had a tendency to be expressed in a tissue-specific manner while “body methylated” genes were constitutively expressed at high levels compared to non-methylated ones. Another study on *A. thaliana* methylome using high-density tiling microarrays revealed that methylation within coding regions is strongly biased away from the promoter and 3' end which is just inverse to the localization of eukaryotic RNA polymerase II. Moderately transcribed genes are also more prone to be methylated than genes with extremely high or low expression levels, whereas, shorter methylated genes tend to be poorly transcribed. These results suggest that gene methylation might be a general property of eukaryotes and in *A. thaliana* is highly correlated with the respective levels of transcription (Zilberman et al., 2007).

A whole genome single base pair resolution DNA methylation map of *A. thaliana* generated by combining bisulfite treatment of genomic DNA with ultra-high-throughput sequencing using the Illumina 1G Genome Analyzer and Solexa sequencing technology showed that methylation in transcribed genic regions not associated with short interfering RNA (siRNA), almost exclusively occurred in a CG sequence context. In contrast, genomic regions producing siRNA showed methylation at CG, CHG and CHH sites. Different DNA methyltransferases of *A. thaliana* responsible for methylation of cytosine in different contexts act redundantly to establish these patterns (Cokus et al., 2008).

DNA methylation is a reversible modification. Passive loss can occur by DNA replication in the absence of DNA methylation maintenance. Active loss of DNA methylation requires enzymatic activities like DNA glycosylases similar to the ones involved in base excision DNA repair and could protect the genome from potentially deleterious methylation (Penterman et al., 2007). The steady state pattern of DNA methylation is shaped by interaction between DNA methyltransferases, chromatin remodelling factors, the RNAi machinery, and DNA methylation removing mechanisms (Saze et al., 2008). Comparison of DNA methylation among different accessions of *A. thaliana* revealed natural polymorphisms in this epigenetic modification mark with an additive mode of inheritance after intraspecific hybridization (Zhang et al., 2008).

1.1.5.1.2. Post-translational histone modifications

In the eukaryote nucleus, genomic DNA is packed into a complex structure called chromatin that includes of DNA, RNA and proteins. Its elementary building block, the nucleosome, consists of ~147 base pairs of DNA that are wrapped around the surface of a histone octamer in 1.65 superhelical turns. This octamer is composed of a central histone (H3-H4)₂ tetramer which is flanked by two histone H2A-H2B dimers (Rocha and Verreault, 2008). All histones involved in nucleosome formation have in general a tripartite structure with a globular structured domain in centre surrounded by flexible and largely unstructured tails. These domains are rich in lysines and arginines and, therefore, rather basic.

The optional presence of numerous diverse chemical modifications (also referred to as histone ‘marks’) in different positions especially at their N-terminal tails, is a characteristic feature of histones (Godde and Ura, 2008). For instance, histones can undergo methylation at their lysine and arginine and phosphorylation at their serine and threonine residues. These modifications are thought to regulate different biological processes like transcription, DNA repair, replication and chromatin condensation (Fischle et al., 2003). The related “histone code hypothesis” states that (i) different modifications on histone tails could change their affinity to chromatin-associated proteins. Accordingly, the recruitment of proteins to particular chromatin domains for regulatory purpose would rely on the presence of specific histone modifications (Strahl and Allis, 2000); (ii) different modifications on the same or different histone tails would possibly cross talk, thereby generating complex combinations of marks with potentially distinct effects.

In consequence (iii), higher order structures of chromatin like hetero- or euchromatic domains would depend on the local frequency and combination of histone modifications (Jenuwein and Allis, 2001). Antibodies specific to particular modifications have been used in immunostaining of nuclei or in chromatin immunoprecipitation coupled with microarray-based hybridization analysis (ChIP on chip) to elucidate the global distribution of these histone modifications. From the studies done so far for a subset of modifications, it has become clear that they are not uniformly distributed across the genome (Kouzarides, 2007).

1.1.5.1.3. Histone methylation

Histone methylation is mediated by enzymes belonging to three distinct families of proteins, the PRMT1 family, the SET-DOMAIN-containing protein family, and the non-SET-domain proteins DOT1/DOT1L (Martin and Zhang, 2005). All known histone methyltransferases use S-adenosyl-methionine as the methyl group donor (Shilatifard, 2008). One of the major sites for histone methylation is the basic amino acid side chain of lysine (Zhang and Reinberg, 2001). Lysine residues can acquire mono- (me1), di- (me2), or trimethylation (me3) *in vivo*. Each methylation state can be controlled by a different methyltransferase enzyme and can be associated with distinct effects and localizations.

Histone H3 dimethylated at lysine 4 (H3K4me2) is generally associated with transcriptionally active chromatin, whereas histone H3 trimethylated at position lysine 27 (H3K27me3) or dimethylated at position lysine 9 (H3K9me2) is associated with silent chromatin (Kouzarides, 2007). A genome-wide high resolution analysis of histone modifications in *Saccharomyces cerevisiae* revealed that active transcription is characteristically accompanied by H3K4me2, which is distributed mainly across the body of active genes (Pokholok et al., 2005). The distribution pattern of H3K4me2 in mouse and human cell lines was also shown to reside in vicinity of active genes (Bernstein et al., 2005). Cytological studies via immunostaining by antibodies against H3K4me2 in several plant species revealed that this modification is exclusively distributed across euchromatic regions independent from the genome size (Houben et al., 2003). H3K9me2 is a mark characteristic for heterochromatin. The study of polytene chromosomes in *Drosophila* proved the presence of this mark in the pericentric heterochromatin regions (Jacobs et al., 2001). In *A. thaliana*, establishment of H3K9me2 is

mainly catalyzed by histone methyltransferase KRYPTONITE (KYP or SUVH4) (Jackson et al., 2004). H3K9me2 is a mark for silencing of transposons and due to its distribution at centromeric and pericentromeric heterochromatin regions is considered as an obligate characteristic of heterochromatin (Tariq and Paszkowski, 2004). In *A. thaliana*, it has been shown that H3K9me2 is an important signal for the maintenance of non-CG methylation (Henderson and Jacobsen, 2007) and that in turn, reduction of DNA methylation is associated with the depletion of H3K9me2 (Jackson et al., 2002; Soppe et al., 2002; Mathieu et al., 2007).

Polycomb group (PcG) proteins act in developmental control of gene activity (Kinoshita et al., 1999). A conserved protein complex known as polycomb repressive complex 2 (PRC2) is known in both plants and animals. It mediates H3K27 trimethylation on its target loci such as silent genes distributed in euchromatic regions (Lachner et al., 2003; Pien and Grossniklaus, 2007). In plants, several PRC2 complexes with overlapping subunits exist which have distinct specialized functions during development (Hsieh et al., 2003). Disruption of H3K27me3 formation is accompanied by severe developmental abnormalities in *A. thaliana* (Schubert et al., 2005). Genome-wide analysis of H3K27me3 distribution in *A. thaliana* by ChIP-on-chip revealed that a large number of genes (~4,400) are targeted by this mark (Zhang et al., 2007). In contrast to *Drosophila* and mammals in which H3K27me3 sometimes covers domains of hundreds of kilobases spanning several genes, in *A. thaliana* H3K27me3 enriched regions are considerably shorter and mainly located at the 5' ends of transcribed regions. In most cases, H3K27me3 enriched regions were confined to single genes. Sequences immediately upstream of gene promoters and downstream of the 3' ends of genes had lower amounts of H3K27me3 relative to the genome average. These differences in H3K27me3 patterning in *A. thaliana* compared with other model organisms suggests that distinct mechanisms may be responsible for the establishment and spread of this mark in plants and animals (Zhang et al., 2007).

1.1.5.2. DNA methylation changes in allopolyploids

A combination of redundant and diverged homoeologous sets of genes in allopolyploid species might trigger DNA methylation changes which could play an important role both gene activation and gene silencing as observed in newly formed *Arabidopsis* allotetraploids (Madlung et al., 2002). Similarly, synthetic allopolyploid individuals of wheat showed that

gene silencing was correlated with repatterning of the cytosine methylation level (Kashkush et al., 2002).

Possible consequences of hybridization and genome duplication for DNA methylation were investigated in *Spartina* allopolyploids by methylation sensitive amplification polymorphism (MSAP), which revealed that approximately 30% of the parental methylation patterns are altered in hybrids and allopolyploids (Salmon et al., 2005). Southern analysis showed changes of cytosine methylation in CG and CHG context in allotetraploid *Brassica* compared with its diploid parents (Song et al., 1995). Changes of DNA methylation induced by wide hybridization were reported for synthetic *Cucumis* allopolyploids (Chen and Chen, 2008). Similarly, differential methylation of CG or CHG context was observed in intraspecific F₁ maize hybrids and their respective parents (Tsaftaris et al., 1997) and rice (Xiong et al., 1999). However, no DNA methylation alterations were detected in the process of allopolyploidization of cotton (Liu et al., 2001).

1.1.6. Activation and inactivation of rRNA genes in allopolyploid species

Ribosomal RNAs constitute the catalytic and architectural components of ribosomes. In eukaryotes, they are encoded by several hundred to thousands of gene copies. These genes are organized in tandem repeats which are clustered in chromosomal domains known as nucleolus organizer regions (NORs). RNA polymerase I (Pol I) is responsible for transcribing these rRNA genes to a pre-mature rRNA (45S rRNA), which is later processed to mature rRNA (18S, 5.8S and 25-27S, the sizes being species-dependent (Chen and Pikaard, 1997). The activity of Pol I is confined to the nucleolus as the nuclear compartment of ribosome biosynthesis. The genes encoding 5S rRNA are located elsewhere in the genome and are transcribed by RNA polymerase III (Pol III). Pre-rRNA processing, modification and the initial steps of pre-ribosome assembly of the four transcribed rRNA with ribosomal proteins take place in the nucleolus (Scheer and Weisenberger, 1994; Shaw and Jordan, 1995). In *A. thaliana*, approximately 8 million base pairs are devoted to rRNA genes at two chromosomal loci, NOR2 and NOR4. The level of transcription of rRNA genes depends on the demand of the cell for protein synthesis. In consequence, in cells with high growth rates most of the transcripts belong to rRNA (Grummt and Pikaard, 2003). Therefore it has been thought that stimulators of cell

growth and/or proliferation directly regulate transcription of ribosomal genes (Warner, 1999; Stefanovsky et al., 2001). Furthermore, it was shown that the size of the nucleolus correlates with the cell proliferation rate in cancer tissue (Derenzini et al., 2000).

Electron microscopic studies have shown that actively transcribing rRNA genes loaded with Pol I molecules and pre-rRNA are localized directly adjacent to rRNA genes without any activity. Also, different inbred lines of maize possessing different copy numbers of rRNA genes ranging from 2500 to 24000 did not show an obvious effect of rRNA copy number on growth rates (Rivin et al., 1986). This is consistent with the view that not all rRNA genes are active at any one time.

Nucleolar dominance is a well known phenomenon that describes the consistent activity of the nucleolus organizer (NORs) of only one of the two parental genomes in hybrid individuals. This occurs in a broad range of organisms like insects, amphibians, mammals and plants (Chen and Pikaard, 1997). The phenomenon was first discovered in the plant genus *Crepis* (Navashin, 1934) and was later found to exist in interspecific hybrids of many plant genera (Pikaard, 2000). This silencing is restricted to the rRNA genes, while neighbouring protein coding genes are not affected. The pattern of rRNA gene silencing in hybrids is heritable and, therefore, nucleolar dominance is considered to be a classic example of epigenetic control of gene expression. Studies in *Brassica* allopolyploids and their diploid progenitors revealed that the rRNA genes of only one progenitor are active in vegetative tissues and are independent of maternal effects, ploidy levels (tetra or hexa allopolyploids), or rRNA gene dosage. Natural and synthetic allopolyploid individuals showed the same parental rRNA gene- silencing pattern (Chen and Pikaard, 1997). In *A. suecica*, the allotetraploid hybrid of *A. thaliana* and *A. arenosa*, the rRNA genes inherited from *A. arenosa* are transcribed whereas rRNA genes derived from *A. thaliana* were transcriptionally silenced (Chen et al., 1998). Loss-of-function and blocking of a histone deacetylase and DNA methylation in *A. suecica* indicated that deacetylation of histones in association with DNA and histone methylation plays an important role for selective silencing of *A. thaliana*-derived rRNA genes (Lawrence et al., 2004; Earley et al., 2006). Subtle regulation of rRNA genes might also happen in intraspecific crosses. The genetic mapping of heterosis-related loci in hybrid offspring of two *A. thaliana* accessions located indeed a relevant QTL in the region of a cluster of rRNA genes (Meyer et al., 2009).

1.2. Heterosis

The term heterosis was introduced by Shull in a lecture at Göttingen University in 1914 to replace the term “heterozygosis” in order to describe the increased vigour that is observed in heterozygous hybrid F₁ organisms compared with their homozygous parents (Shull, 1948). Offspring of intraspecific crosses often show superior phenotypes compared to their inbred parents for adaptive traits such as fertility, biomass, growth rate, yield and flowering time (Birchler et al., 2003; Meyer et al., 2004). Maximum heterosis is observed in the first generation, while the superiority of the progeny over their parents is progressively lost in subsequent selfed generations. The importance of heterosis in agriculture is evident from the significant yield increases over the past 50 years following the introduction of hybrid crops (Duvick, 1999 and 2001).

Yet despite almost a century of investigation and the great benefit to agriculture of hybrid crops, the principles underlying heterosis are still not understood at the molecular level. Although several hypotheses have been proposed to explain this enigmatic phenomenon, there is no generally accepted theory for the genetic mechanisms underlying heterosis. Among the current models, the most prevailing and non-exclusive ones are genome-wide dominant complementation, locus-specific overdominance, and epistasis (Fig. 3) (Tsaftaris, 1995).

The dominance hypothesis states that each inbred parent line contains in addition to beneficial alleles some slightly deleterious alleles. In hybrid offspring, these inferior recessive alleles are compensated by beneficial or superior dominant alleles originating from the other parent. An early argument against this hypothesis was that if the dominance complementation was the reason for heterosis, then it should be possible to create an inbred line which contains mostly superior alleles and, therefore, would show little or no hybrid vigour. However, the practical generation of such a homozygous line would be difficult, if not impossible, because of the inevitable genetic linkage between at least some superior and deleterious alleles. On the other hand, according to the dominance hypothesis the complementation of deleterious by superior alleles from either inbred parent in hybrid heterozygotes would need to have a cumulative effect on the level of heterosis. Otherwise, one would expect the hybrid to be equivalent to the better of the two homozygous lines for the effect of any individual gene. Nevertheless, although

the quality of inbred lines has been improved during the last century, the amount of heterosis by hybridization of these inbred lines has not decreased. In other words, the amount of heterosis can not be predicted by solely looking at the quality of inbred parents.

Progressive heterosis in hybrids between polyploids, which appears to improve according to the numbers of combined genomes, is another argument against the dominance theory. Increased heterosis at each stage of combining genomes would demand an increase of superior alleles to compensate for pre-existing deleterious alleles without introducing new deleterious alleles at other loci. The likelihood for this seems rather low and other mechanisms probably explain heterosis in addition to simple complementation.

The overdominance theory postulates that in the heterozygous F₁ offspring the interaction between different alleles at single loci will generate a synergistic effect on vigour which exceeds both homozygous parents. Consequently heterozygous genotypes show superior fitness over homozygous genotypes (Lippman and Zamir, 2007). Evidence from different studies suggested that dominance and overdominance may have both a role in heterosis, but their respective contributions remain unclear (Cockerham and Zeng, 1996; Li et al., 2001; Lippman and Zamir, 2007).

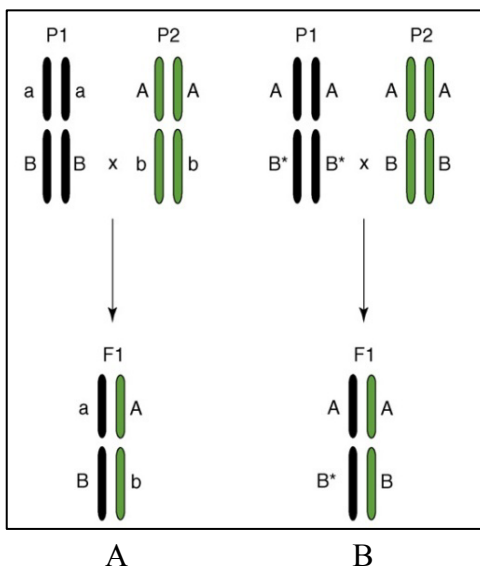


Figure 3 Genetic models of heterosis.

A: Genome-wide dominance complementation indicates a situation in which independent slightly deleterious alleles at different loci ('a' in parent 1 (P1) and 'b' in P2) in each inbred

parental line are complemented by superior alleles ('A' in P2 and 'B' in P1) in the other parental line. B: The overdominance theory proposes that interactions between different alleles in a single locus (here allele 'B*' in P1 and 'B' in P2, assuming no dominance or recessive relationship) results in a performance that surpasses the combination of both parental homozygous alleles. (copied from Lippman and Zamir, 2007).

In addition to these "classic" hypotheses, various other physiological and molecular mechanisms have been proposed to explain heterosis. Some examples are changes in the dosage of transacting regulatory elements, differences in the interaction networks of regulatory elements in hybrid offspring, changes in gene expression patterns, and alteration of epigenetic modifications (Comings and MacMurray, 2000; de Vienne et al., 2001; Birchler et al., 2003).

At the molecular level, we can imagine two different modes for the expression of alleles in hybrids. Either they will retain their parental expression level and thereby show additive expression or their expression is altered by the new interactions between *cis* and *trans* factors from the two different genomes resulting in a new expression pattern deviating from the midparent value (Birchler et al., 2003).

The present challenge in the study of heterosis is to make a causative link between an observed heterotic phenotype and molecular events associated with genome hybridization in F₁ offspring. For instance, general yield as an output of plant production is composed of multiple quantitative traits such as rate of vegetative growth, flowering time, inflorescence number, flowers per inflorescence, fruit or grain set, and fruit or grain weight. Any of these quantitative traits could have a different effect on general yield, and at the same time, be controlled by different molecular mechanisms.

2. Aims of the thesis

In this study I have investigated changes in chromatin modifications and other nuclear features in response to intraspecific hybridization and to evaluate their potential contribution to heterosis. Inbred accessions of *A. thaliana* and their intraspecific hybrids were used as a model. Analyzes were done with different methods at the whole genome and the single gene level to address the following questions:

2.1. Is endopolyploidization altered by intraspecific hybridization?

Flow cytometry was employed to investigate whether intraspecific hybridization is accompanied by a higher level of endopolyploidization in hybrid offspring plants. Endopolyploidy levels of inbred parental lines (Col-0 and C24) and their reciprocal hybrid offspring (Col-0xC24 and C24xCol-0) were compared at four different developmental stages.

2.2. Is rRNA gene activity altered in intraspecific hybrids?

Nucleolus size in interphase nuclei was determined by silver staining to assess rRNA gene activity in inbred parental lines (Col-0 and C24) and their reciprocal hybrid offspring (Col-0xC24 and C24xCol-0). Nuclei with defined ploidy levels were separated via flow cytometry from 6DAS seedlings.

2.3. Does microscopic analysis detect altered DNA methylation, histone H3K4me2, H3K27me3, or H3K9me2 in intraspecific hybrids?

Indirect immunostaining using specific antibodies against different chromatin modifications was performed on isolated interphase nuclei with defined ploidy level. The global distribution and intensity of DNA methylation, histone H3 dimethylation at position K4 and K9, and histone H3 trimethylation at position K27 in inbred parental lines (Col-0, C24 and Cvi) and their reciprocal hybrids (Col-0xC24, C24xCol-0, Col-0xCvi and CviXCol-0) was investigated.

2.4. Do DNA methylation patterns change in response to intraspecific hybridization?

Methylation sensitive amplified polymorphism (MSAP) analysis was used to survey DNA methylation patterns of inbred parental lines (Col-0 and C24) compared with their reciprocal hybrid offspring (Col-0xC24 and C24xCol-0).

2.5. Is DNA methylation in 5' regulatory regions of selected genes altered in intraspecific hybrids?

Analysis of DNA methylation in the 5' regulatory region was analyzed for seven candidate genes for which differential expression in inbred parental lines (Col-0 and C24) and their reciprocal hybrid offspring (Col-0xC24 and C24xCol-0) had been detected by micro array-based transcript analysis (T. Altmann, personal communication). Sequence differences in the 5' regulatory and coding regions of these genes in Col-0 and C24 were determined. Based on this information, allele-specific tests for DNA methylation in 5' regulatory regions of these genes by methylation sensitive restriction cleavage and cleavage amplified polymorphic sequence (CAPS) analysis were established.

2.6. Are allele-specific transcript levels of selected genes altered in intraspecific hybrids?

The expression levels of the respective alleles of seven selected genes were determined in inbred parental lines (Col-0 and C24) and their hybrid offspring (Col-0xC24 and C24xCol-0) using quantitative RT-PCR and cleavage amplified polymorphic sequence (CAPS) analysis.

2.7. Is intraspecific hybridization accompanied by establishment of new histone modification patterns?

This question was addressed by using chromatin immunoprecipitation combined with microarray hybridization (Chip on chip) analysis in inbred parental lines (Col-0, C24 and Cvi) and their reciprocal hybrid offspring (Col-0xC24, C24xCol-0, Col-0xCvi and CvixCol-0).

2.7.1. Comparative genomic hybridization (CGH) analysis of different *A. thaliana* accessions

The microarray systems adopted for *A. thaliana* were based on the genomic sequence of accession Col-0. As substantial sequence polymorphism exists between different accessions, prior calibration by comparative genome hybridization (CGH) was required for the analyzes involving material from other accessions. Sequence polymorphisms between the accessions Col-0, C24, and Cvi were determined using chromosome 4 tiling microarrays and *Arabidopsis* whole genome NimbleGen arrays.

2.7.2. ChIP on chip comparison of histone modifications in parental inbred lines and their hybrid offspring

Histone modifications H3K4me2 and H3K27me3 were compared between parental lines (Col-0, C24, and Cvi) and their hybrid offspring (Col-0xC24 and Col-0xCvi) by Chip on chip analysis.

3. Material and methods

3.1. Plant material

3.1.1. Production of hybrid seeds for *A. thaliana* accessions Col-0, C24 and Cvi

Arabidopsis thaliana accessions Col-0, C24, and Cvi and their reciprocal F₁ hybrid offspring Col-0xC24, C24xCol-0, Col-0xCvi, and CvixCol-0 were used as a model system. Seeds of inbred parental lines were sown on soil and stratified for 3 days at 4°C. After 15 days under short day conditions (8 hour light, 16 hour dark, 21°C), culture was continued under long day conditions (16 hour light, 8 hour dark, 21°C) until plants started to produce inflorescences. For crosses between accessions (Col-0xC24, C24xCol-0, Col-0xCvi, and CvixCol-0), five flower buds on the primary shoot were emasculated and manually cross-pollinated, while all other flower buds were removed. The same procedure was applied to the plants selected to produce parental seeds (Col-0, C24, and Cvi) except that the remaining five flower buds were permitted to self-pollinate. After ripening of siliques, seeds from crossed and self-pollinated plants were reaped and pooled according to their genotype.

3.1.2. Cultivation of *A. thaliana* seedlings in soil

Seedlings to be harvested 4 and 6 days after “sowing” (DAS, actually, more correct “after stratification” excluding time at 4°C), were cultivated on square plastic plates (120 mm x 120 mm) filled with a mix of 50% of gardening soil “Einheitserde Typ GS 90” with 50% “Vermiculite Körnung 2” and covered with nylon mesh. Each plate was divided into 4 sectors for simultaneous cultivation of two inbred parental lines and two reciprocal F₁ hybrids on the same plate. Per sector, 2 mg (equalling 80-100 seeds) of seeds were sown. The plates were sealed with Leukopor BSN Medical tape, stored at 4°C for 3 days for stratification and then incubated with 16 hours light per day with a light intensity of approximately 120 µE, 20°C day temperature and 18°C night temperature. The lids of the plates were removed after 3 days incubation time and the seedlings were from then irrigated daily for the remaining cultivation time until they were harvested and used directly or stored at -80°C. Seedlings to be harvested at 10, 15 and 21 DAS were grown directly on soil without nylon mesh with each genotype being grown in a separate pot.

3.1.3. Soil-free cultivation of *A. thaliana* seedlings for chromatin extraction

Approximately 400 seeds were surface sterilized and cultured in liquid medium under controlled conditions. For sterilization, seeds were washed two times for 5 minutes in 5 ml of [4.8% (w/v) NaClO + 3 drops of Tween 20] and subsequently rinsed three times for 5 minutes in distilled water. Liquid medium containing 2.2 g/l Murashige and Skoog basal salt mixture (Sigma M5524), 0.5 g/l MES, and 7 g/l sucrose was adjusted to pH 5.8 with KOH and sterilized by autoclaving. After adding vitamins from a 1000x stock solution (Murashige and Skoog, vitamin powder, x1000, M7150, Sigma) to a 1x concentration, 60 ml of medium were combined with each sterilized seed batch in a sterile Petri dish of 15 cm diameter. After sealing with parafilm, dishes were incubated for 10 days in a controlled growth chamber under controlled conditions with 16 hours light per day, light intensity of approximately 150 μ E, 22°C day temperature, and 18°C night temperature.

3.2. Determination of the endopolyploidization and flow sorting of nuclei

The endopolyploidization levels of nuclei isolated from seedlings, cotyledons or leaves at 4, 6, 10 and 15 DAS, respectively, were measured using a FACStarPLUS device (BD Biosciences) according to Barow and Meister (2003). Cycle values indicating the mean number of endoreduplication cycles per nucleus in a cell population were calculated based on the formula: $\text{cycle value} = (0 \cdot n_{2C} + 1 \cdot n_{4C} + 2 \cdot n_{8C} + 3 \cdot n_{16C} \dots) / (n_{2C} + n_{4C} + n_{8C} + n_{16C} \dots)$ where n_{2C} , n_{4C} , n_{8C} , . . . are the numbers of nuclei with the corresponding C-value (2C, 4C, 8C, . . .) and coefficient numbers (0, 1, 2, . . .) are the number of endoreduplication cycles necessary to reach the corresponding ploidy level.

For silver staining and immunostaining experiments, 2C and 4C nuclei from soil-grown 6 DAS seedlings were flow sorted using a FACSARIA device (BD Biosciences) and “dropped” onto slides as described in Pecinka et al. (2004).

3.3. Silver staining of size-fractionated nuclei for nucleolus measurement

Nucleoli of 2C and 4C nuclei from 6 DAS seedlings were silver-stained according to Hizume et al. (1980) and Rufas et al. (1982). Slides with attached nuclei were washed in borate buffer (0.01M boric acid, pH 9.2) and stained by incubation with freshly prepared 50% (w/v) silver

nitrate solution (in distilled water, adjusted to pH 4-5 with formic acid) for 1-2 hours at 65°C. Images were taken with an Olympus BX61 microscope equipped with an ORCA-ER CCD camera. The darkly stained nucleolus regions in the images were encircled manually and their areas were measured using SIS software (Olympus).

3.4. Indirect immunostaining of histone modifications in sorted nuclei

3.4.1. Immunostaining for histone H3K4, H3K9 and H3K27 methylation

Polyclonal rabbit antibodies against histone H3K27me₃, H3K4me₂ and, H3K9me₂ (Millipore, cat. nos. 07-449, 07-030 and 07-441, respectively) were used to detect the distribution of these marks in flow-sorted 4C nuclei. The slides with fixed nuclei were stored in 100% glycerol at 4°C. For immunostaining, they were washed with 1X phosphate buffered saline (PBS) and subsequently were incubated with blocking solution (0.1% (v/v) Tween-20 or Triton X-100, 5% (w/v) BSA in 1X PBS) for 45 minutes at ambient temperature. After two further washes for 5 minutes in 1X PBS, 20 µL of primary antibody solution (antibody stock diluted 1:200 in PBS containing 1% BSA) were applied to each slide. The slides were covered with parafilm and incubated in a moist chamber overnight at 4°C. After washing with 1X PBS buffer at ambient temperature, incubation with 25 µL secondary antibody, anti-rabbit-Cy3 (antibody stock diluted 1:100 in 1X PBS with 1% BSA), was performed at 37°C for 1 hour. Counterstaining was done with 20 µL antifade solution (70% (v/v) glycerol, 30 mM Tris pH 9, 2.5% (w/v) n-propylgallate) containing 10 mg/L 4', 6-diamidino-2-phenylindole (DAPI). The slides were covered with cover slips and analyzed with an Olympus BX61 fluorescence microscope equipped with an ORCA-ER CCD camera. Images were analyzed using the SIS software (Olympus).

3.4.2. Immunostaining for 5-methylcytosine

5-methylcytosine immunodetection was performed according to Castiglione et al. (2002). In nuclei fixed to slides, chromosomal DNA was denatured by immersing the slides for 2 min in 70% (v/v) formamide (in 2X SSC) at 70°C. Subsequently, the slides were dehydrated by insertion into ice cold 70% and 100% ethanol, and finally air dried. After blocking at 37°C in blocking solution (0.1% (v/v) Tween-20 or Triton X-100, 5% (w/v) BSA in 1X PBS), slides were incubated with 20 µL mouse monoclonal 5mC-antibody solution (Eurogentec, cat. no.

MMS-900P-A, antibody stock diluted 1:100 in PBS containing 1%BSA), covered with parafilm, and incubated in a moist chamber overnight at 4°C. Afterwards, 20 µL secondary FITC-conjugated anti-mouse antibody (antibody stock diluted 1:50 in 1% BSA in PBS) were applied and slides were incubated for 1h at 37°C in darkness. Counterstaining was done with 20 µL antifade solution (70% (v/v) glycerol, 30 mM Tris pH 9, 2.5% (w/v) n-propylgallate) containing 10 mg/L 4', 6-diamidino-2-phenylindole (DAPI). The slides were covered with cover slips and analyzed with a Olympus BX61 fluorescence microscope equipped with an ORCA-ER CCD camera. Images were analyzed using the SIS software (Olympus).

3.5. Extraction of genomic DNA and restriction cleavage

Plant genomic DNA was extracted using the QIAGEN DNeasy Plant DNA extraction system according to manufacturer's instructions. Cleavage with restriction enzymes from NEB was performed according to manufacturer's instructions.

3.6. Methylation sensitive amplified polymorphism (MSAP) assay

DNA methylation patterns in pooled samples of Col-0, Col-0xC24, C24xCol-0, and C24 plants harvested 6 and 21 DAS were compared by MSAP analysis according to Cervera et al. (2002) in four biological replicates. Briefly, genomic DNA of each genotype was cleaved with restriction endonucleases *EcoRI* (recognition sequence GAATTC) and either *HpaII* [in 23.5 µL volume] (recognition sequence CCGG, cleavage inhibited by methylation of the inner C, but not inhibited by hemi-methylation of the outer C) or *MspI* [in 22.5 µL volume] (recognition sequence CCGG, cleavage not inhibited by methylation of the inner C, but inhibited by methylation of the outer C). After restriction cleavage, *EcoRI* and *HpaII* / *MspI* adapters (Table 1) were ligated by adding to each cleavage reaction 5 µL of a mix containing 5 pmol of *EcoRI* adapter, 50 pmol of *HpaII/MspI* adapter, 40 mM Tris-HCl pH 7.8, 10 mM MgCl₂, 10 mM DTT, 0.5 mM ATP and 1U of T4 DNA ligase (Fermentas) followed by incubation of reactions for 3 h at 37°C and overnight at 4°C. DNA fragments with ligated adapters were diluted 5-fold and served as templates for primary PCR amplifications using pre-selective primers (*EcoRI*-A and *HpaII/MspI*-0) complementary to the *EcoRI* and *HpaII* / *MspI* adapters, the former with one additional selective nucleotide at the 3' end (Table 2). The PCRs were performed in a 20 µL volume of 1X PCR buffer (Qiagen 10X PCR buffer), 1.5 mM MgCl₂, 0.2 mM of each dNTP,

0.6 μ L of each primer [10 pmol] *EcoRI*+A and *HpaII/MspI*+A, 2.5 U of *Taq* DNA polymerase (QIAGEN) and 3 μ L of 5-fold diluted ligation products. PCR was done with a protocol including 3 minutes 95°C denaturation followed by 20 cycles of 30 seconds at 94°C, 1 minute annealing at 60°C and 1 minute extension at 72°C with ramping of 1°C/s. PCR products were diluted 15-fold with double distilled water and used as templates for secondary selective amplification with combinations of selective primers complementary to the *EcoRI* and one *HpaII* / *MspI* adaptors, but this time with three (*HpaII/MspI*-AAT, *HpaII/MspI*-ATC, and *HpaII/MspI*-ACT) or two (*EcoRI* –AC, *EcoRI* -AA *EcoRI* –AG, and *EcoRI* –AT) selective nucleotides, respectively, at the 3' end (Table 3). The PCR was carried out in 20 μ L volume of 1X PCR buffer (Qiagen 10X PCR buffer), 1.5 mM MgCl₂, 0.2 mM of each dNTP, 0.1 μ L of *EcoRI* primer [10 pmol] (5' end-labelled with IRD700 fluorescent dye), 0.6 μ L of *HpaII/MspI* primer [10 pmol], 2.5 U of *Taq* DNA polymerase (QIAGEN), and 5 μ L of 15 times diluted pre-amplified DNA. The PCR protocol included [1] 94°C for 2 minutes, [2] 13 cycles of 30 seconds 94°C, 30 seconds annealing at 60°C (touch-down profile 60°C - 0.7°C/cycle) and 1 minute extension at 72°C, and [3] 23 cycles of 30 seconds 94°C, 30 seconds annealing at 56°C and 1 minute extension at 72°C with ramping of 1°C/s. PCR products were combined with the differentially labeled size marker microSTEP 15a (Microzone, Lewes, UK) and analyzed by capillary electrophoresis in a BECKMAN COULTER CEQ 8000 Genetic Analysis System. Obtained data were evaluated using CEQ 8000 software. In total, eight different primer combinations were used to compare DNA methylation patterns of Col-0, C24 and their reciprocal hybrids (Table 4).

Table 1 Adapters used in MSAP assays.

adapter name	sequence
<i>EcoRI</i> -adapterI	5'CTCGTAGACTGCGTACC 3'
<i>EcoRI</i> -adapterII	5' AATTGGTACGCAGTC 3'
<i>HpaII/MspI</i> -adapterI	5' GACGATGAGTCTCGAT 3'
<i>HpaII/MspI</i> -adapterII	5' CGATCGAGACTCAT 3'

Table 2 Pre-selective primers used in MSAP assays.

name	sequence
<i>EcoRI</i> -A	5' GACTGCGTACCAATT <u>C</u> A 3' ¹
<i>HpaII/MspI</i> -0	5' ATGAGTCTCGATCGGA 3'

Table 3 Selective primers in MSAP assays.

name	sequence
<i>HpaII/MspI</i> -AAT	5' ATGAGTCTCGATCGGA <u>AAT</u> 3' ¹
<i>HpaII/MspI</i> -ATC	5' ATGAGTCTCGATCGGA <u>ATC</u> 3'
<i>HpaII/MspI</i> -ACT	5' ATGAGTCTCGATCGGA <u>ACT</u> 3'
² <i>EcoRI</i> -AC	5' GACTGCGTACCAATT <u>CAC</u> 3'
<i>EcoRI</i> -AA	5' GACTGCGTACCAATT <u>CAA</u> 3'
<i>EcoRI</i> -AG	5' GACTGCGTACCAATT <u>CAG</u> 3'
<i>EcoRI</i> -AT	5' GACTGCGTACCAATT <u>CAT</u> 3'

¹ Selective nucleotides are underlined.

² Primers 5' labeled with IRD700 fluorescence dye are in set in bold font.

Table 4 Primer combinations used in MSAP assays.

Primer combination
<i>EcoRI</i> +AC/ <i>HpaII/MspI</i> +AAT
<i>EcoRI</i> +AC/ <i>HpaII/MspI</i> +ATC
<i>EcoRI</i> +AC/ <i>HpaII/MspI</i> +ACT
<i>EcoRI</i> +AA/ <i>HpaII/MspI</i> +AAT
<i>EcoRI</i> +AA/ <i>HpaII/MspI</i> +ATC
<i>EcoRI</i> +AA/ <i>HpaII/MspI</i> +ACT
<i>EcoRI</i> +AG/ <i>HpaII/MspI</i> +ATC
<i>EcoRI</i> +AT/ <i>HpaII/MspI</i> +ACT

3.7. Primers

3.7.1. Design of primers

Primers for all types of quantitative PCR were designed with public domain software Primer3 (http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi) using the parameter settings optimal primer length 20 nt, optimal T_m 60°C, length of PCR fragment 100 – 200 bp, limit of 3 C/Gs in the final five bases (not sequential). BLASTN (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) and Electronic PCR (<http://evry.inra.fr/e-PCR>) analysis to genomic sequences of *A. thaliana* were used to confirm the uniqueness of primer binding sites. Primers specific for C24 sequences were designed based on differences in Col-0 and C24 sequences (part 3.8.).

3.7.2. Calculation of PCR efficiency

PCR efficiencies for each primer pair were experimentally determined to ensure equal amplification of templates derived from accessions Col-0, Cvi, and C24. PCR reactions were performed with serial dilutions of genomic DNA or cDNA (5 ng, 0.5 ng, 0.05 ng, 0.005 ng) of each genotype as templates in three technical replicates and Ct values for the reactions were plotted against the respective template concentration. PCR efficiency was calculated as $E = (10^{-1/\text{slope}} - 1) \times 100$. The regression coefficient between the values obtained in the technical replicates was calculated as R². Only primers that had R² > 0.99 and an efficiency between 90 < x < 110 % were used for quantitative PCR analysis. “Melting curve” determination after quantitative PCR was used to test for specificity of the PCR amplification.

3.8. Cloning and sequencing of C24 alleles of selected genes

For seven genes (Table 6) that were identified as being specifically regulated in Col-0xC24 and reciprocal hybrid seedlings (Maria von Korff Schmiesing and Thomas Altmann, personal communication), sequence differences between accessions Col-0 and C24 were determined. Sequence information for *A. thaliana* accession Col-0 available from <http://www.arabidopsis.org> was used to design primers (Table 5) to amplify the sequences around 700 bp upstream of the translational start point (5' regulatory region) and 700 bp of the transcribed region within the selected genes. The sizes of obtained PCR products were verified by agarose gel electrophoresis. Products derived from accession C24 were cloned into pGEM-T Easy vector (Promega) and sequenced at the sequencing facilities of IPK Gatersleben.

Sequences were processed using the software ‘Editseq’ and aligned using ‘MegAlign’ (Lasergene 6 package). Obtained C24-derived sequences were submitted to the DDBJ/EMBL/GenBank DNA sequence database under accession numbers FJ899135, FJ899136, FJ899137, FJ899138, FJ899139, FJ899140, and FJ899141.

3.9. Allele-specific DNA methylation assay

The sequences of the 5’ regulatory region of seven selected genes in Col-0 and C24 were compared to identify restriction sites suitable for allele-specific DNA methylation assays. This would require a recognition site for a methylation-sensitive restriction enzyme, that is common to both Col-0 and C24, and in addition a second restriction site which is unique to one of the two accessions (Table 6). Genomic DNA extracted from seedlings 6 DAS was incubated with cytosine methylation-sensitive restriction enzymes specific for a common site. As a control, the same amount of DNA was incubated under the same conditions without the restriction enzyme. Restriction enzymes were inactivated and DNA samples were used as templates for PCR with primers indicated in Table 7. After PCR and PCR product purification (QIAquick PCR Purification Kit, Qiagen), aliquots were cleaved with the allele-specific restriction enzymes and size fractionated side by side with uncleaved controls by electrophoresis on 3% Metaphor agarose gels in TBE. Allele-specific methylation levels were estimated from the fluorescence intensity of the corresponding bands after ethidium bromide staining.

3.10. Extraction of plant RNA and cDNA synthesis

Total RNA was isolated from plant tissue according to the Trizol method (Chomczynski and Mackey, 1995). RNA concentrations were determined spectrophotometrically and RNA quality was monitored by electrophoresis in denaturing MOPS-formaldehyde gels (Sambrook J., 2001). Potentially contaminating genomic DNA was removed by incubating with TURBO DNase (Ambion) for 1 hour followed by inactivation via incubation with inactivation reagent (Ambion) for 2 minutes. The absence of DNA contamination in extracted RNA was confirmed by PCR using primers specific to the *ACTIN7* promoter which would specifically generate a product from genomic DNA, but not from *ACTIN7* transcripts (Table 9). From RNA preparations passing these tests, cDNA was synthesized using a RevertAid H Minus first strand cDNA synthesis kit (Fermentas) according to manufacturer’s instructions.

Table 5 Primers used for amplification of candidate genes from accession C24.

Name of primer	Sequences F: Forward R: Reverse
At1g03420	F: 5' AATCCCTTCTTTCATCCTCTC 3'
	R: 5' ATTCCACGTTCCGTTGACTG 3'
At1g29270	F: 5' CTCAGACTTCACCTTTCAACTTC 3'
	R: 5' TGCATCTAACTCCGCGTC 3'
At2g33220 5' UTR	F: 5' TTCACAAGACTGGCACCAAC 3'
	R: 5' TACATGCCCAAGCAAAAGC 3'
At2g33220 3' UTR	F: 5' AGCAATTCTACCGATTCTTCAAG 3'
	R: 5' AGATCAAACTGTGCTACTC 3'
At3g19520 5' UTR	F: 5' CCAAATGGATACCCGCTTAC 3'
	R: 5' GAAGTTTGTCCCCTGAGAATG 3'
At3g19520 3' UTR	F: 5' TTGTAGATGGGCAGATGAAG 3'
	R: 5' AGAAGATCAAATGTAGCATAGTG 3'
At3g25905	F: 5' CAATTCCTTTCACACAATGCTAC 3'
	R: 5' AGGGTTCATACTCATCCTCTTC 3'
At4g29200 5' UTR	F: 5' GCGACAACGAAAGTCTGTTAG 3'
	R: 5' GAGCGTCCTGGTAGAGATG 3'
At4g29200 3' UTR	F: 5' TCATTTGCGAATCCTGGTTTTAG 3'
	R: 5' TCCATCTCGACCGTTCATTATG 3'
At5g15360 5' UTR	F: 5' GAGTCTCCAAAAGCTCTCAAC 3'
	R: 5' TTCATTGTCTCGTCCCCAC 3'
At5g15360 3' UTR	F: 5' CTTTTGTGAGGCTCCCTACG 3'
	R: 5' GCTATCGTGTATTGTTGTCAGTG 3'

Table 6 Restriction sites for allele-specific DNA methylation analysis.

genes number and accession number for C24 sequence	product size	methylation - sensitive restriction enzyme	parent - specific restriction enzyme (DNA methylation assay)	pattern of cleavage
At1g03420 FJ899135	Col-0: 451 bp C24: 461 bp	<i>HpaII</i>	parent - specific primers	-
At1g29270 FJ899136	Col-0, C24: 338 bp	<i>Sau3AI</i>	<i>Tsp509I</i>	Col-0: 3 band; 144bp, 86bp, 108bp. C24: 2 bands; 230bp, 108bp.
At2g33220 FJ899137	Col-0: 360 bp C24: 357 bp	<i>HpyCH4IV</i>	<i>XbaI</i>	Col-0: 1 band; 360 bp C24:2 bands; 143bp, 214bp.
At3g19520 FJ899138	Col-0: 418 bp C24: 427 bp	<i>HpaII</i>	<i>EcoRI</i>	Col-0: 2 bands; 356bp, 62bp. C24: 1 band; 427bp.
At3g25905 FJ899139	Col-0, C24: 292 bp	<i>HpyCH4IV</i>	<i>HpyCH4III</i>	Col-0: 3 bands; 71bp, 11bp, 210bp. C24: 2 bands; 71bp, 221bp.
At4g29200 FJ899140	Col-0, C24: 262 bp	<i>HpaII</i>	<i>MboII</i>	Col-0: 3 bands; 146bp, 65bp, 51bp. C24: 2 bands; 146bp, 116bp.
At5g15360 FJ899141	Col-0: 356 bp C24: 355 bp	<i>HpyCH4IV</i>	<i>DraI</i>	Col-0: 2 bands; 78bp, 278bp. C24: 1 band; 355bp.

Table 7 Primers for allele-specific DNA methylation assays.

name of primer	sequence F: forward R: reverse
Pro-At1g03420 Col-0	F: 5' TTGACTTGAACGATCACTTG 3'
	R: 5' ATCTGTAAAACTGGAAAAAC 3'
Pro-At1g03420 C24	F: 5' TTGCTTGAACGATCACTTG 3'
	R: 5' AACAGAGGCAGAAAAAACTTCAAAG 3'
Pro-At1g29270	F: 5' TATGATGTTTCAATAAGCTACCAG 3'
	R: 5' ATGCTCTGTTTCATGTATACTC 3'
Pro-At2g33220	F: 5' TTGGAGACAAGCCGATAC 3'
	R: 5' TGAGAAAGACAATCAGAAATG 3'
Pro-At3g19520	F: 5' TCTATATCTGAACTGATCCAAAC 3'
	R: 5' ACGCAGGTCTGAAAGATTTATTC 3'
Pro-At3g25905	F: 5' ATTCCAGAACAAGAGACAAAAC 3'
	R: 5' AGAGAGAGAGCAAAGATTGAAG 3'
Pro-At4g29200	F: 5' AATGCACGTTTTAGATTCATG 3'
	R: 5' ATTA CTGGTTAGAGCACAG 3'
Pro-At5g15360	F: 5' ATTGCTTTTGAATAATATATCATCG 3'
	R: 5' AACTGATCATCCTCAAGTC 3'

3.11. Transcript determination

3.11.1 Quantitative RT-PCR

Primers for quantitative RT-PCR were designed based on C24 sequence information generated in this study and publicly available Col-0 sequences to amplify transcripts of seven candidate genes equally from Col-0 and C24 alleles (Table 8). Quantitative RT-PCR was performed in a BioRad iQ5 Real-Time PCR system using a BioRad iQ SYBR Green Supermix kit with primers specific for the selected candidate genes and in addition for housekeeping genes for data calibration (Table 9). C_T values (the cycle at which the fluorescence measured rises above a threshold) obtained for the candidate genes were normalized to the expression levels of housekeeping gene *ACTIN2* (At3g18780, Table 9) according to the formula $\Delta C_T = C_{T, \text{target gene}} - C_{T, \text{Actin2}}$. Normalized expression levels were then related to the values measured in accession Col-0 as a reference genotype in a $\Delta\Delta C_T$ calculation (Livak and Schmittgen, 2001) using the equation $\Delta\Delta C_T = ((C_{T, \text{Target}} - C_{T, \text{Actin2}})_{\text{Col-0 or Col-0xC24 or C24xCol-0 or C24}} - (C_{T, \text{Target}} - C_{T, \text{Actin2}})_{\text{Col-0}})$. Finally,

the magnitude of transcript level changes of candidate genes in different genotypes was calculated as $2^{-(\Delta\Delta CT)}$.

3.11.2 Allele-specific semiquantitative RT-PCR

The primers in Table 8 were used on cDNA preparation for amplification of exemplary genes for various numbers of cycles until PCR product levels clearly visible in agarose gel electrophoresis were obtained. For allele-selective transcript analysis, amplification products were purified from PCR setups and incubated with “allele-specific” restriction enzymes identifying restriction site polymorphisms between Col-0 and C24 (Table 10). After cleavage, the amplification products were size-fractionated by electrophoresis on 3% Metaphor agarose gel and visualised by ethidium bromide staining.

Table 8 Primers used for transcript analysis of candidate genes by quantitative and semiquantitative RT-PCR.

name of primer	sequence F: forward R: reverse
At1g03420-mRNA	F: 5' TTGAAGGTTGCTTTGCTTTTG 3'
	R: 5' AAATCTCTTACCAATGAGCTG 3'
At1g29270-mRNA	F: 5' TTCTATACATAGCCATGCTCAAAC 3'
	R: 5' TTCTCTACCTTCACCTCCTGAAATTG 3'
At2g33220-mRNA	F: 5' TGAAGAAGATGAAAGGTTTGTG 3'
	R: 5' AGGAGCCATTGAAATTTTACCAG 3'
At3g19520-mRNA	F: 5' ACCTTGAGTTACTAAAATACAACG 3'
	R: 5' TTCTCGGTGGTCACTTTCTC 3'
At3g25905-mRNA	F: 5' TTGTCGGTAAAGGGATCAG 3'
	R: 5' AACGAGAAAAACGTAGGTATTTTG 3'
At4g29200-mRNA	F: 5' TATCCTTTTCCGTGGGTATTG 3'
	R: 5' TTTGAATTCCATGAGAGCTTTGAG 3'
At5g15360-mRNA	F: 5' TCGCTGTGGTCGTGAGAG 3'
	R: 5' AAGGTGACAATCACCGTCCTC 3'

Table 9 Primers used for controls in transcript analysis by quantitative and semiquantitative RT-PCR.

name of primer	sequence
	F: forward R: reverse
Actin 2 (At3g18780)	F: TGAGAGATTCAGATGCCCAGAAG 3'
	R: TGGATTCCAGCAGCTTCCAT 3'
Actin 7 Promoter (At5g09810)	F: CTCGTTTCGCTTTCCTTAGTGTTAGCT 3'
	R: AGCGAACGGATCTAGAGACTCACCTTG 3'

Table 10 Candidate genes and restriction sites for allele-specific gene expression analysis.

TAIR genes number and accession number for C24 sequence	product size	parent specific restriction enzyme (allele expression assay)	pattern of cleavage
At1g03420 FJ899135	Col-0, C24: 181bp	<i>BfaI</i>	Col-0: 1 band; 181bp C24: 2 bands; 81bp, 100bp
At1g29270 FJ899136	Col-0, C24: 115bp.	-	-
At2g33220 FJ899137	Col-0, C24: 176bp	<i>HpaII</i>	Col-0: 2 bands; 132 bp, 44bp. C24:1 band; 176bp.
At3g19520 FJ899138	Col-0, C24: 193bp	<i>BsiHKAI</i>	Col-0: 2 bands; 45bp, 148bp C24: 1 band; 193bp
At3g25905 FJ899139	Col-0, C24: 180bp	-	-
At4g29200 FJ899140	Col-0,C24: 199bp	<i>XhoI</i>	Col-0: 1 band; 199 bp C24: 2 bands; 113bp, 86bp
At5g15360 FJ899141	Col-o, C24: 179bp	<i>SalI</i>	Col-0: 2 bands; 134bp, 45bp. C24: 1 band; 179 bp

3.12. Quantitative PCR analysis of DNA from immunoprecipitated chromatin

To validate ChIP on chip results, quantitative PCR was performed on immunoprecipitated (IP) DNA in a Light Cycler 480 device (Roche) using SYBR® Green Master Mix (Roche). The Ct values of regions of interest (Table 13) in the single IP preparations were normalized according to the formula $\Delta C_T = C_{T, \text{IP candidate region}} - C_{T, \text{IP reference region}}$ with the Ct values of reference regions depending on the histone modification analyzed (Table 11). The relative levels in a given candidate region in the four genotypes analyzed were set into relation with the relative levels in Col-0 in a $\Delta\Delta C_t$ calculation using the formula $\Delta\Delta C_T = \Delta C_{T, \text{genotypes}} - \Delta C_{T, \text{Col-0}}$. Finally, the magnitude of changes for any given modification mark in a particular target regions compared with the that target region in Col-0 was calculated as $2^{(-\Delta\Delta C_T)}$. For the calculation of IP efficiency, the Ct value of reference sequences known to be enriched or depleted for the histone modification of interest (Table 11) in the IP preparations and the “no antibody” controls were normalized to the Ct values of input controls using the formula $\Delta C_T = C_{T, \text{IP or “no antibody” controls}} - C_{T, \text{input}}$. Based on this, the percentage of enrichment for that reference region in the analyzed genotypes for a given modification mark were calculated using the formula $(1/ (2^{(\Delta C_T)})) * 100$.

Table 11 Primers for IP efficiency determination and quantitative PCR of reference regions in ChIP-PCR.

application	name of primer	sequences F: Forward R: Reverse
amplification of a regions enriched in H3K4me2	At4g04910-3	F: 5' TCCTCTTAAGCTGGGTCGAA 3'
		R: 5' TTGATCTGGCTATGCTGACG 3'
	At5g13440	F: 5' TCGGCCATCTCTAACCTTAC 3'
		R: 5' GAAGAGGAGTTGCCTCCATC 3'
	AG-3	F: 5' TGAAGTTTTGAGCGATGACG 3'
		R: 5' CCTAAGCCTTTGGAGCAATG 3'
AT4G04920-1	F: 5' ATACCTGAGCGCCCTACTGA	
	R: 5' GGACAAGAAGTGGGAGACCA	
FWA5	F: 5' ATCTTGCCATTACGGCTTTG 3'	
	R: 5' AGCTCTTGAAGCCTCCACAA 3'	
amplification of a regions enriched in H3K27me3	AG-7	F: 5' CAAATTTTCCTGCAGAATGTCA 3'
		R: 5' TCAAGTTGGGCAATCACTCA 3'

3.13. Comparative Genomic Hybridization (CGH) assay

To discriminate potential genomic sequence differences between *A. thaliana* accessions and histone modification differences both of which could cause changes in hybridization signal intensity in ChIP on chip experiments, comparative genomic hybridization (CGH) assays were performed. High quality genomic DNA (gDNA, with an A_{260}/A_{280} ratio of at least 1.8, and an A_{260}/A_{230} ratio of at least 1.9) was isolated from leaves of *A. thaliana* accessions Col-0, C24 and Cvi 21 DAS using the DNeasy Plant Maxi Kit (Qiagen). Extracted gDNA was sheared to fragments with an average size between 500 and 800 bp by sonication (BIORUPTOR; http://www.cosmobio.co.jp/index_e.asp). The efficiency of shearing by sonication was tested by comparative electrophoresis of 50 ng of unsonicated and sonicated DNA in a 1% agarose gel in 1x TAE (Appendix Fig. 1).

3.13.1. CGH experiment using chromosome 4 DNA microarrays

The design of the chromosome 4 tiling microarray used was based on the entire sequence of chromosome 4 of *A. thaliana* (Martienssen et al., 2005). It comprised 21,815 printed tiles, each consisting of a 0.3 - 1.2 kb PCR product derived from sequences along chromosome 4 and in addition several other genomic regions. More than 50% of these tiles represent single copy regions according to BLAST analysis in sequential 100 bp windows against the *A. thaliana* genome sequence. CGH analysis with genomic DNA of accessions Col-0 and C24 was done as a control for ChIP on chip analysis of H3K4me2 (part 3.14.1.). Fragmented gDNA of Col-0 and C24 was directly labeled in a dye-swape manner (labeling exchange of compared genotypes to correct for dye biases) with Cy3 or Cy5 by Klenow DNA polymerase and random hexamers (Lippman et al., 2005) and purified using a QIAquick PCR purification kit (Qiagen). Equal amounts of labeled Col-0 DNA and labeled C24 DNA were mixed with hybridization buffer containing 4X SSC and 0.2% SDS. After hybridization with the probe at 65°C for 12-16 h, microarrays were washed one time with 2X SSC and 0.1% SDS for 5 minutes at 42°C, one time with 1X SSC at ambient temperature for 5 minutes, once with 0.2X SSC at ambient temperature for 5 minutes, and once with 0.05X SSC at ambient temperature for 5 minutes. Arrays were immediately scanned (GenePix 4000A scanner, Axon Instruments) and the spot intensities in the images quantified with GenePix Pro. Data obtained for each array were stored as GenePix reader (.gpr) files and analyzed according to published procedures (Turck et al., 2007).

3.13.2. CGH experiment using *Arabidopsis* whole genome NimbleGen arrays

Arabidopsis whole genome arrays provided by NimbleGen (<http://www.nimblegen.com/products/cgh/index.html>) consist of sets of two slides, each containing 360,000 isothermal probes (tiles) with a length between 50 - 70 bp. When combined in hybridization, they allow whole genome analysis at a resolution of 165 nucleotides. CGH analysis with genomic DNA of accessions Col-0, C24, and Cvi was performed as a control for ChIP on chip analysis of H3K4me2 and H3K27me3 (part 3.14.2.). Sheared genomic DNA (3 μ g) from accessions Col-0, C24, and Cvi was prepared (as described in part 3.13.) and submitted to NimbleGen for labelling and hybridization. Raw data analysis was performed in hidden Markov model approach according to Seifert et al. (2009).

3.14. ChIP on chip experiments

ChIP experiments were performed according to Gendrel et al. (2005) in three biological replicates for each genotype analyzed. Seedlings were cultured until 10 DAS in liquid medium, harvested in 50 ml falcon tubes, and rinsed twice by gentle shaking in 40 ml double distilled water. After the water was removed, seedlings were vacuum-infiltrated with a solution of 1% formaldehyde (freshly prepared in water) at ambient temperature for 15 min to cross-link their chromatin. Cross-linking was stopped by adding glycine to a final concentration of 0.125 M to the reaction setup and vacuum-infiltration for 5 more minutes. Then, seedlings were rinsed with water three times for 5 minutes, frozen in liquid nitrogen, and ground to a fine powder using a mortar and pestle. For chromatin extraction, 1 g tissue powder was suspended in 30 mL of a solution freshly prepared by dissolving one complete protease inhibitor (PIs) tablet (Roche Biochemicals) in 50 ml of 0.4 M sucrose, 10 mM Tris-HCl (pH 8.0), 10 mM MgCl₂, and 5 mM β -mercaptoethanol, 0.1 mM phenylmethanesulfonyl fluoride (PMSF). The resulting slurry was filtered through four layers of Miracloth (Calbiochem). The filtrate was centrifuged for 20 min at 3000 x g at 4°C. The supernatant was removed and the sedimented chromatin was resuspended in 1 mL of a solution freshly prepared by dissolving one complete protease inhibitor (PIs) mini-tablet (Roche) in 10 mL 0.25 M sucrose, 10 mM Tris-HCl (pH 8.0), 10 mM MgCl₂, 1% Triton X-100, 5 mM β -mercaptoethanol, 0.1 mM PMSF. The chromatin was again sedimented by centrifugation for 10 min at 12,000 x g at 4°C and then homogenized by “pipetting” in 300 μ L of a medium freshly prepared by dissolving one complete protease

inhibitor (PIs) mini-tablet (Roche) in 10 mL 1.7 M sucrose, 10 mM Tris-HCl (pH 8.0), 0.15% Triton X-100, 2 mM MgCl₂, and 5 mM β-mercaptoethanol, 0.1 mM phenylmethanesulfonyl fluoride (PMSF). The homogenized chromatin was layered over an equal volume of this buffer and centrifuged for 1 hour at 16,000 x g at 4°C. The resulting chromatin sediment was resuspended in 300 μL of nuclei lysis buffer freshly prepared by dissolving one complete protease inhibitor (PIs) mini-tablet (Roche) in 10 mL 50mM Tris-HCl (pH 8.0), 10 mM EDTA, 1% SDS and was sonicated seven times for 20 seconds on 7% power, pausing 20 seconds between cycles (BIORUPTOR; http://www.cosmobio.co.jp/index_e.asp) to achieve average DNA fragment sizes of approximately 0.8 - 1.5 kb. The success of sonication was tested by electrophoresis of a small (1-2 μL) aliquot of sonicated and non-sonicated chromatin (both reverse-crosslinked and deproteinized by the method indicated below in this paragraph) on a 1% agarose gel in 1X TAE (Appendix Figure 2). Any remaining cellular debris was removed by centrifugation at 16,000 x g for 5 min at 4°C. From each chromatin preparation, 20 μL of initial volume of 300 μL were kept as an input control. For each immuno-precipitation (IP), 100 μL of chromatin solution were diluted 10-fold with CHIP dilution buffer containing 1.1% (v/v) Triton X-100, 1.2 mM EDTA, 16.7 mM Tris-HCl (pH 8.0), 167 mM NaCl, and then incubated with 40 μL of a suspension (in CHIP dilution buffer) of protein-A agarose beads (Roche Biochemicals) for 1 hour at 4°C with rotation. Protein-A agarose beads were removed by centrifugation of chromatin solutions at 12,000 x g for 2 min at 4°C. For immuno-precipitation of histone-DNA complexes, batches of 1000 μL of chromatin preparation were combined with 5 μL of undiluted anti-histone H3K9me2 (cat.-no. 07-441, Upstate), anti-histone H3K4me2 (cat.-no. 07-030, Upstate) or anti-histone H3K27me3 (cat.-no. 07-449, Upstate) antibody preparations followed by incubation overnight at 4°C. Equal volumes of chromatin preparations without addition of any antibody were treated identically in parallel to serve as a “no antibody” controls. The histone-DNA complexes bound to the respective antibodies were then extracted from the solution by incubation with 50 μL of protein A-agarose beads for 1 hour at 4°C. A-agarose beads then sedimented by centrifugation and were washed twice (a first time briefly and a second time for 5 minutes) with 1 ml of low salt wash buffer containing 150 mM NaCl, 0.1% (w/v) SDS, 1% (v/v) TritonX-100, 2 mM EDTA, 20 mM Tris-HCl (pH 8.1), high salt wash buffer containing 500 mM NaCl, 0.1% (w/v) SDS, 1% (v/v) TritonX-100, 2 mM EDTA, 20 mM Tris-HCl (pH 8.1), LiCl wash buffer containing 0.25 M LiCl, 1% (w/v) NP40, 1% (v/v) sodium deoxycholate, 1 mM EDTA, 10 mM Tris (pH 8.0), and finally TE buffer containing 10

mM Tris-HCl (pH 8.0), 1 mM EDTA. The immunocomplexes were eluted from the beads in two steps with 250 μ L of elution buffer containing 1% SDS, 0.1M NaHCO₃, each. The input control samples were filled to a final volume of 500 μ L with the same elution buffer. The obtained immunocomplex preparations and input control samples were submitted to reverse-crosslinking by adding 20 μ L of 5M NaCl followed by an incubation for 6 - 8 hours at 65°C. Residual protein in the samples was degraded by the addition of 2 μ L of 10 g/L proteinase K (Abcam) in 10 mM EDTA, 40 mM Tris (pH 6.5) followed by incubation at 45°C for 1 hr. DNA was purified by phenol/chloroform/isoamyl alcohol extraction and recovered by ethanol precipitation in the presence of 0.3 M sodium acetate (pH 5.2) and 2 μ L of glycogen carrier (10mg/mL). The recovered DNA sediment was washed with 70% EtOH, sedimented again, vacuum-dried, and dissolved in 100 μ L TE buffer (10mM Tris-HCl pH 8, 1mM EDTA). The efficiency of immuno-precipitation in CHIP-preparations compared with input and “no antibody” controls was checked via semiquantitative or quantitative PCR amplification of preselected genomic sequences enriched or depleted for the histone modification of interest compared with (Table 11, Appendix Fig.3).

In contrast to CGH experiments, because the amounts of immunoprecipitated DNA by CHIP were not sufficient to allow the use of directly-labeled probes for microarray hybridization, two rounds of amplification (Round A and B, Lippman et al., 2005, adapted from Bohlander et al., 1992) were performed by PCR using random primers (Table 12) and Sequenase DNA polymerase (US Biochemical, 70775). After 27 - 30 cycles, amplification products were purified using a Qiaquick PCR purification kit (Qiagen) and checked by electrophoresis of 3 μ L of amplified DNA preparation on a 1.2 % agarose gel in 1X TAE buffer.

Table 12 Primers for amplification of immunoprecipitated DNA for microarray hybridization.

name of primer	sequence
Primer A	5' GTTCCAGTCACGATCNNNNNNNNN 3'
Primer B	5' GTTCCAGTCACGATC 3'

3.14.1 ChIP on chip experiments using chromosome 4 tiling microarrays

Chromosome 4 tiling microarrays were used to determine differences in the level and distribution of histone H3K4me2 among *A. thaliana* genotypes Col-0, Col-0xC24, and C24. DNA preparations recovered after ChIP (IP fractions) with antibody specific to H3K4me2 were amplified, differentially labeled, and co-hybridised to microarrays in a dye-swap manner to correct for possible dye biases, in two biological replicates (4 hybridizations total) according to the experimental loop design outlined in Table 19 in Results. Labeling, hybridization and data acquisition were performed as described for CGH (3.13.1.). Data were analyzed according to published procedures (Turck et al., 2007).

3.14.2. ChIP on chip experiments using *Arabidopsis* whole genome NimbleGen arrays

Arabidopsis whole genome NimbleGen arrays were used to determine differences in the level and distribution of histone H3K4me2 among *A. thaliana* genotypes Col-0, Col-0xCvi, and Cvi and of histone H3K27me3 among *A. thaliana* genotypes Col-0, Col-0xC24, and C24, as well as Col-0, Col-0xCvi, and Cvi. Hybridizations were performed in two biological replicates in a dye switch manner to correct for possible dye biases (two arrays, four slides total for each comparison) according to the experimental loop design outlined in Table 19 in Results. For each array hybridization experiment 3 µg (concentration of 300 ng/µL) of amplified IP fraction obtained with antibodies specific for H3K4me2 or H3K27me3, respectively, were sent to the NimbleGen Company for labeling and hybridization processes. Raw data were analyzed according to Seifert et al. (2009) using a hidden Markov model.

3.15. Validation of ChIP on chip results via quantitative ChIP-PCR

Results from ChIP on chip experiments indicating substantial changes of H3K4me2 in particular regions of the *A. thaliana* genome were examined by quantitative ChIP-PCR. Primer pairs were designed for the regions of interest (Table 13) and quantitative PCR was performed for two biological replicates using unamplified immuno-precipitated DNA as template.

Table13 Primers for quantitative PCR examination of ChIP on chip results from chromosome 4 tiling microarrays.

name of primer	sequences F: Forward R: Reverse
Ta06c12	F: 5' CCTGGCTCTCCACCTTACAC 3'
	R: 5' AGCTTGGACAAAACCCACCTG 3'
Ta06d01	F: 5' GCTTTTGGATCAGGCAAAGA 3'
	R: 5' AGCCACCTTCTCCAAGTTGA 3'
Ta06d02	F: 5' TCAGAGGCAACAATGTGAGC 3'
	R: 5' ATGATGATGATGGCGATTAC 3'
Ta14g03	F: 5' TGCCTTGCTTTCCAAGGATG 3'
	R: 5' TCCAACCTCCGCCAAACTTAG 3'
Ta14g04	F: 5' CCTGAGCTGCAGTTTTTCTCC 3'
	R: 5' CGATTGCTTCTCGAATCCTG 3'
Ta14g05	F: 5' TCATGTTCCCCTGGATTTTC 3'
	R: 5' TTCTCTCCCACCTTTGAAGC 3'
Ta52d08	F: 5' AATGGGACGACCAAGAACAG 3'
	R: 5' GGAAGCTCAAGGAGCAACTG 3'
Ta52d09	F: 5' AAACCAGGATTCTGCCAAAC 3'
	R: 5' ACCTCCACGAGCTGACATTC 3'
Tb51a05	F: 5' TAGAATCTGACGAGAGAAGTATCG 3'
	R: 5' TTGCTGGATCTGGACGTATG 3'
Tb51a06	F: 5' TCGTCCTCGTATGTGTTTGC 3'
	R: 5' TACGAGCTCATGACCCAACC 3'

Results

4.1. The level of endoreduplication and nucleolus size does not differ between inbred parents and their intraspecific hybrids

Somatic plant cells can undergo increasing ploidy levels by successive rounds of chromosomal DNA replication without intervening mitoses in a process called endoreduplication. In *A. thaliana*, endoploidy levels ranging from 4C to 32C are commonly found (Sugimoto-Shirasu and Roberts 2003). To test whether the level of endopolyploidization differs between inbred parents and their hybrid progeny, flow cytometric analysis was conducted on three biological replicates of each genotype. The ploidy levels were determined for nuclei from leaves (15 DAS), cotyledons (15 and 10 DAS) or seedlings (6 and 4 DAS) and used to calculate cycle values according to Barow and Meister (2003) as an indicator of endoreduplication (Fig. 4). No differences in endoreduplication were obvious between Col-0, C24 and their reciprocal hybrids.

The nucleolus area correlates with the activity of rRNA genes (Hubbell 1985) and was used to compare the relative rRNA gene activity of inbred parental accessions and their hybrids (Fig. 5). To ensure that the nucleolus areas were determined in nuclei of the same ploidy level, flow-sorted 2C and 4C nuclei from seedlings 6 DAS were used. The average nucleolus area showed a high degree of variability within all samples tested and the nucleolus area of intraspecific hybrids was similar to that of their parent nuclei. Hence, the nucleolus area does not indicate an increased rRNA gene activity in hybrid offspring.

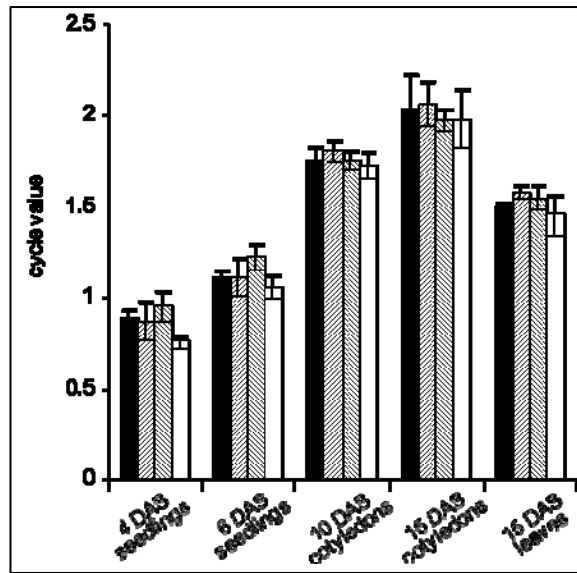


Figure 4 Endopolyploidization levels in inbred parents and their hybrid offspring. Cycle values were determined in accessions Col-0 (black bars) and C24 (white) and their reciprocal hybrid offspring (hatched bars) at different growth stages. Seedlings were analyzed 4 DAS (days after sowing) and 6 DAS, cotyledons 10 DAS and 15 DAS, and leaves 15 DAS. The error bars indicate standard deviations among three biological replicates. Cycle values differ among time points and different organs, but no differences between inbred parents and their hybrid offspring were detected. Approximately 10.000 nuclei per sample were analyzed (Appendix Table 1).

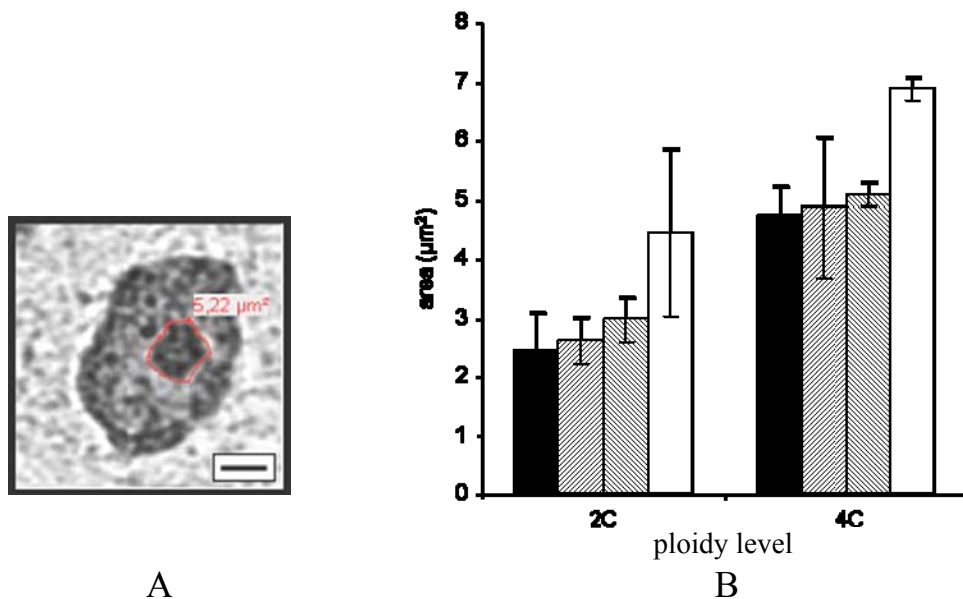


Figure 5 Nucleolus size in inbred parents and their hybrid offspring.

A: Nucleoli (surrounded by a red line) appeared darker than the surrounding nucleoplasm. The scale bar indicates 2 μm . B: Nucleolus area measurements were performed on 2C and 4C nuclei of accessions Col-0 (black bars) and C24 (white bars) and their reciprocal hybrid offspring (hatched bars) (Appendix Table 2). The nucleolus area (μm^2) of nuclei with a defined ploidy level varied among individuals (bars show standard deviation), with nucleolus areas of 4C nuclei being larger than the ones of 2C nuclei.

4.2. The nuclear distribution of DNA methylation and histone H3-methylation at positions K4, K9 and K27 in inbred parents and intraspecific hybrids is similar

DNA methylation (5-methylcytosine, 5mC), histone H3 dimethylation of lysines at position K4 (H3K4me2) or K9 (H3K9me2) and H3 trimethylation of lysine at position K27 (H3K27me3) were selected as representative chromatin marks for microscopic analysis. Indirect immunolabeling was performed on isolated 2C and 4C nuclei and fluorescence signal distribution recorded (Fig. 6). As previously reported by Fuchs et al. (2006) co-localization of immunofluorescence signals with DAPI-stained heterochromatic chromocenters was observed for H3K9me2 and 5mC, while signals for H3K4me2 and H3K27me3 were dispersed in euchromatin of interphase nuclei and excluded from chromocenters. Regardless of which chromatin modification was analyzed, there were no obvious differences in the distribution and intensity of immunofluorescence signals between Col-0, C24 and their reciprocal hybrids.

4.3. A limited gain of DNA methylation appeared in hybrid offspring compared to inbred parents

To investigate possible changes in the overall distribution of DNA methylation in intraspecific hybrid offspring (Col-0xC24 and C24xCol-0) compared with inbred parents (Col-0 and C24) Methylation sensitive amplified polymorphism (MSAP) analysis was used. To address the question of how DNA methylation responds to ageing, DNA methylation patterns of inbred parents and their hybrid offspring harvested 6 and 21 DAS were compared. DNA preparations from pooled seedlings 6 DAS and 21 DAS were digested by differentially methylation-sensitive isoschizomeres (*MspI* and *HpaII*) and amplified with 8 primer pairs (Table 4). In Fig. 7, MSAP patterns are grouped according to their information content in the context of methylation

changes connected with hybrid formation. Approximately 97% (355 of 368 at 6 DAS and 282 of 291 at 21 DAS) of all MSAP signals analyzed showed patterns that did not indicate any methylation changes in hybrid offspring compared with their inbred parents (Fig. 7, pattern category A and D).

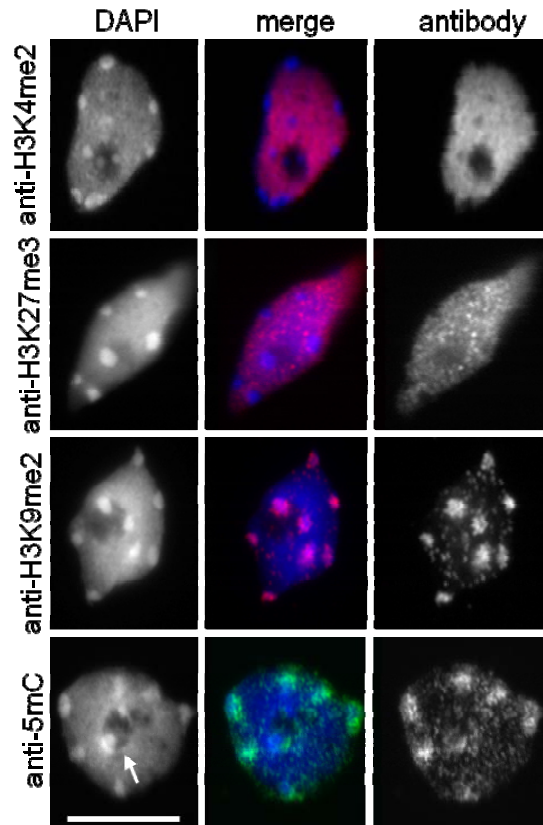


Figure 6 Nuclear distribution of histone H3-methylation at positions K4, K9 and K27 and DNA methylation.

Indirect immunolabeling of histone H3K4me2, H3K9me2, H3K27me3 and 5mC was done on 4C nuclei isolated from seedlings 6 DAS. Images from nuclei of Col-0xC24 hybrid progeny are shown as representatives. Images from Col-0 and C24 inbred parents and C24xCol-0 hybrid progeny showed identical signal patterns. DAPI-bright spots in nuclei represent heterochromatic chromocenters. Note that 5mC and H3K9me2-signals typical for heterochromatin co-localize with chromocenters. Histone H3K4me2-specific signals were found in euchromatic regions only. Histone H3K27me3-specific signals which correspond to silenced genes are dispersed within the euchromatin regions. The dark round area indicated with a white arrow represents a nucleolus. The scale bar indicates 2.5 μ m.

In more than half of these cases, no methylation polymorphism between Col-0 and C24, and the reciprocal hybrids was detected. The patterns indicated either an absence of methylation at either cytosine in all samples (Fig. 7, pattern A1), methylation at the inner cytosine of the CCGG recognition site (Fig. 7, pattern A2) or, for one single signal, methylation at the outer cytosine (Fig. 7, pattern category D). In approximately 10% of cases, the presence of a signal after *MspI* treatment was observed in both parents and their reciprocal hybrids, in combination with the presence of a signal after *HpaII* treatment in all samples except one parent. This indicated that methylation was present at the inner C of the CCGG recognition site in one but not both parents, (Fig. 7, pattern A3 and A4). In the remaining cases, the absence of a signal for one parent after *MspI* and *HpaII* treatments made it impossible to determine whether the restriction site in this parent was either methylated at both the inner and outer cytosine of the recognition site, or whether the restriction site was absent (Fig. 7, pattern A5 to A8). Nevertheless, the absence of methylation at the outer and inner cytosine indicated by cleavage of the site by *MspI* and *HpaII* (Fig. 7, pattern A5 and A6) of the other parent was always inherited in the hybrids. This applied to the absence of methylation at the outer cytosine and presence of methylation at the inner cytosine, which was indicated by cleavage by *MspI*, but not *HpaII* (Fig. 7, pattern A7 and A8). Only approximately 3% (13 of 368 at 6 DAS and 9 of 291 at 21 DAS) of the analyzed MSAP signals showed signs of altered methylation in hybrids (Fig. 7, pattern category B and C). None of the hybrids had a methylation status different from both parents, either by gaining (Fig. 7, pattern B1) or losing (Fig. 7, pattern B2) methylation. Rather, methylation of the inner cytosine of the CCGG recognition site always increased in both reciprocal hybrids, as indicated by the loss of *HpaII* cleavage. In one of the parents, methylation at the same position was present (Fig. 7, pattern B3 and B4), or at least likely to be present (Fig. 7, pattern B5 and B6). In the latter, lack of cleavage by *MspI* and *HpaII* in one of the parents once again could not discriminate between the presence of methylation at the outer and the inner cytosine at these sites and the possible absence of the recognition site (compare also Fig. 7, pattern A5 to A8). Finally, in one case one of the two reciprocal hybrids did lose the *HpaII* cleavage site indicating an increase in methylation at the inner cytosine (Fig. 7, pattern C1). As above, in one of the parents methylation was already present at the same position. Furthermore, the total number of scorable MSAP signals declined from 368 to 291 during plantlets development (Fig. 7, column 6 DAS and 21 DAS), although the same set of primer pairs (Table 4) was used for amplification for both points of time.

pattern		<i>MspI</i>				<i>HpaII</i>				total	
		Col-0	Col-0 x C24	C24x Col-0	C24	Col-0	Col-0 x C24	C24x Col-0	C24	6 DAS	21 DAS
A	1	■	■	■	■	■	■	■	■	167	134
	2	■	■	■	■					50	40
	3	■	■	■	■		■	■	■	25	19
	4	■	■	■	■	■	■	■		5	4
	5		■	■	■		■	■	■	19	19
	6	■	■	■		■	■	■		19	23
	7		■	■	■					16	20
	8	■	■	■						53	22
B	1	■	■	■	■	■			■	0	0
	2	■	■	■	■		■	■		0	0
	3	■	■	■	■	■				0	0
	4	■	■	■	■				■	6	4
	5	■	■	■		■				2	2
	6		■	■	■				■	4	3
C	1	■	■	■	■			■	■	1	0
	2	■	■	■	■	■	■			0	0
D	1					■	■	■	■	1	1
									368	291	

Figure 7 Methylation sensitive amplified polymorphism (MSAP) analysis of inbred parents and their hybrid offspring.

Results from four biological replicates are summarized according to the pattern of presence (black bar) or absence (empty cell) of the corresponding signals in samples derived from inbred parents and their hybrid offspring. The presence of a signal indicates cleavage at a particular *MspI* / *HpaII* site, whereas absence indicates inhibition of cleavage. The presence of a signal in *MspI*- and *HpaII*-treated DNA from the same sample indicates an absence of methylation. The presence of a signal in *MspI*-treated DNA in combination with the absence of a signal in *HpaII*-treated DNA indicates methylation of the inner cytosine of the CCGG recognition site (pattern category A, B, C). The absence of a signal in *MspI*-treated DNA in combination with the presence of a signal in *HpaII*-treated DNA from the same sample indicates hemi-methylation of the outer cytosine of the CCGG recognition site (pattern category D). Methylation at a particular site in only some DNA copies in a sample cannot be detected by this method as it will be masked by a positive signal from the unmethylated copies. The absence of a signal in both *MspI*- and *HpaII*-treated DNA is inconclusive, as this could be caused either by simultaneous methylation at the outer and inner cytosine of the CCGG recognition site, or by the absence of the restriction site (Cervera et al., 2002). Patterns of category B and C indicate changes between parents and hybrids, while patterns of category A and D indicate no changes. The numbers in the columns 6 DAS and 21 DAS indicate the numbers of MSAP signals for the patterns that were determined at these time points.

4.4. No changes in allele-specific methylation of 5′ regulatory regions of selected genes were detected in intraspecific hybrid offspring

To supplement our overall DNA methylation data obtained by the MSAP technique with exemplary gene- and also allele-specific information, seven genes were selected for analysis by methylation-sensitive restriction cleavage followed by PCR. In microarray-based analysis, transcript levels of At1g03420, At1g29270, At2g33220, At3g19520, At3g25905, At4g29200 and At5g15360 had been found to differ significantly between the parental lines Col-0 and C24, and possibly to deviate from mid-parent values in the respective reciprocal hybrids (Maria von Korff Schmiesing and Thomas Altmann, personal communication). Partial sequences of the 5′ regulatory regions and open reading frames of the C24 alleles of these genes were determined. After aligning the sequences with sequence information for Col-0 from public databases, restriction enzymes and PCR primers for allele-specific methylation analysis of the 5′ regulatory regions of these genes were selected (Tables 6 and 7). Analyzes were performed on 6 DAS seedlings of inbred parents and reciprocal hybrids (Fig. 8, Table 14). Methylation levels in the restriction site of a methyl-sensitive enzyme within approximately 300 bp upstream of translation start site as 5′ regulatory element for each genes was investigated in inbred parents and hybrid offspring to address this question whether a link exists between DNA methylation and transcription level. Amplification with primers specific for the 5′ regulatory region of At2g33220 resulted in PCR products of 360 bp for Col-0, 357 bp for C24, and additively, 360 bp and 357 bp, for reciprocal hybrids (Fig. 8, panel A). After cleavage of genomic DNA with *HpyCH4IV*, no PCR products were obtained, indicating that the restriction site is unmethylated in all tested samples (Fig. 8, panel B). Incubation with *XbaI* cleaved the C24-specific PCR product into two fragments of 143 bp and 214 bp, but left the Col-0-specific PCR product un-cleaved. PCR products from reciprocal hybrids showed partial cleavage by *XbaI*, confirming equal amplification of both the Col-0- and the C24-derived allele (Fig. 8, panel C). As in Fig. 8, panel B, PCR products were absent after incubation of genomic DNA with *HpyCH4IV* (Fig. 8, panel D). Therefore, there was no methylation at the *HpyCH4IV* site in the 5′ regulatory region of gene At2g33220.

Amplification with primers specific for the 5′ regulatory region of At4g29200 resulted in PCR products of 262 bp for Col-0, C24, and their reciprocal hybrids (Fig. 8, panel E). After

incubation of genomic DNA with *HpaII*, equal amounts of PCR products were obtained for all tested samples. Nevertheless, compared with Fig. 8E, the amounts of PCR products were lower and indicated the presence of partial, approximately equal methylation at the *HpaII* site in genomic DNA of Col-0, C24, and their reciprocal hybrids (Fig. 8, panel E). Incubation with *MboII* cleaved the Col-0-specific PCR product into three fragments of 146 bp, 65 bp and 51 bp and the C24-specific PCR into two fragments of 146 bp and 116 bp. PCR products from reciprocal hybrids showed partial cleavage by *MboII* with fragment sizes typical for both parents, confirming equal amplification of both the Col-0- and the C24-derived alleles (Fig. 8, panel G). *MboII*-cleavage of the PCR product from genomic DNA resistant to incubation with *HpaII* resulted in band patterns specific for Col-0 and C-24. As in Fig. 8G, *MboII*-treatment resulted in partial cleavage for the reciprocal hybrids. In both hybrids, fragments of 65 bp and 51 bp specific for Col-0 and the fragment of 116 bp specific for C24 were present (Fig. 8, panel H). So, partial methylation at the *HpaII* site is present in approximately equal levels in genomic DNA of Col-0, C24 and their reciprocal hybrids, and there is no shift in the relative methylation levels of Col-0- and C24-derived alleles in the hybrids compared with the inbred parental lines.

Methylation in the 5' regulatory region was detected in genes At4g29200, At1g29270 and At3g19520 on both Col-0- and C24-derived alleles. In the other genes that were analyzed, methylation was absent. Table 14 summarizes the results for all genes that were examined. For all genes, no change of methylation status was observed in the reciprocal hybrids compared with the inbred parental lines.

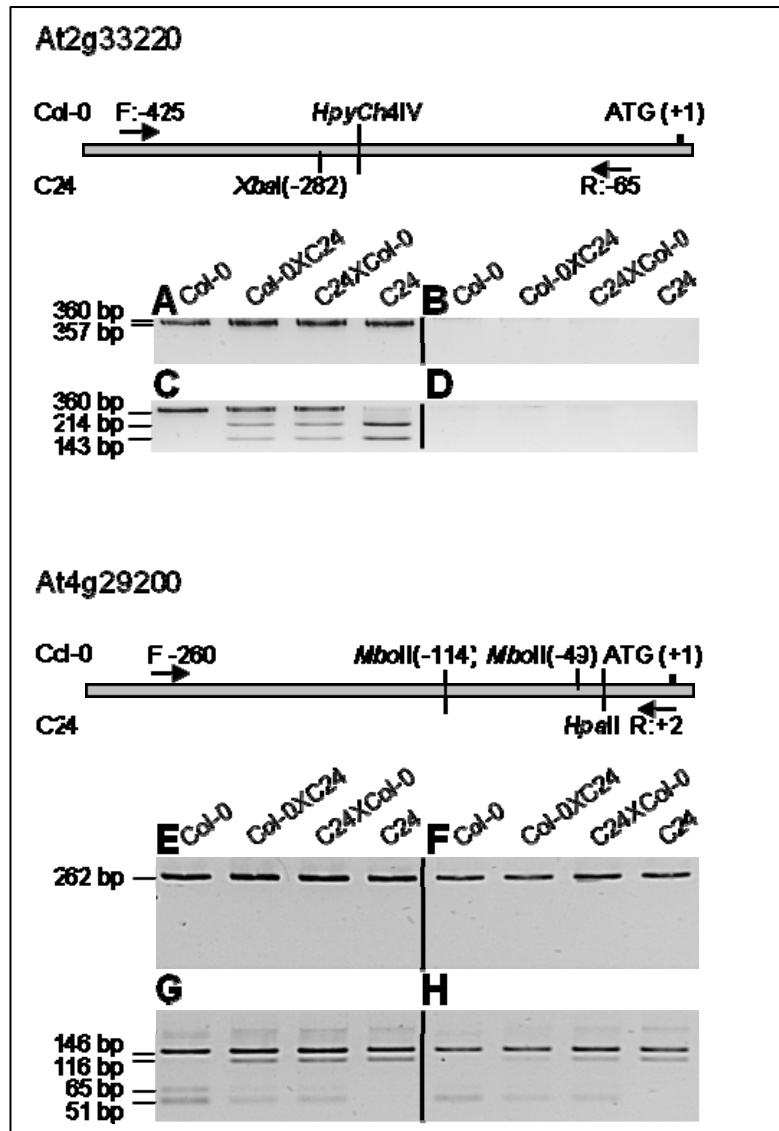


Figure 8 Semiquantitative allele-specific analysis of methylation in 5' regulatory regions of exemplary genes in inbred parents and their hybrid offspring. DNA cytosine methylation in the 5' regulatory region of genes was probed by incubation of genomic DNA from with a methylation-sensitive restriction enzyme followed by PCR with primers flanking the respective restriction site (see gene models). In the presence of methylation at the recognition site, cleavage would be inhibited and a PCR product would be obtained, while in the absence of methylation, cleavage would occur and no PCR product would be obtained. In addition, primers were designed so that PCR products would include a restriction site polymorphic between Col-0 and C24 to discriminate PCR products according to their parental origin. Results are provided in detail for examples of the absence of methylation (At2g33220; A to D) and presence of methylation (At4g29200; E to H) at the methylation-sensitive restriction site. Inverse images of fluorescence signals after agarose gel electrophoresis and ethidium bromide staining are shown.

Table 14 Summary of 5' regulatory region methylation analysis.

gene	methylation in 5' regulatory region
At1g03420	not determined
At1g29270	methylation at both alleles
At2g33220	no methylation
At3g19520	methylation at both alleles
At3g25905	no methylation
At4g29200	methylation at both alleles
At5g15360	no methylation

4.5. Expression of exemplary genes is additive in intraspecific hybrid offspring

Comparative microarray-based gene expression studies (Maria von Korff Schmiesing and Thomas Altmann, personal communication) resulted in the identification of genes with expression levels that were significantly different between the parental lines Col-0 and C24, and possibly deviated from mid-parent values in the respective reciprocal hybrids. Based on these data, an in-depth study of the status of expression levels of 'suspect' genes (At1g03420, At1g29270, At2g33220, At3g19520, At3g25905, At4g29200 and At5g15360) was performed. To reassess the reported "hybrid-regulated" expression levels, quantitative RT-PCR was performed on the respective genes in relation to a housekeeping gene (*ACTIN2*, AT3g18780) as an internal standard on cDNA made from three biological replicates of 6 DAS seedlings. In accordance with the microarray-based data, these genes displayed different expression levels in Col-0 and C24 (Fig. 9). Nevertheless, within the resolution limits of the method, transcript

levels in hybrid offspring were found to be equal to the intermediate of the transcript levels in inbred parental lines.

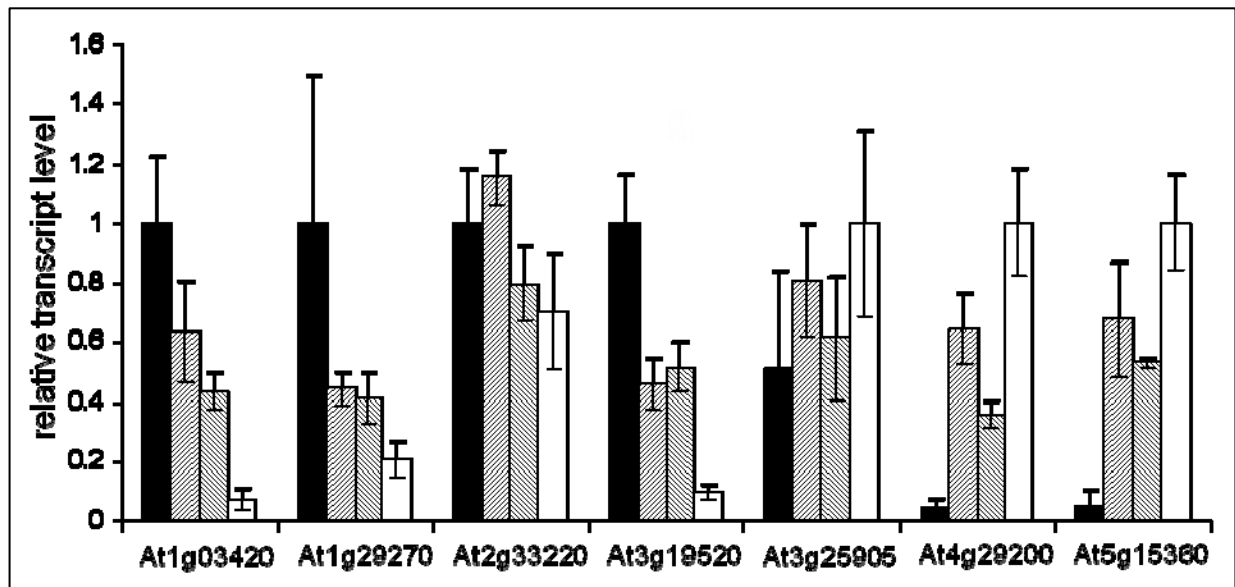


Figure 9 Relative transcript levels of exemplary genes obtained by quantitative RT-PCR. The relative transcript levels are compared with those of *ACTIN2* (At3g18780) as a control gene. Data for Col-0 (black bars) and C24 (grey bars) and their reciprocal hybrid offspring (hatched bars) are shown. To facilitate comparison, the median relative transcript level of the parental line showing the higher RNA abundance was assigned the value 1.0 for each gene. The corresponding values for the relative transcript levels of the other parental line and the reciprocal hybrids were calculated accordingly. Error bars show standard deviations among three biological replicates.

4.6. Allele-specific transcript levels of exemplary genes are unaltered in intraspecific hybrid offspring

Allele-specific expression analysis via cleaved amplified polymorphic sequence – reverse transcription - PCR (CAPS-RT-PCR) was performed in three biological replicates for exemplary genes to investigate whether alleles from Col-0 and C24 alter their expression in intraspecific hybrid offspring. PCR primers (Table 8 and 9) and restriction enzymes for allele-specific transcript analysis (Table 10) were selected according to sequence alignments between Col-0 and C24 genomic sequences. Semiquantitative RT-PCR in combination with restriction enzyme cleavage was performed with cDNA from 6 DAS seedlings. Amplification with

primers specific for the transcript of gene At2g33220 resulted in PCR products of 176 bp for Col-0, C24, and their reciprocal hybrids. Band intensities were similar, indicating about equal transcript levels in all samples (Fig. 10, panel A). Incubation with *HpaII* cleaved the Col-0-derived PCR product into two fragments of 132 bp and 44 bp, but left the C24-derived PCR product uncleaved. PCR products from reciprocal hybrids showed partial cleavage by *HpaII*, with relative band-intensities of the Col-0- and C24-specific fragments almost reflecting the relative band intensities of the respective fragments from the inbred parental lines (Fig. 10, panel B). Amplification with primers specific for transcripts of gene At4g29200 resulted in PCR products of 199 bp for C24 and the reciprocal hybrids, but no PCR product for Col-0. Band intensities were similar for C24 and the reciprocal hybrids, indicating about equal transcript levels in these samples (Fig. 10, panel C). After incubation with *XhoI*, only fragments of 113 bp and 86 bp typical for transcripts of the C24 allele were detected for C24 and the reciprocal hybrids. No uncleaved 199 bp PCR product, which would indicate transcripts from the Col-0 allele, was detected (Fig. 10, panel D). Therefore, transcription states observed in the parental inbred lines, i.e. bi-parental transcription from Col-0 and C24 alleles for gene At2g33220 or mono-parental transcription from the C24 allele, for gene At4g29200, remained unaltered in the reciprocal hybrids.

As summarized in Table 15, transcripts for exemplary genes were detected from both parental alleles (At2g33220, At1g03420), or from only one parent (At4g29200, At5g15360, At3g19520). The relative transcript levels of parental inbred lines and their reciprocal hybrids determined by quantitative RT-PCR (Fig. 9) and semiquantitative RT-PCR (Fig. 10) were largely consistent. No changes in the allele-specific transcript status were observed in the reciprocal hybrids compared with the inbred parental lines.

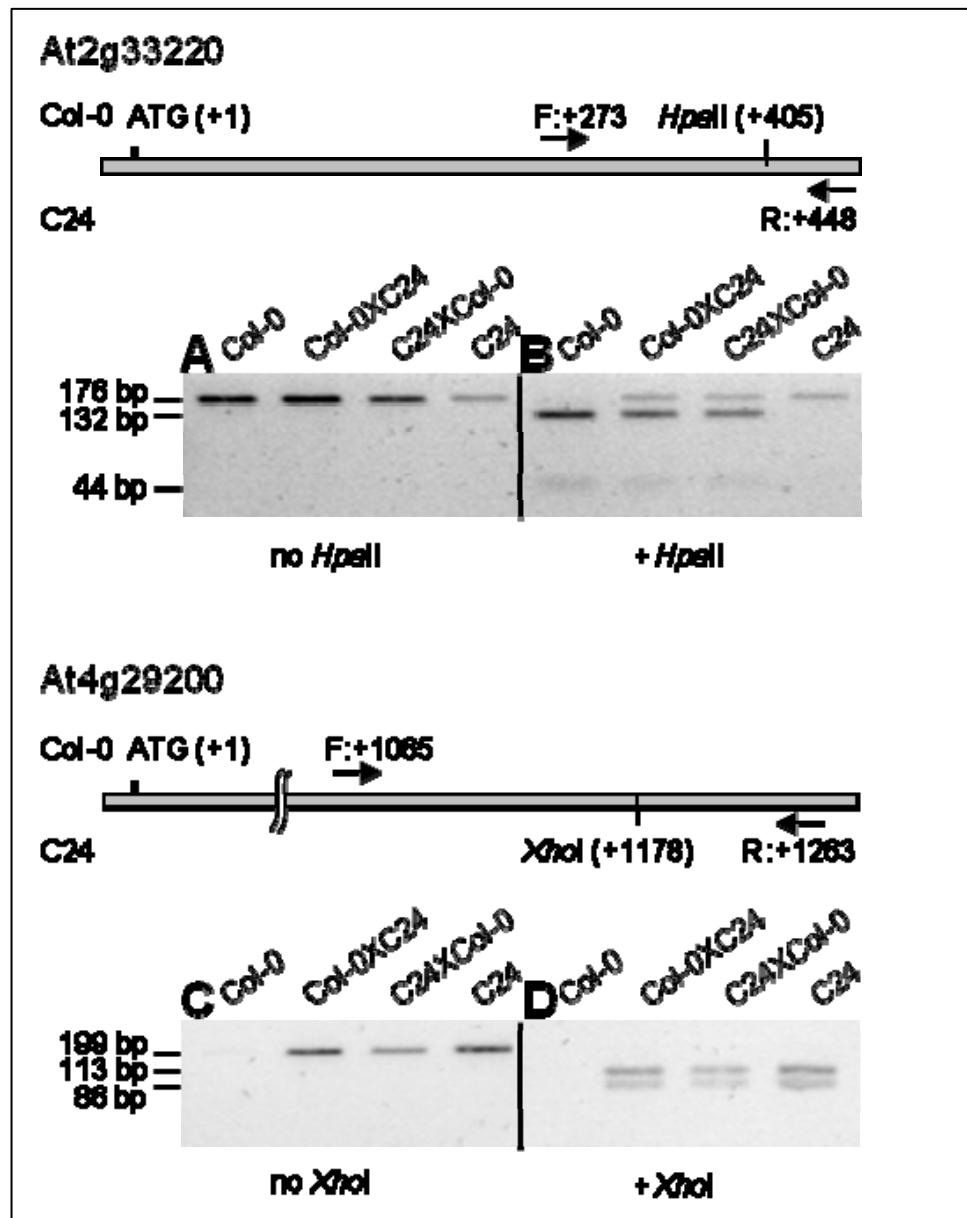


Figure 10 Semiquantitative allele-specific analysis of transcript levels of selected genes in inbred parents and their hybrid offspring.

The primers were designed so that PCR products included a restriction site polymorphic between transcripts of Col-0 and C24 alleles to differentiate PCR products according to their parental origin (see cDNA models). Sample results present details for a case of presence of transcripts in both parental lines (At2g33220; A and B) and transcripts in one parent, but not in the other (At4g29200; C to D). Inverse images of fluorescence signals after agarose gel electrophoresis and ethidium bromide staining are shown.

Table 15 Summary of allele-specific expression assays.

gene	allele expression pattern
At1g03420	from both parents
At1g29270	not determined
At2g33220	from both parents
At3g19520	from one parent, Col-0
At3g25905	not determined
At4g29200	from one parent, C24
At5g15360	from one parent, C24

4.7. Comparative genomic hybridization hybridization (CGH)

Both microarray systems used, chromosome 4 tiling microarrays (Martienssen et al., 2005) and *Arabidopsis* whole genome NimbleGen arrays (<http://www.nimblegen.com/products/cgh/index.html>), were design based on the reference sequence of *A. thaliana* accession Col-0. The determination of sequence differences between accessions Col-0, and C24 and Cvi used as crossing partners for intraspecific hybridization, by comparative genome hybridization (CGH) was therefore a prerequisite for interpretation any chromatin immuno-precipitation - microarray hybridization (ChIP on chip) experiment involving accessions other than Col-0. Otherwise, it would not have been clear whether differences observed in hybridization signal intensities were due to differences in genome sequences or in histone modifications. Differences in signal ratios observed in CGH experiments indicate copy number or partial sequence differences among the compared genotypes (Table 16).




Table 16 Terms for the interpreting CGH and ChIP on chip results.

type of experiment	term	description	interpretation
CGH	CGH polymorphic tile	tiles with significant signal ratio between inbred parents in CGH experiments	genomic regions differing in sequence between compared accessions
ChIP on chip for H3K4me2 or H3K27me3	H3K4me2 polymorphic tile or H3K27me3 polymorphic tile	tiles with significant signal ratio between inbred parents in ChIP on chip experiments for H3K4me2 or H3K27me3	regions with similar sequences, but different levels of H3K4me2 or H3K27me3 between compared accessions
ChIP on chip for H3K4me2 or H3K27me3	H3K4me2 hybridization responsive tile or H3K27me3 hybridization responsive tile	tiles with the indication of a response to intraspecific hybridization	regions with levels of H3K4me2 or H3K27me3 in hybrid offspring exceeding or lower than the levels of both inbred parents

Sequences of Col-0 and C24 were compared using the chromosome 4 tiling microarrays in a dye-swap manner in two technical replicates. Sequences of the accessions C24 and Cvi were compared with the sequence of Col-0 as the reference genotype using the *Arabidopsis* whole genome NimbleGen arrays. Comparison between Col-0 and C24 was done with labeling in one direction (Col-Cy5 + C24-Cy3), comparison between Col-0 and Cvi was done with reciprocal labeling in a dye-swap manner for one technical replicate (2 array hybridizations, 4 slides in total, Table 17).

Data of whole genome CGH experiments were analyzed by the hidden Markov model according to Seifert et al. (2009). Groups of contiguous tiles with significantly different signal ratios for each set of compared accessions were interpreted as representing regions of sequence polymorphism between the reference genotype Col-0 and C24 or Cvi, respectively.

Table 17 Design of CGH experiments to compare inbred parental lines.

accessions involved	type of chip	design of experiment
Col-0 and C24	chromosome 4 tiling microarray	dye-swap* in two technical replicates Col-0  C24
Col-0 and C24	<i>Arabidopsis</i> whole genome NimbleGen array	one direction labeling in one technical replicate Col-0  C24
Col-0 and Cvi	<i>Arabidopsis</i> whole genome NimbleGen array	dye-swap in one technical replicate Col-0  Cvi

*: Each arrow represents a hybridization, with its start and head denoting Cy5- and Cy3-labeling, respectively. Two arrows in opposite orientation indicate a dye-swap experiment.

4.7.1. CGH revealed sequence polymorphisms between *A. thaliana* accessions Col-0 and C24

4.7.1.1 Chromosome 4 tiling microarrays

Overall, 6.43% of tiles present on chromosome 4 tiling microarrays (1368 of 21,296 tiles total) revealed signal ratios of Col-0-derived to C24-derived signals that significantly differed from the mean signal ratio and were therefore considered CGH polymorphic tiles. Out of these 1368, 1139 (5.35% of the total tiles) had a positive value (of the \log_2 ratio between Col-0- and C24-derived signal) that was indicative for sequences with copy numbers in C24 being lower than in Col-0. In contrast, 229 CGH polymorphic tiles (1.08% of total tiles) had a negative value indicative for sequences with a higher copy number in C24 (Fig. 11). As the used microarrays were based on Col-0 sequences, any sequences which solely exist in C24 but not Col-0 could not be detected by this method.

Sequence polymorphisms were not evenly distributed across chromosome 4, but clustered around the heterochromatic knob and pericentromeric regions of chromosome 4. Overall, 15.6% and 12.1% of tiles covering the heterochromatic knob and the pericentromeric regions, respectively, indicated sequence polymorphisms between accessions Col-0 and C24. In contrast, only 7% and 4.5% of tiles in euchromatic regions of the short and long arm of

chromosome 4, respectively, displayed significant differences between both accessions (Fig. 11).

Of the 1368 CGH polymorphic tiles, 345 tiles were “singletons”. The remaining tiles could be categorized into 254 domains with two to 59 contiguous tiles (Table 18). Most domains had a length of less than five kilo base pairs (Fig. 12). The longest domain, which included 59 contiguous CGH polymorphic tiles with lower copy number in C24 (from position 11,045,830 to 11,104,948), coincides in the sequence of Col-0 with several copies of receptor-like kinase-related proteins (AT4G20530, AT4G20540, AT4G20550, AT4G20560, AT4G20570, AT4G20580, AT4G20590, AT4G20600, AT4G20610, AT4G20620, AT4G20630, AT4G20640, AT4G20650, and AT4G20680), a putative pollen coat receptor kinase (AT4G20670) and several hypothetical proteins (AT4G20690, AT4G20700, AT4G20710 and AT4G20715). The complete list of genes localized in domains with a higher or lower copy number in C24 compared with Col-0 are listed in Excel File 1.

To address the question whether particular sequence types are more prone to CGH-detectable polymorphisms than others, tiles were classified according to their underlying sequence annotation. A comparison between the total set of tiles on the used microarray and CGH polymorphic tiles revealed that most sequence polymorphisms between Col-0 and C24 occurred in sequences annotated as transposable elements, while genic region seemed more conserved (Fig. 13).

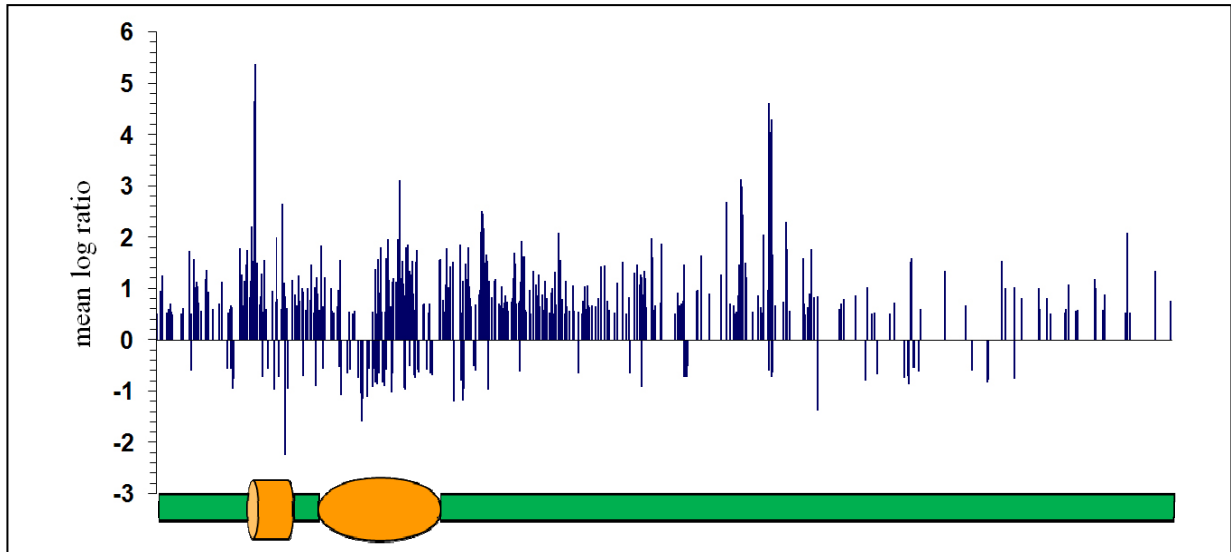


Figure 11 Distribution of CGH detected sequence polymorphisms between Col-0 and C24 on chromosome 4.

The location of CGH polymorphic tiles between Col-0 and C24 along chromosome 4 of *A. thaliana* showed an uneven distribution. Sequence differences were clustered around the heterochromatic pericentromere (orange oval) and the knob region (orange cylinder). Mean \log_2 signal ratios for each tile were calculated by dividing signal intensity of Col-0 by the one of C24. Tiles with a positive value represent sequences which have a lower copy number or are absent in C24. Tiles with a negative value represent sequences with a higher copy number in C24.

Table 18 Number of domains with at least two contiguous CGH polymorphic tiles in CGH experiment using chromosome 4 tiling microarrays.

successive tiles per domain	2	3	4	5	6	7	8	9	10	11	13	16	18	19	59
number of domain	109	47	29	26	15	11	5	2	1	3	1	2	1	1	1

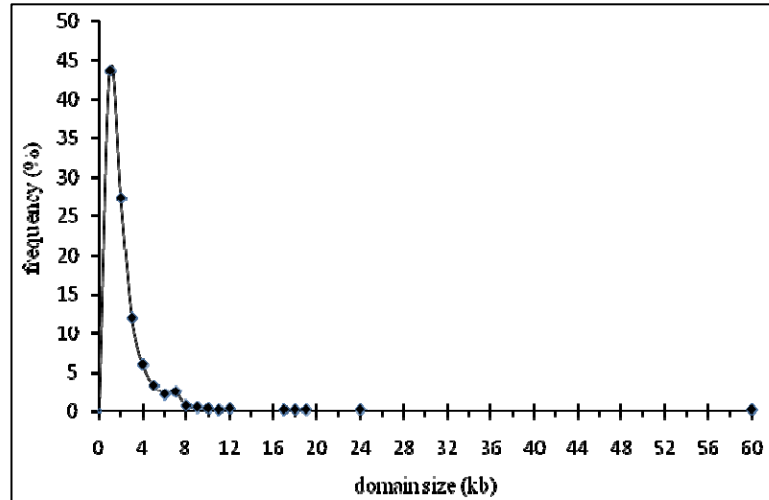


Figure 12 Length distribution of domains on chromosome 4 containing contiguous CGH polymorphic tiles between accessions Col-0 and C24. Domain sizes were calculated by subtracting the respective start and stop positions of the first and last CGH polymorphic tile.

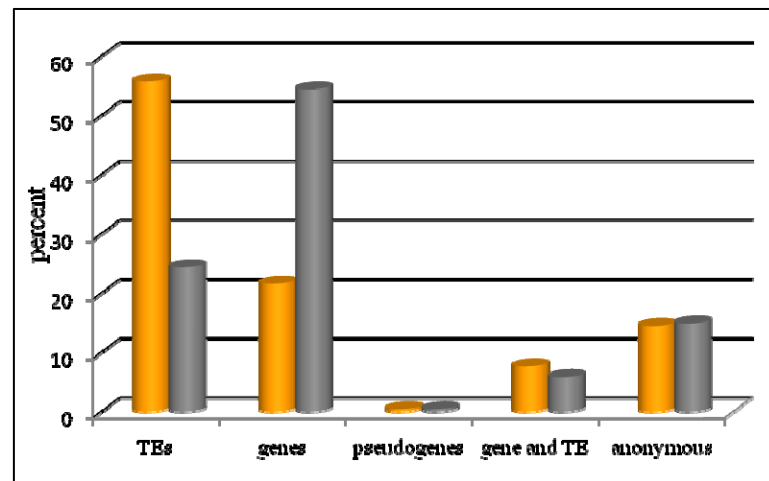


Figure 13 Sequence type classification of CGH polymorphic tiles on chromosome 4 between Col-0 and C24. Grey bars indicate the tiles in a particular class as a percentage of the total tiles on the chromosome 4 tiling microarray. Hence, genes account for 54%, transposable elements (TEs) for 24% and anonymous sequences for 15% of the total tiles. Orange bars represent the percentage CGH polymorphic tiles between Col-0 and C24 of a particular classification; transposable elements account for 56%, genes for 22% and anonymous sequences for 14%.

4.7.1.2. *Arabidopsis* whole genome NimbleGen arrays

Out of 717,235 tiles included in *Arabidopsis* whole genome NimbleGen arrays, 39,776 (5.54% of total tiles) distributed over all chromosomes were identified as CGH polymorphic tiles between Col-0 and C24. 2,562 of these (0.35% of total tiles) covered sequences with probably a higher copy number in C24; 885 were singletons, and the remaining tiles could be assigned to domains with two to 54 contiguous tiles. The average size of the polymorphic regions with a higher copy number in C24 was 1.9 kb, with most (90%) of the polymorphic regions being less than 4 kb long. Nevertheless, several domains had larger sizes of up to 9 kb, containing 54 contiguous tiles (Fig. 14A). The genes localized in the domains with a higher copy number in C24 compared with Col-0 are listed in Excel File 2.

CGH polymorphic tiles comprised 37,214 (5.18%) of the total number of tiles and covered sequences with a lower copy number in C24 than Col-0. They can be classified into 6996 singletons and domains with two to 372 contiguous tiles. The average size of the polymorphic regions with lower copy number in C24 was 3.9 kb. The vast majority of these domains (90%) had a length of less than 8 kb, but some larger domains including one of approximately 57 kb were also detected (Fig. 14A). The genes present in the domains of sequences with lower copy number in C24 compared to Col-0 are listed in Excel File 2.

To address the question whether sequence polymorphisms are associated with particular sequence classes, CGH polymorphic tiles were classified according to their sequence annotation. Equivalent numbers of randomly picked tiles were classified using the same annotation categories to serve as controls. Random selection was repeated 500 times and the numbers of tiles in each category were calculated. The counts in each category for the random selection were then used to determine a p-value for the counts of the original selection. The p-value is the relative frequency of how many times the 500 counts of a category in the random selection are equal or greater than (equal or smaller than) the corresponding count in the original selection. Based on a p-value cutoff of 0.05, categories containing significantly more (or significantly less) tiles in the original data than in the random selection have been selected. Results revealed that sequence polymorphisms mainly affect transposable elements while genic regions, and 5' and 3' untranslated regions (UTRs) are more conserved between Col-0 and C24 (Fig. 15A). Ontology categorization of the genes associated with CGH polymorphic regions

between Col-0 and C24 indicated that they preferentially involve genes with a function in signal transduction or response to stress (Fig. 16A). For the above interpretation of CGH data, only contiguous CGH polymorphic tiles which were consistently found in both directions of dye-swap experiments were included.

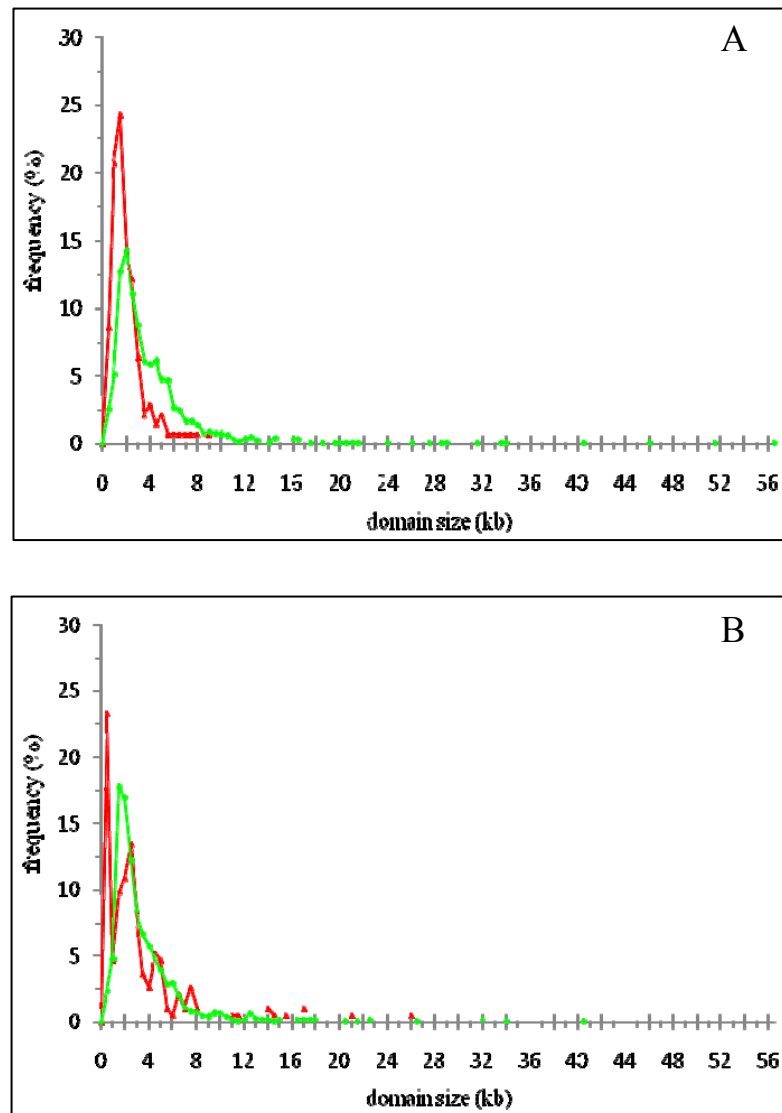


Figure 14 Size distributions of CGH polymorphic domains between Col-0, and C24 or Cvi. Domain sizes were calculated by subtracting the sequence position of the start of the first and the end of the last contiguous CGH polymorphic tile on the respective chromosome. The frequency of domain size classes covering sequences with higher (in red) or lower copy number (in green) in C24 (A) or Cvi (B) compared with Col-0 are depicted.

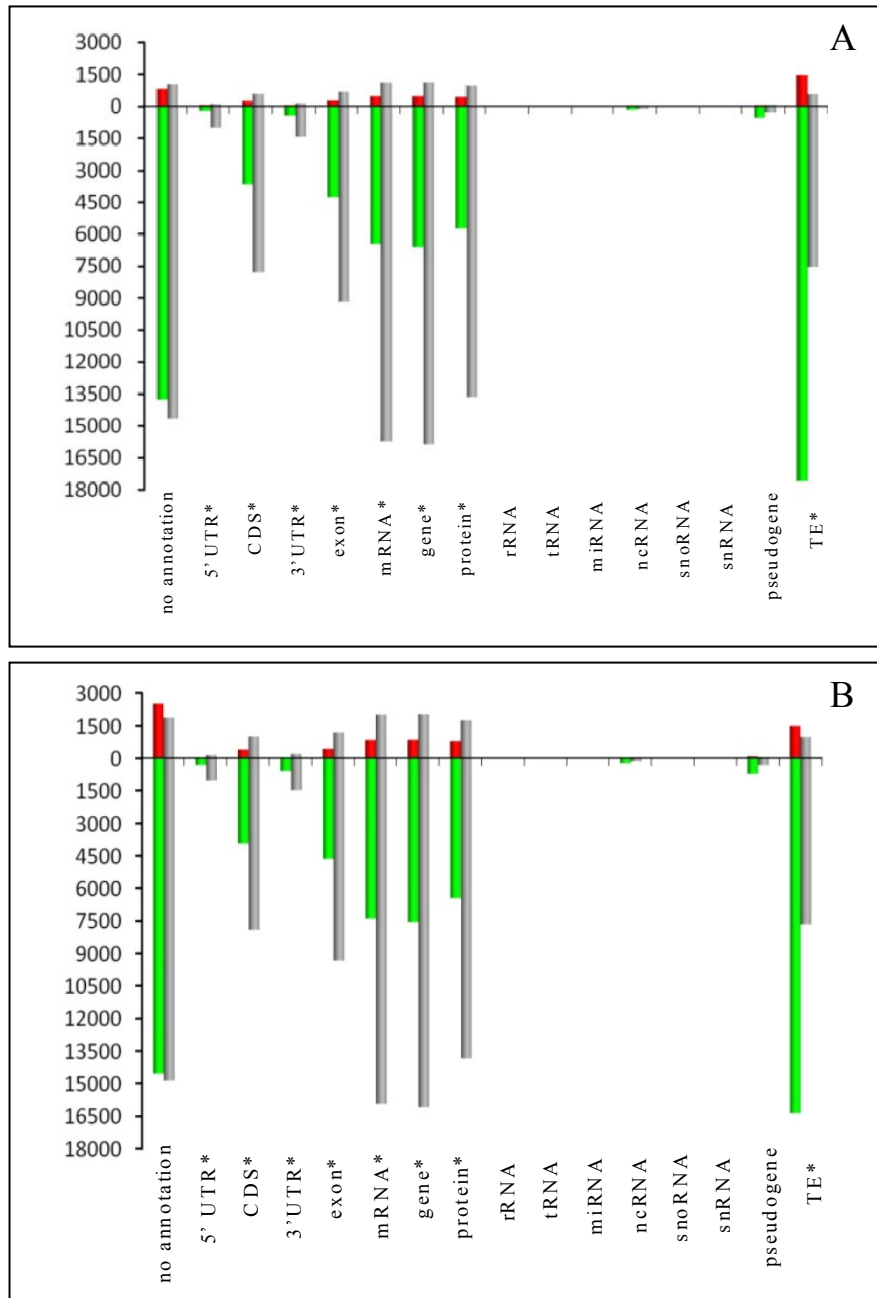


Figure 15 Sequence type classifications of CGH polymorphic tiles of Col-0 versus C24 and Cvi using *Arabidopsis* whole genome NimbleGen arrays. CGH polymorphic tiles between Col-0, C24 and Cvi were classified based on the annotation of their underlying sequences. Red and green bars represent tiles with higher and lower copy number in C24 (A) and Cvi (B) relative to Col-0, respectively. Grey bars indicate random count controls. Cases with significant differences between CGH data (green or red bars) and random counts (grey bars) as indicated by a p-value of 0.05 are marked with an asterisk.

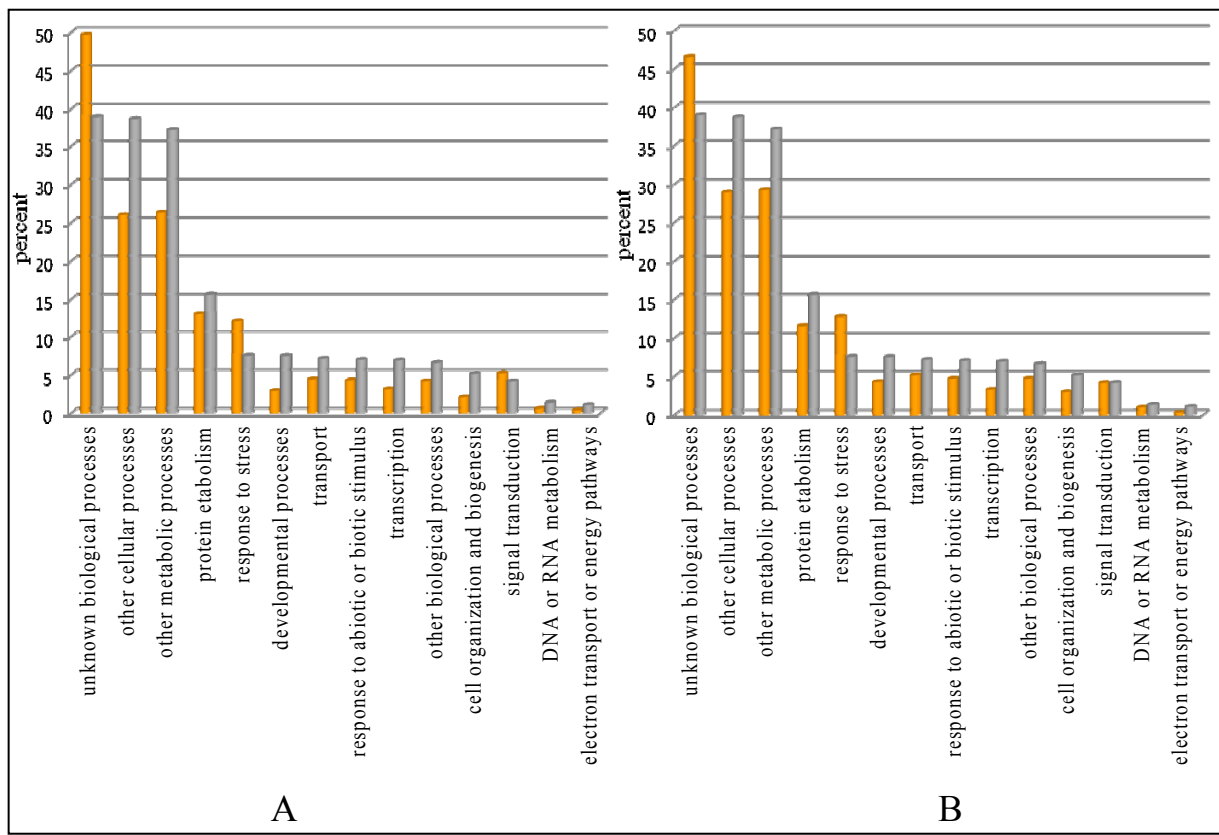


Figure 16 Ontology classification of genes localized in CGH polymorphic regions. A: Col-0 versus C24, B: Col-0 versus Cvi. Orange bars indicate frequency of ontology classes of genes in CGH polymorphic regions. Grey bars represent the random frequencies of genes in the same classification based on the whole genome categorization of *A. thaliana*.

A comparison between CGH polymorphic regions identified by chromosome 4 tiling microarray and *Arabidopsis* whole genome NimbleGen array analysis revealed a considerable overlap (Fig. 17). Out of 433 genes located in CGH polymorphic regions detected by chromosome 4 tiling microarrays and 220 genes located in CGH polymorphic regions on chromosome4 detected by *Arabidopsis* whole genome NimbleGen arrays, 132 genes (25%) were common. Similarly, out of 652 transposable elements located in CGH polymorphic regions detected by chromosome 4 tiling microarrays and 637 transposable elements in CGH polymorphic regions on chromosome 4 detected by *Arabidopsis* whole genome NimbleGen arrays, 396 (45%) were common. The overlap of less than 50% between genes and TEs identified by the different arrays could be explained by the difference in the length of tiles in

chromosome 4 tiling microarrays and *Arabidopsis* whole genome NimbleGen arrays (one kb versus 50-70 bp, respectively). Nevertheless, the general pattern of sequence polymorphisms found by the two arrays is similar. For the subsequent interpretation of ChIP on chip data of Col-0, C24 and their intraspecific hybrids using *Arabidopsis* whole genome NimbleGen arrays, all CGH polymorphic tiles (including singletons and those that were only detected in one out of two CGH technical replicates) were excluded.

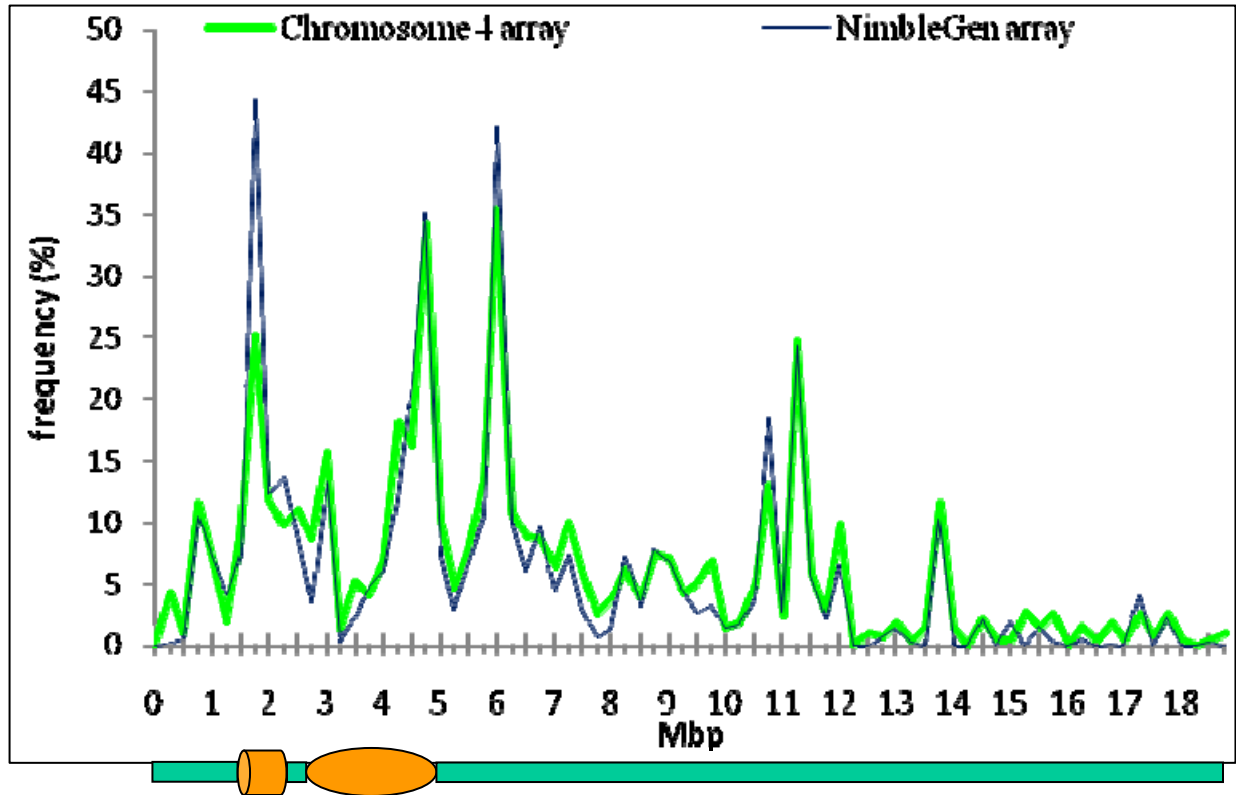


Figure 17 CGH polymorphic regions of chromosome 4 detected with different hybridization array platforms.

CGH polymorphisms between Col-0 and C24 were detected by using either a whole genome NimbleGen array (blue line) or a chromosome 4 tiling array (green line). Frequencies of CGH polymorphic tiles in 250 kb intervals are plotted against the sequence of chromosome 4. The heterochromatic pericentromere is indicated by an orange oval and the heterochromatic knob region by an orange cylinder.

4.7.2. CGH revealed sequence polymorphisms between *A. thaliana* accessions Col-0 and Cvi

CGH experiments with genomic DNA of *A. thaliana* accessions Col-0 and Cvi were only performed on *Arabidopsis* whole genome NimbleGen arrays in a dye swap manner with one technical replicate (Table 17). Out of 717235 tiles on the microarray, 43141 tiles (6.01 % of total tiles) showed significant CGH polymorphisms for Col-0 labeled with Cy5 and Cvi labeled with Cy3; 45570 tiles (6.35 % of total tiles) showed significant CGH polymorphisms for the inverse labeling. Comparison of both technical replicates revealed 97.43% identity of CGH polymorphism in both experiments (Fig. 18). Hence, CGH analysis is highly reproducible. For interpreting CGH data only contiguous CGH polymorphic tiles that were consistently found in both directions of dye-swap experiments were included.

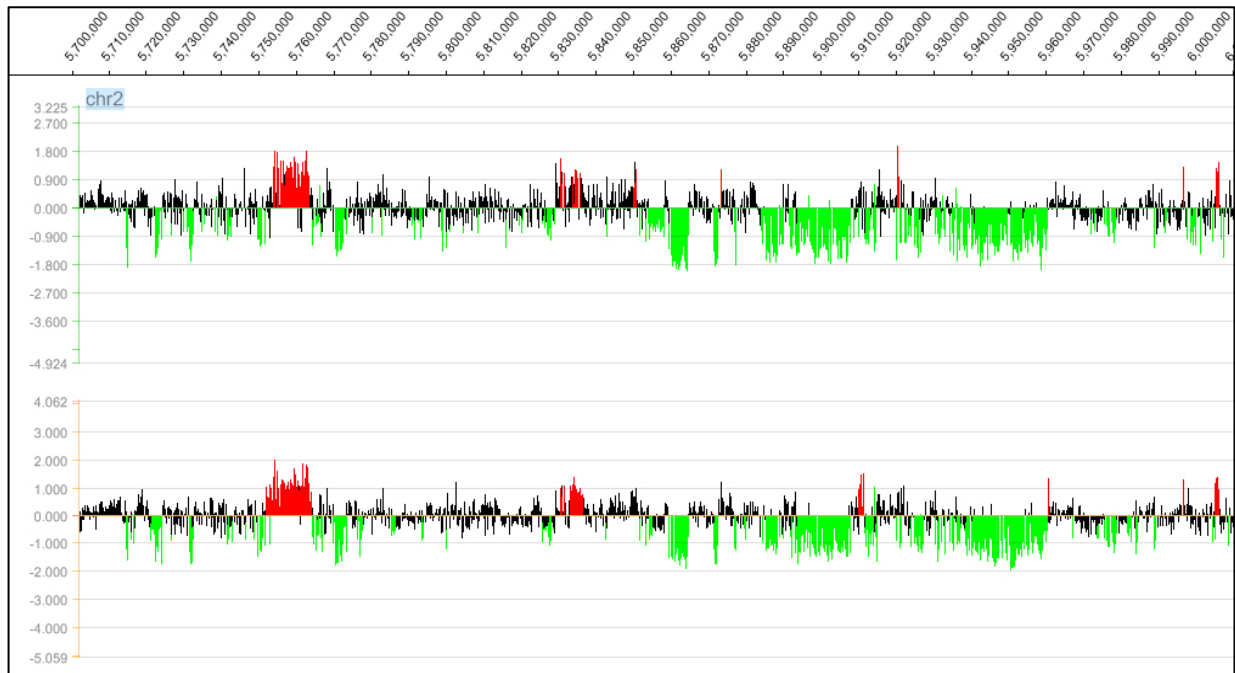


Figure 18 Identification of CGH polymorphic regions (e.g. Col-0 versus Cvi) in dye-swap experiments is highly reproducible.

CGH signals along a 300 kb-long region of chromosome 2 (chr2) are displayed using SignalMap software (Roche NimbleGen). CGH polymorphic tiles covering sequences with higher (in red) or lower copy number (in green) in Cvi compared with Col-0 are depicted. Black bars indicate tiles without significant CGH polymorphisms. Upper panel: Cvi - Cy3 / Col-0 - Cy5 labeling. Lower panel: Cvi - Cy5 / Col-0 - Cy3 labeling. The vertical axes show the mean \log_2 signal ratios of tile in the respective microarray read outs.

From 43,141 CGH polymorphic tiles found in the Col-0 - Cy5 / Cvi - Cy3 hybridization setup, 4,785 tiles (0.67% of 717,235 tiles in total) covered sequences with a probable higher copy number in Cvi compared with Col-0. Of the total, 1,122 were singletons, while the remaining tiles could be categorized in domains with two to 169 tiles. The average size of CGH polymorphic regions was 3 kb. The size distribution of domains revealed that the majority (90%) of domains was less than 6.5 kb long. The largest domain had a size of 26 kb (Fig. 14B). Of the total number of tiles, 38,383 (5.35%) covered sequences with a lower copy number in Cvi compared with Col-0. These could be assigned into 7,094 singletons and domains with two to 369 tiles. The average size of the domains was 3.3 kb. The size distribution of the domains showed that most domains (90%) had a length of less than 6 kb. Nevertheless, several larger domains including one with a size of ~40.5 kb were found (Fig. 14B). A list of genes localising to the CGH polymorphic regions between Col-0 and Cvi is provided in Excel File 3.

CGH polymorphic tiles in Cvi compared with Col-0 were categorized based on the annotation of their associated sequences to see whether particular sequence classes are enriched. Transposable elements were significantly more common than expected for a random distribution. Genes, and 5' and 3' UTRs were clearly underrepresented compared with randomly selected sequences (Fig. 15B). Ontology categorization of genes associated with CGH polymorphic regions between Col-0 and Cvi indicated that they preferentially involve genes with a function in signal transduction or response to stress (Fig. 16B).

For the subsequent analysis of ChIP on chip data of Col-0, Cvi, and their intraspecific hybrids, all CGH polymorphic tiles (including singletons and those tiles that were detected only in one out of two CGH technical replicates) were excluded.

Among CGH polymorphic tiles identified using *Arabidopsis* whole genome NimbleGen arrays for Col-0 versus Cvi and Col-0 versus C24, 93% (93.15% for the first and 92.84% for the second technical replicate) were common. However, the direction of the polymorphism was not necessarily identical. The common CGH polymorphic tiles included sequences with a higher or lower copy number in C24 than Col-0 and a lower or higher copy number, respectively, in Cvi compared with Col-0.

In conclusion, results of CGH experiments using two different microarray platforms were considerably reproducible. CGH data revealed sequence polymorphisms among the accessions of *A. thaliana* in both combinations (Col-0 versus C24 and Col-0 versus Cvi). These polymorphisms occur in all type of sequences, but transposable elements showed a higher level of polymorphism than genic sequences.

4.8. Analysis of histone H3K4me2 and H3K27me3 distribution between Col-0, C24, Cvi, and their intraspecific hybrids

ChIP on chip is a powerful technique to determine interactions between DNA sequences and particular proteins of interest in a genome-wide scale. It combines chromatin immunoprecipitation, which can capture DNA/protein complexes by using specific antibodies, with DNA array hybridization, which then can identify the DNA sequences that are associated with the protein, or protein modification, which is targeted by the antibody. The levels and distribution of histone H3K4 and H3K27 methylation were comparatively analyzed between *A. thaliana* accessions Col-0, Cvi and C24 using two different DNA array platforms, chromosome 4 tiling arrays and *Arabidopsis* whole genome NimbleGen arrays. The mode of inheritance of these modifications after intraspecific hybridization was analyzed by comparison of histone modification patterns in intraspecific hybrids (Col-0xCvi and Col-0xC24) with the patterns in their corresponding inbred parents.

By ChIP on chip, the distribution of histone H3K4me2 was investigated along chromosome 4 in inbred parents (Col-0 and C24) and their hybrid progeny (Col-0xC24). After chromatin immunoprecipitation (ChIP) with a histone H3K4me2-specific antibody, DNA of H3K4me2-enriched chromatin was isolated and labelled to obtain hybridization probes for chromosome 4 tiling arrays. Hybridization was performed with immuno-precipitated (IP) fractions from inbred parents and hybrid offspring in a dye swap manner for two biological replicates (in total 12 hybridization experiments, Table 19).

ChIP on chip experiments at the whole genome level were performed using antibodies specific for histone H3K4me2 and H3K27me3 in combination with *Arabidopsis* whole genome

NimbleGen arrays as a platform. The distribution of H3K4me2 was studied in Col-0, Cvi and their intraspecific hybrid Col-0xCvi. The distribution of histone H3K27me3 was investigated in Col-0, Cvi, C24, and their intraspecific hybrids Col-0xCvi and Col-0xC24. The experiments were conducted in a dye switch manner for two biological replicates to reduce possible dye effects (in total 6 hybridization experiments with 12 slides, Table 19). Data were analyzed using a hidden Markov model approach according to Seifert et al. (2009) to identify successive tiles with differential signal ratios between compared IPs. Tiles polymorphic for CGH between the inbred parental lines were excluded from ChIP on chip data interpretation. In contrast to what was observed in CGH analysis, the data from biological replicates in ChIP on chip analysis showed a high degree of variability. Therefore, data from each biological replicate were analyzed separately.

4.8.1. Natural differences in the distribution of histone H3K4me2 along chromosome 4 of Col-0 and C24

H3K4me2 polymorphic tiles were defined as those that produced significantly different signal ratios between IPs prepared from material of different genotypes. 4.6% of total tiles (984 of 21,296 tiles on chromosome 4 tiling microarrays) were selected as H3K4me2 polymorphic tiles between Col-0 and C24 (Fig. 19). To exclude tiles for which the observed differential signal ratios were caused by DNA sequence differences rather than by differences in the level of histone H3K4 dimethylation associated with their target sequences, CGH polymorphic tiles between Col-0 and C24 were excluded from the interpretation of the results (Fig. 19). After exclusion, 548 tiles (2.6% of total tiles) were identified to display significant differences in the level of histone H3K4 dimethylation between inbred parental lines Col-0 and C24 (Fig. 19).

H3K4me2 polymorphisms were evenly distributed along chromosome 4. Approximately equal numbers of tiles had a positive or a negative value, indicating a lower or higher level of H3K4me2 in C24 compared with Col-0 (Fig. 20). Out of 548 H3K4me2 polymorphic tiles, 367 tiles were singletons, while the remaining 201 reminding could be categorized into domains with two to maximum four contiguous tiles (Table 20). The majority of domains had a size of 2 - 2.5 kb (Fig. 21). Two of the domains with 4 successive tiles which had a higher level of H3K4me2 in C24 compared with Col-0, colocalized with the positions of genes, At4G04510 (a

protein kinase family protein) and of At4G08970 (a member of a Mutator-like transposase family), respectively. A domain with four successive tiles in which C24 displayed a lower level of H3K4me2 than Col-0 colocalized with the gene At4G08110 (a member of a CACTA-like transposase family (Ptta/En/Spm)). Further domains with three successive

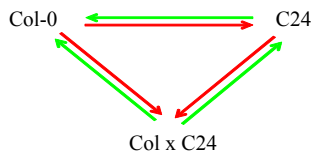
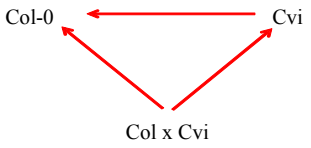
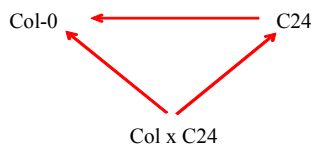
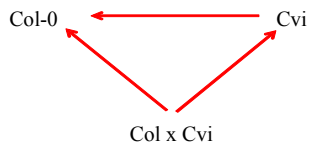
H3K4me2 polymorphic tiles colocalized with the position of other genes and members of retrotransposon families. Genes which are localized in H3K4me2 polymorphic domains are summarized in Excel File 4. These results suggest that in addition to sequence polymorphisms, chromatin modification polymorphisms can naturally occur between different accessions of *A. thaliana*.

The 548 H3K4me2 polymorphic tiles which indicated variation in the level of histone H3K4me2 between Col-0 and C24 were classified according to the sequences with which they are associated. In contrast to the sequence polymorphism detected in CGH experiments, histone H3K4me2 polymorphisms between Col-0 and C24 did not reveal any bias towards a particular sequence class (Fig. 22).



Figure 19 Tiles with significant signal ratios in CGH and H3K4me2 ChIP on chip experiments. Tiles on chromosome 4 tiling arrays with differential signal ratios comparison of Col-0 and C24 by CGH (yellow circle) and by H3K4me2 ChIP on chip (green circle). The ones common to both groups were excluded from the further interpretation of ChIP on chip data.

Table 19 Design of ChIP on chip experiments

histone modification	accessions involved	type of chip array	design of experiment
H3K4me2	Col-0 and C24	chromosome 4 tiling microarray	 <p>dye-swap design* two biological replicates</p>
	Col-0 and Cvi	the <i>Arabidopsis</i> whole genome NimbleGen array	 <p>dye-switch design** two biological replicates</p>
H3K27me3	Col-0 and C24	the <i>Arabidopsis</i> whole genome NimbleGen array	 <p>dye-switch design two biological replicates</p>
	Col-0 and Cvi	the <i>Arabidopsis</i> whole genome NimbleGen array	 <p>dye-switch design two biological replicates</p>

*: each arrow represents a hybridization, with its start and head denoting Cy5- and Cy3-labeling, respectively. Two arrows in opposite orientation indicate a dye-swap experiment.

** : in the dye-switch design, labeling of IPs with the same antibody was inverted between the two biological replicates.

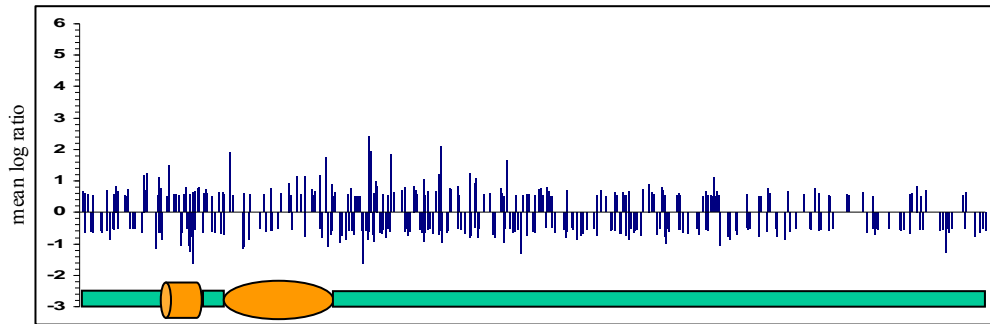


Figure 20 H3K4me2 polymorphisms between Col-0 and C24 along *A. thaliana* chromosome 4. The horizontal axis indicates the positions of tiles on chromosome 4 (model) of *A. thaliana*. Positive values indicate lower levels of H3K4me2 in C24 compared with Col-0, while negative values indicate higher levels of H3K4me2 in C24 compared with Col-0.

Table 20 Domains of contiguous H3K4me2 polymorphic tiles between C24 and Col-0.

number of successive tiles per domain	2	3	4
number of domains	62	15	3

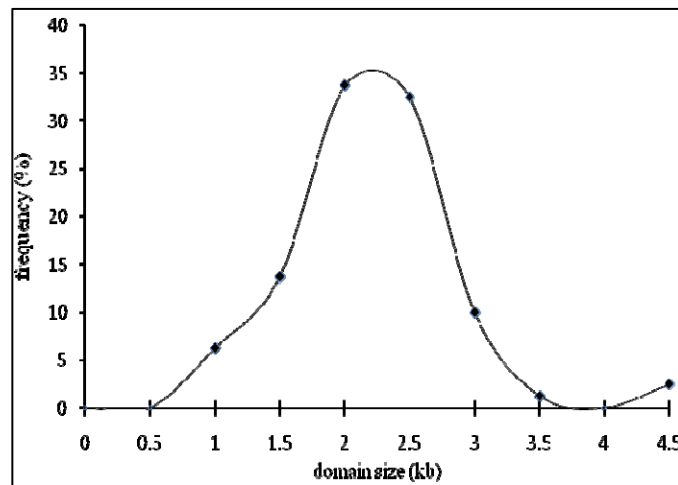


Figure 21 Size distribution of domains of contiguous tiles polymorphic for H3K4me2 levels in Col-0 compared with C24.

The sizes of domains were calculated by subtracting start and end positions of the first and last contiguous tile, respectively, on the chromosome sequence.

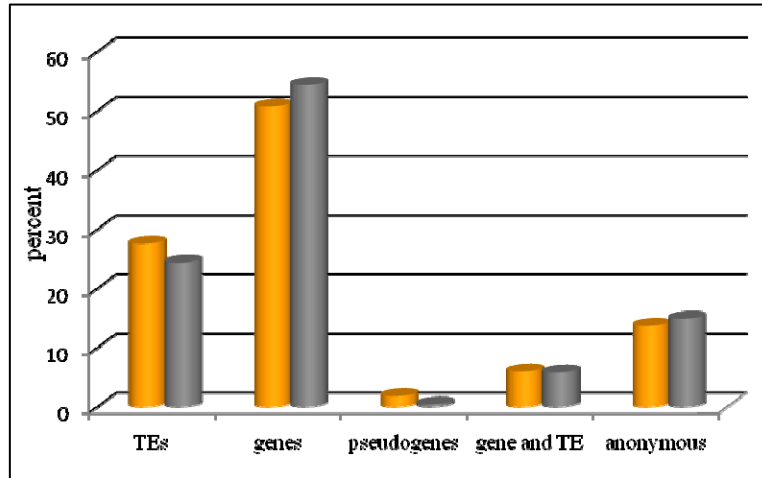


Figure 22 Sequence type classification of H3K4me2 polymorphic tiles in C24 compared with Col-0 along *A. thaliana* chromosome 4.

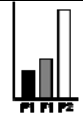
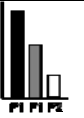
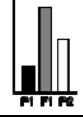
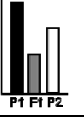
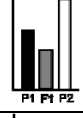
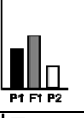
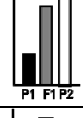
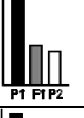
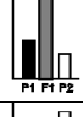

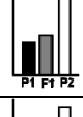
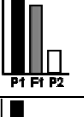
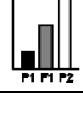
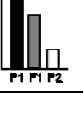
Grey bars indicate the percentages of the respective sequence classes on chromosome 4 tiling microarrays: 55% genes, 24% transposable elements, and 15% anonymous. Orange bars represent percentages among H3K4me2 polymorphic tiles: 51% genes, and 27% transposable elements.

4.8.2. Histone H3K4me2 is a stable modification which is additively inherited in intraspecific hybrids between *A. thaliana* accessions Col-0 and C24

To investigate whether new patterns of histone H3K4me2 would arise after intraspecific hybridization between Col-0 and C24, an intersection analysis of tiles showing significant H3K4me2 differences in Col-0 versus C24, Col-0 versus Col-0xC24 and Col-0xC24 versus C24 (after exclusion of CGH polymorphic tiles) was performed (Fig. 23). The analysis revealed that at only very few tiles met the criteria for H3K4me2 hybridization responsive tiles, that is, the level of histone H3K4me2 in Col-0xC24 hybrids rose above or fell below the levels found in both inbred parental lines, Col-0 and C24. Mere 12 tiles (0.06% of total tiles) in the hybrid progeny showed histone H3K4me2 modification levels higher (8 tiles) or lower (4 tiles) than both parental values (Fig. 23E). These tiles were all singletons and randomly distributed along chromosome 4 (Appendix Table 3). The vast majority of tiles displayed H3K4me2 close to mid-parent values.

In order to confirm the data from ChIP on chip analysis, quantitative PCR experiments were performed for a subset of the 12 H3K4me2 hybridization responsive tiles and their immediate

flanking tiles on anti-H3K4me2 immuno-precipitated DNA. Because of the inherent variability of immuno-precipitation assays, only differences in H3K4me2 levels between the tested genotypes of at least two fold were considered to be meaningful. The results did not indicate any difference in the H3K4me2 levels between both parents (Fig. 24). This was not unexpected, as tiles with a significant H3K4me2 difference between Col-0 and C24 in ChIP on chip analysis were not included in this section (Fig. 23 part E). However, for all candidate tiles reassessed by ChIP-quantitative PCR, the expected differences in the levels of histone H3K4me2 in hybrid offsprings (Col-0xC24, C24xCol-0) compared with the inbred parents Col-0 and C24 could not be detected (Fig. 24). Therefore, it has to be concluded that the candidate H3K4me2 responsive tiles identified by ChIP on chip represent experimental noise and are to be considered as false positives. In conclusion, our results identify histone H3K4me2 as a stable histone modification, which seems to hardly change in response to intraspecific hybridization of Col-0 and C24.

category	higher (+) than Col-0	lower (-) than Col-0	description
A			Col-0 (P1) and C24 (P2) are different
B			Col-0 and Col-0xC24 (F1) are different
C			C24 and Col-0xC24 are different
D			Col-0 and C24 are different; Col-0 and Col-0xC24 are different
E			Col-0 and Col-0xC24 are different; C24 and Col-0xC24 are different
F			Col-0 and C24 are different; C24 and Col-0xC24 are different
G			Col-0 and C24 are different; Col-0 and Col-0xC24 are different; C24 and Col-0xC24 are different

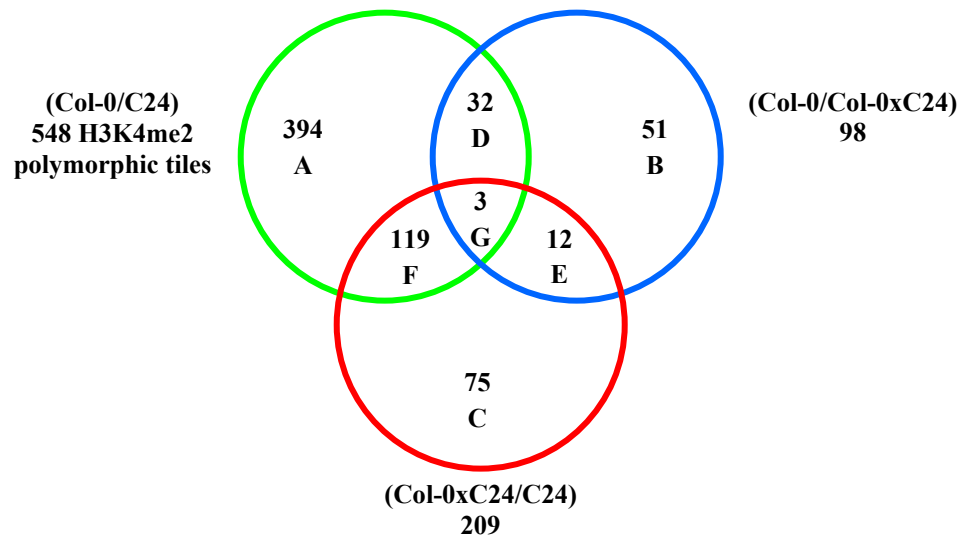


Figure 23 Intersection analysis of tiles showing differences in H3K4me2 in direct comparison between Col-0 / C24, Col-0 / Col-0xC24, and Col-0xC24 / C24.

The green circle includes tiles with significant differences in H3K4me2 ChIP on chip experiments for the combination Col-0 / C24. Blue and red circles include the respective tiles for the combinations Col-0 / Col-0xC24 and Col-0xC24 / C24. Section A defines tiles covering sequences with different levels of histone H3K4me2 between Col-0 and C24. The level of histone H3K4me2 in hybrid progeny does not differ from any of both parents. Section B defines tiles covering sequences for which histone H3K4me2 levels in both parents are not different, but the level in hybrid progeny is different from Col-0, but not from C24. Section C defines tiles covering sequences for which the histone H3K4me2 levels in both parents, are not different, but the level in hybrid progeny is different from C24, but not from Col-0. Section D indicates tiles covering sequences for which the histone H3K4me2 levels differ in both parents but the level in hybrid progeny is different from Col-0. This suggests that the histone H3K4me2 level for these regions in hybrid progeny is more similar to the histone H3K4me2 level of equivalent sequences in C24. Section E defines tiles covering sequences for which the histone H3K4me2 levels are similar in both parents, but differ in hybrid progeny from the parents. Eight of these tiles show higher histone H3K4me2 levels in the hybrid, and four show higher histone H3K4me2 levels in parents. Part F defines tiles covering sequences with different histone H3K4me2 levels in both parents and a level in the hybrid that is similar to Col-0 but different from C24. Part G represents tiles with unrelated histone H3K4me2 levels in each of both parents and the hybrid progeny.

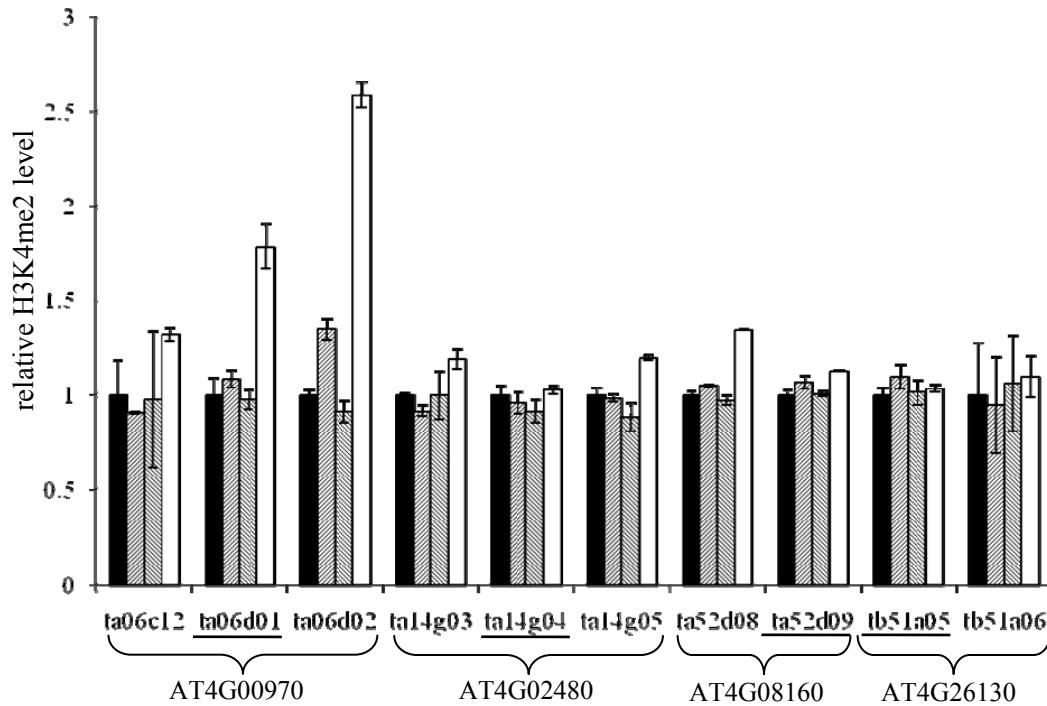


Figure 24 ChIP-quantitative PCR results for candidate regions showing responsiveness of H3K4me2 in intraspecific hybrids in Chip on chip analysis.

Candidate tiles (underlined in legend of the horizontal axis) and their immediate flanking tiles were selected based on the results from ChIP on chip experiments using chromosome 4 tiling microarray. Primer pairs were designed to have similar efficiency in all tested genotypes. H3K4me2-enriched immuno-precipitated DNA was prepared from inbred parental lines and hybrid offspring in two biological replicates and subjected to quantitative PCR in two technical replicates. The measured levels of H3K4me2 for the tiles of interest were normalized to the average levels of H3K4me2 in three control regions (Table 11). To facilitate the comparison between different tiles, data were recalibrated by assigning the value 1.0 to the median of the relative H3K4me2 level of Col-0. Error bars indicate absolute deviation between biological replicates. The differences in the levels of H3K4me2 between parents and their offspring determined by ChIP-quantitative PCR were in all cases less than two fold. Therefore, changes in H3K4me2 in response to the intraspecific hybridization seen in ChIP on chip analysis could not be confirmed.

4.8.3. Natural differences in the distribution of histone H3K4me2 between *A. thaliana* accessions Col-0 and Cvi

Out of 717,235 tiles present on *Arabidopsis* whole genome NimbleGen microarrays, ChIP on chip analysis revealed 20,441 tiles (2.85% of total tiles) in the first biological replicate (1BR) and 5,166 tiles (0.72% of total tiles) in second biological replicate (2BR) as H3K4me2 polymorphic between Col-0 and Cvi after exclusion of the CGH polymorphic tiles. Of the total

of tiles, 1.71% in 1BR and 0.38% in 2BR covered sequences with a higher level of H3K4me2 in Cvi compared with Col-0. Of these tiles 45.11% in 1BR and 18.65% in 2BR were assigned to singletons; the remaining tiles could be categorized in domains of contiguous tiles with tile numbers ranging from two to 114 for 1BR and two to 67 for 2BR.

The average size of contiguous regions with a higher level of histone H3K4me2 in Cvi compared with Col-0 was 3.2 kb in 1BR and 2.5 kb in 2BR. Of the regions with a higher level of H3K4me2 in Cvi 90% were less than 6 kb long in 1BR and 4.5 kb long in 2BR. However, larger domains with e.g. 33.5 kb in 1BR or 56 kb in 2BR could also be determined (Fig. 25). Genes which are localized in domains with a higher level of histone H3K4me2 in Cvi are summarized in Excel File 5 and Excel File 6.

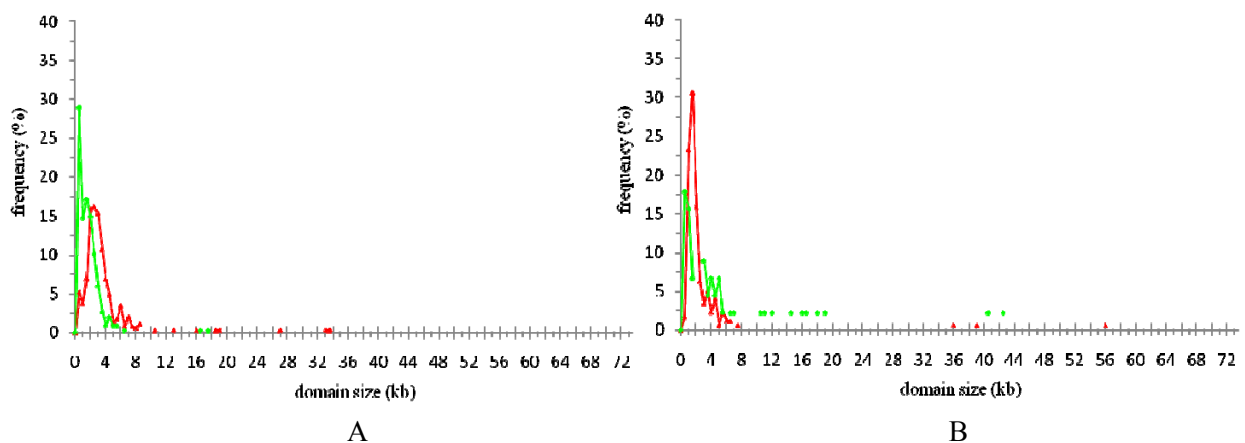


Figure 25 Sizes of genomic regions with higher (in red) or lower level (in green) of histone H3K4me2 in Cvi versus Col-0.

The sizes of domains was calculated by subtracting start and end positions of the first and last contiguous H3K4me2 polymorphic tile, respectively, on the related chromosome sequence. A: 1BR, B: 2BR.

Of the total tiles, 8,159 (1.14%) in 1BR and 2,453 (0.34%) in 2BR covered sequences with a lower level of H3K4me2 in Cvi compared with Col-0. Classifying of them into domains with contiguous tiles resulted in 4,145 singletons in 1BR and 768 singletons in 2BR. The remaining tiles could be assigned to domains with two to 115 successive tiles in 1BR and two to 258 successive tiles in 2BR. The average size of regions with lower levels of H3K4me2 in Cvi was 1.4 kb in 1BR and 6.2 kb in 2BR. Of the domains 90% were less than 3 kb long in 1BR and less

than 16.5 kb long in 2BR. Larger domains with sizes up to 17.5 kb in 1BR and 42.5 kb in 2BR could also be detected (Fig. 27). Genes which are localized in domains with a lower level of H3K4me2 in Cvi are summarised in Excel File 5 and Excel File 6.

To address the question whether histone H3K4me2 level polymorphisms between Col-0 and Cvi were randomly distributed, H3K4me2 polymorphic tiles were classified according to their sequence annotation. Tiles selected randomly as described in part 4.7.1.2 served as a control. The results revealed that histone H3K4me2 level differences between Col-0 and Cvi mainly affect transposable elements (Fig. 26).

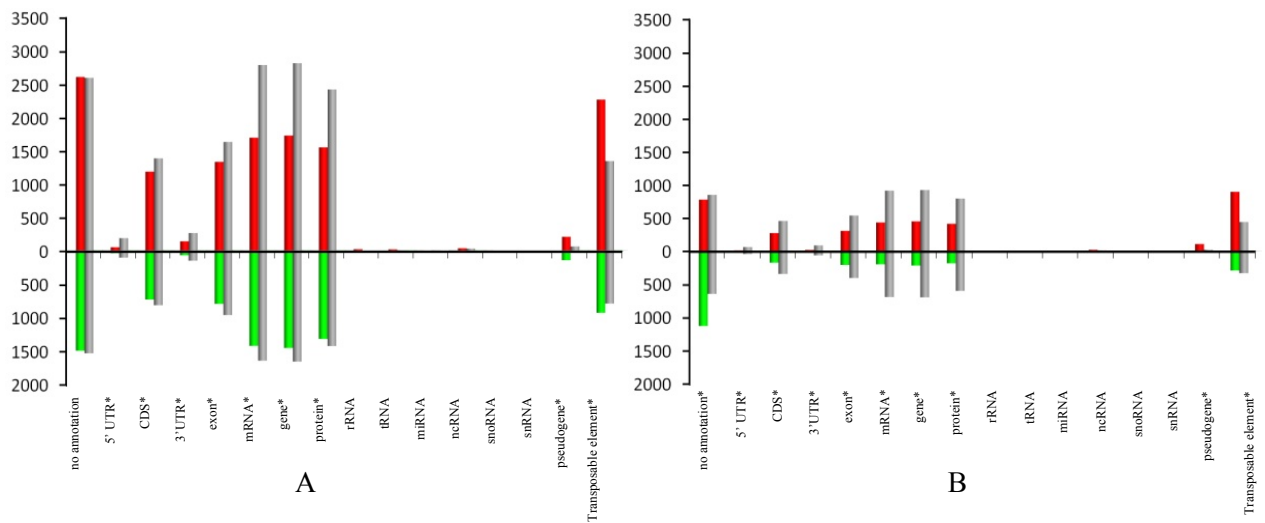


Figure 26 Class distribution of sequences covered by tiles with significant H3K4me2 differences between Col-0 and Cvi.

Red and green bars represent the numbers of tiles covering sequences with higher and lower level of H3K4me2 in Cvi relative to Col-0, respectively, assigned to particular sequence classes. Grey bars indicate controls generated by random selection of tiles. Asterisks identify classes for which frequencies in experimental results (green or red bars) differed from random controls (grey bars) significantly with a p-value of 0.05. A: 1BR (Cvi-Cy5_vs_Col-Cy3). B: 2BR (Cvi-Cy3_vs_Col-Cy5).

4.8.4. Histone H3K4me2 shows stability in response to intraspecific hybridization between *A. thaliana* accessions Col-0 and Cvi

To investigate whether a new pattern of histone H3K4me2 is formed in hybrid offspring after intraspecific hybridization between Col-0 and Cvi, an intersection analysis of the tiles showing

significant differences of H3K4me2 levels of the covered sequences in the comparisons Cvi versus Col-0, Col-0xCvi versus Col-0 and Col-0xCvi versus Cvi was performed (Fig. 27). The analysis revealed that in intraspecific hybrids of *A. thaliana* accessions Col-0 and Cvi the level of histone H3K4me2 of almost all genomic regions did not change significantly in response to intraspecific hybridization. Only 442 tiles in first biological replicate (0.06% of total tiles) showed in the hybrid progeny histone H3K4me2 levels higher than in both inbred parental lines. Also only 83 and 69 tiles (0.01% of total tiles) in the first and second biological replicate, respectively, showed in the hybrid progeny histone H3K4me2 levels lower than both parental values (Fig. 27 part E). However, there was no overlap between the H3K4me2 hybridization responsive tiles between the two biological replicates. As in the case of the hybridization of Col-0 with C24, we would rather interpret the found responsive tiles as experimental noise.

category	higher (+) than Col-0	lower (-) than Col-0	description	number of H3K4me2 polymorphic tiles
A			Col-0 (P1) and Cvi (P2) are different	+1982,-5297, 1BR; +548,-739, 2BR; +427, - 373 , combine
B			Col-0 and Col-0xCvi (F1) are different	+28316 ,-9748, 1BR; +304 ,-9284 ,2BR; +147 ,- 89 , combine
C			Cvi and Col-0xCvi are different	+6584, -4702,1BR; +723, -2354, 2BR; +178 ,- 551, combine
D			Col-0 and Cvi are different; Col-0 and Col-0xCvi are different	+9986, -2743, 1BR; +2134, -1714, 2BR; +1618 ,- 119, combine
E			Col-0 and Col-0xCvi are different; Cvi and Col-0xCvi are different	+442, -83, 1BR; 0, -69, 2BR; 0, 0, combine
F			Col-0 and Cvi are different; Cvi and Col-0xCvi are different	0,0, 1BR; 0, 0, 2BR; 0, 0, combine
G			Col-0 and Cvi are different; Col-0 and Col-0xCvi are different; Cvi and Col-0xCvi are different	+61, 0, 1BR; 0, 0, 2BR; 0, 0, combine

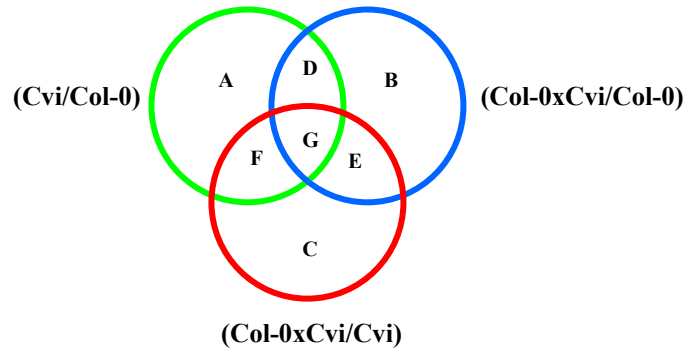


Figure 27 Intersection analysis of tiles showing differences in H3K4me2 in direct comparison between Cvi / Col-0, Col-0xCvi / Col-0, and Col-0xCvi / Cvi.

The green circle containing sections A, D, F and G represents tiles showing differences in H3K4me2 ChIP on chip experiments for Cvi versus Col-0. Blue (sections B, D, E, G) and red circles (sections C, E, F, G) represent tiles showing differences in H3K4me2 ChIP on chip experiments for Col-0xCvi versus Col-0 and Col-0xCvi versus Cvi, respectively.

Section A defines tiles covering sequences with different levels of histone H3K4me2 between Cvi and Col-0. The level of histone H3K4me2 in hybrid progeny does not differ from either of the parents. Section B defines tiles covering sequences for which histone H3K4me2 levels in both parents are not different but the level in hybrid progeny is different from Col-0, but not from Cvi. Section C defines tiles covering sequences for which the histone H3K4me2 levels in both parents are not different, but the level in hybrid progeny is different from Cvi, but not from Col-0. Section D indicates tiles covering sequences for which the histone H3K4me2 levels differ in both parents but, the level in hybrid progeny is different from Col-0. This suggests that the histone H3K4me2 level for these regions in hybrid progeny is more similar to the histone H3K4me2 level of equivalent sequences in Cvi. Section E defines tiles covering sequences for which the histone H3K4me2 levels are similar in both parents, but differ in hybrid progeny from the parents. Of these tiles 442 in 1BR, but none of them in 2BR show histone H3K4me2 levels higher in the hybrid, and 83 in 1BR and 69 in 2BR showed histone H3K4me2 level in parents higher than in the progeny. Part F defines tiles covering sequences with different histone H3K4me2 levels in both parents and a level in the hybrid that is similar to Col-0, but different from Cvi. Part G represents tiles with unrelated histone H3K4me2 levels in each of both parents and the hybrid progeny.

4.8.5. Natural differences in the distribution of histone H3K27me3 between *A. thaliana* accessions Col-0 and C24

Out of 717,235 tiles covering the whole genome, 74,509 tiles (10.39% of total tiles) in 1BR and 112,323 tiles (15.66% of total tiles) in 2BR differed in histone H3K27me3 levels between Col-0 and C24 according to ChIP on chip experiments after exclusion of CGH polymorphic tiles. In the 1BR, 55975 of these (7.8% of total tiles) and in the 2BR 74484 of these (10.38% of total tiles) covered sequences with a higher level of H3K27me3 in C24. While 16,226 of these tiles

in 1BR and 16,673 of these tiles in 2BR were singletons, the remainder could be categorized in domains with numbers of contiguous tiles ranging from two to 449 for 1BR and two to 149 tiles for 2BR.

The average size of contiguous regions with a higher level of histone H3K27me3 in C24 was 3 kb in the 1BR and 3.4 kb in 2BR. 90% of the regions with a higher level of H3K27me3 in C24 were less than 5 kb long in 1BR and less than 5.5 kb long in 2BR. However, several larger domains of up to 72 kb in 1BR and 26 kb in 2BR) could be also determined (Fig. 28). Genes which are localized in domains with a higher level of histone H3K27me3 in C24 are summarized in Excel File 7 and Excel File 8.

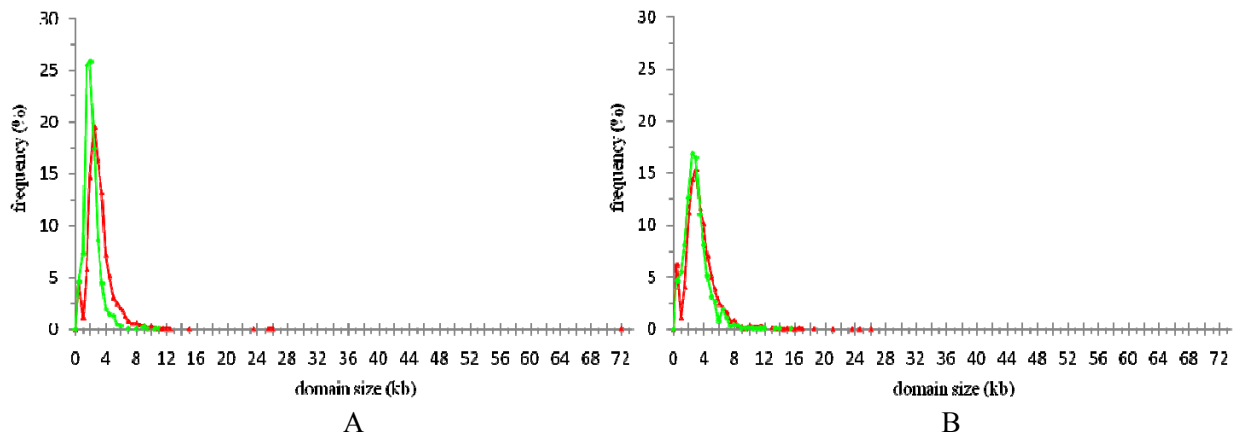


Figure 28 Sizes of regions with higher (in red) or lower levels (in green) of histone H3K27me3 in C24 versus Col-0.

Sizes of domains were calculated by subtracting start and end positions of the first and last contiguous tile on the respective chromosome sequence. A: 1BR. B: 2BR.

Of the total H3K27 polymorphic tiles, 18534 (2.58% of total tiles) in the 1BR and 37,839 (5.28% of total tiles) in 2BR covered sequences with a lower level of H3K27me3 in C24. Classification of these into domains of contiguous tiles identified 6661 singletons in 1BR and 10,051 in 2BR. Nevertheless, domains could be identified ranging from two to 112 successive tiles in 1BR and two to 141 successive tiles in 2BR. The average size of regions with lower levels of H3K27me3 in C24 was 1.9 kb in 1BR and 2.8 kb in 2BR. Of the contiguous domains 90% were less than 3 kb long in the 1BR and less than 4.5 kb long in 2BR. Larger domains of

up to 11 kb in 1BR and up to 15.5 kb in 2BR could also be detected (Fig. 28). Genes that are localized in domains with a lower level of H3K27me3 in C24 are summarized in Excel File 7 and Excel File 8.

To address the question whether differences in histone H3K27me3 levels between Col-0 and C24 are randomly distributed across sequences of different functions, H3K27me3 polymorphic tiles were classified based on the annotation of the sequences covered by them. A random selection of tiles served as a control (as described in part 4.7.1.2.). Results revealed that histone H3K27me3 polymorphisms level between Col-0 and C24 mainly affect genic regions. Also, transposable elements with higher H3K27me3 level in C24 compared to Col-0 were overrepresented (Fig. 29). The distribution of differences in histone H3K27me3 levels between Col-0 and C24 is almost similar in both biological replicates.

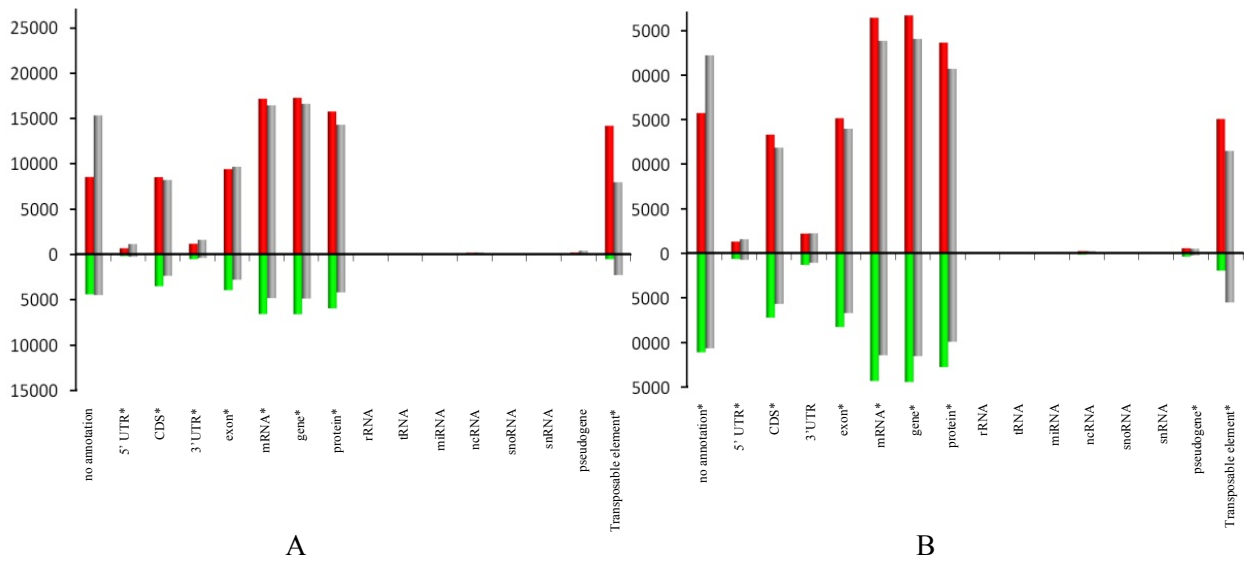
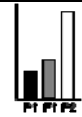
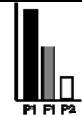

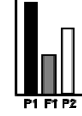
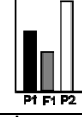
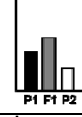
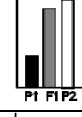
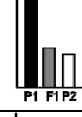
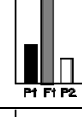
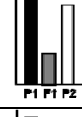
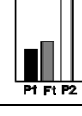
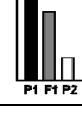
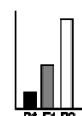
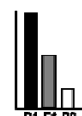


Figure 29 Classifications of tiles covering sequences with differences in H3K27me3 levels between Col-0 and C24.

Tiles were classified based on the annotation of their underlying sequences. Red and green bars represent number of tiles covering sequences with a higher or lower level of H3K27me3 in C24 relative to Col-0, respectively. Grey bars represent controls that were generated by random selection of tiles. Differences between experimental results (green or red bars) and random controls (grey bars) in classes marked with asterisks were significant with a p-value of 0.05. A: ChIP on chip experiments for 1BRs (C24-Cy5_vs_Col-Cy3). B: ChIP on chip experiments for 2BR (C24-Cy3_vs_Col-Cy5). Note that the scale of the graphs differs from one of the graphs for the classification of tiles with differences in H3K4me2 (Fig. 26).

4.8.6. Histone H3K27me3 shows dynamics in response to intraspecific hybridization between Col-0 and C24

To investigate whether a new pattern of histone H3K27me3 arose in hybrid offspring after intraspecific hybridization between Col-0 and C24, an intersection analysis of tiles covering sequences with differences in H3K27me3 levels in C24 versus Col-0, Col-0xC24 versus Col-0, and Col-0xC24 versus C24 was performed (Fig. 30).

category	(+) higher than Col-0	(-) lower than Col-0	description	number of H3K27me3 polymorphic tiles
A			Col-0 (P1) and C24 (P2) are different	+47448,-13394, 1BR; +57426,-27159, 2BR; +15294, - 5402 , combine
B			Col-0 and Col-0xC24 (F1) are different	+12242 ,-14224, 1BR; +12593 ,-13606 ,2BR; +2140 ,- 1983 , combine
C			C24 and Col-0xC24 are different	+21517, -48198,1BR; +22826, -26768, 2BR; +4947 ,- 6161, combine
D			Col-0 and C24 are different; Col-0 and Col-0xC24 are different	+7402, -4610, 1BR; +15435, -10151, 2BR; +512 ,- 153, combine
E			Col-0 and Col-0xC24 are different; C24 and Col-0xC24 are different	+5233, -8754, 1BR; +7220, -7374, 2BR; +195, -344, combine
F			Col-0 and C24 are different; C24 and Col-0xC24 are different	+6,-14, 1BR; +44, -14, 2BR; 0, -1, combine
G			Col-0 and C24 are different; Col-0 and Col-0xC24 are different; C24 and Col-0xC24 are different	+177, -268, 1BR; +559, -135, 2BR; +1, 0, combine

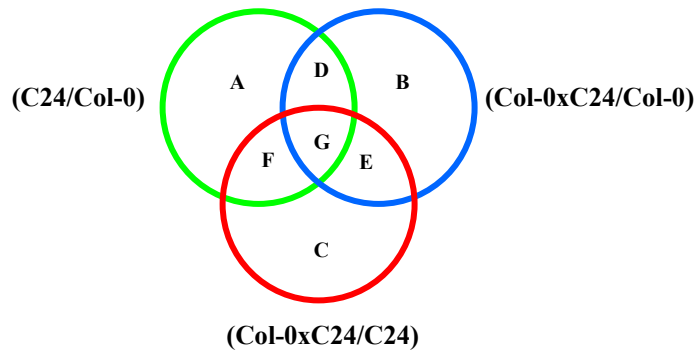


Figure 30 Intersection analysis of tiles showing differences in H3K27me3 in direct comparison between C24 / Col-0, Col-0xC24 / Col-0, and Col-0xC24 / C24.

The green circle containing sections A, D, F and G represents tiles showing differences in H3K27me3 ChIP on chip experiments for C24 versus Col-0 combination. Blue (sections B, D, E, G) and red circles (sections C, E, F, G) represent tiles showing differences in H3K27me3 ChIP on chip experiments for Col-0xC24 versus Col-0 and Col-0xC24 versus C24, respectively.

Section A defines tiles covering sequences with different levels of histone H3K27me3 between C24 and Col-0. The level of histone H3K27me3 in hybrid progeny does not differ from either of the parents. Section B defines tiles covering sequences for which histone H3K27me3 levels in both parents are not different, but the level in hybrid progeny is different from Col-0, but not from C24. Section C defines tiles covering sequences for which the histone H3K27me3 levels in both parents are not different, but the histone H3K27me3 level in hybrid progeny is different from C24, but not from Col-0. Section D indicates tiles covering sequences for which the histone H3K27me3 levels differ in both parents, but the level in hybrid progeny is different from Col-0. This suggests that the histone H3K27me3 level for these regions in hybrid progeny is more similar to the histone H3K27me3 level of equivalent sequences in C24. Section E defines tiles covering sequences for which the histone H3K27me3 levels are similar in both parents, but differ in hybrid progeny from the parents. Of these tiles 5233 in 1BR and 7220 in 2BR showed histone H3K27me3 levels higher in the hybrid than in the parents, and 8754 in 1BR and 7374 in 2BR showed histone H3K27me3 level in parents higher than in the hybrid. Part F defines tiles covering sequences with different histone H3K27me3 levels in both parents and a level in the hybrid that is similar to Col-0, but different from C24. Part G represents tiles with unrelated histone H3K27me3 levels in each of both parents and the hybrid progeny.

The analysis revealed that in Col-0xC24 hybrids the levels of histone H3K27me3 of the majority of genomic regions did not show any indications of hybridization-related changes. Rather, they displayed H3K27me levels indistinguishable from mid-parent values. Only 5,233 tiles in 1BR and 7,220 in 2BR (0.73% and 1% of total tiles in 1BR and 2BR, respectively) showed histone H3K27me3 modification levels higher in the hybrids than the parents. Similarly, only 8,754 tiles in 1BR and 7,220 in 2BR (1.22% and 1.03% of total tiles in 1BR and 2BR, respectively) showed histone H3K27me3 modification levels lower in hybrids than in

both parents (Fig. 30 part E). Between 1 BR and 2 BR H3K27me3 hybridization responsive tiles, an overlap of 195 and 344 tiles showing higher or lower H3K27me3 levels, respectively, in Col-0xC24 compared with parents was found. These tiles could be categorized into 72 domains with two to 11 contiguous tiles with a lower level of H3K27me3 and 29 domains with two to 16 contiguous tiles with a higher level of H3K27me3 in hybrids compared with the parents. The domains were dispersed along the genome and colocalized mainly with genes of diverse functions (Excel File 9). Ontology categorization of genes associated with H3K27me3 hybridization responsive regions in Col-0xC24 common between both biological replicates indicated that they preferentially involve genes with a function in developmental processes, transport, and DNA or RNA metabolism (Fig. 31).

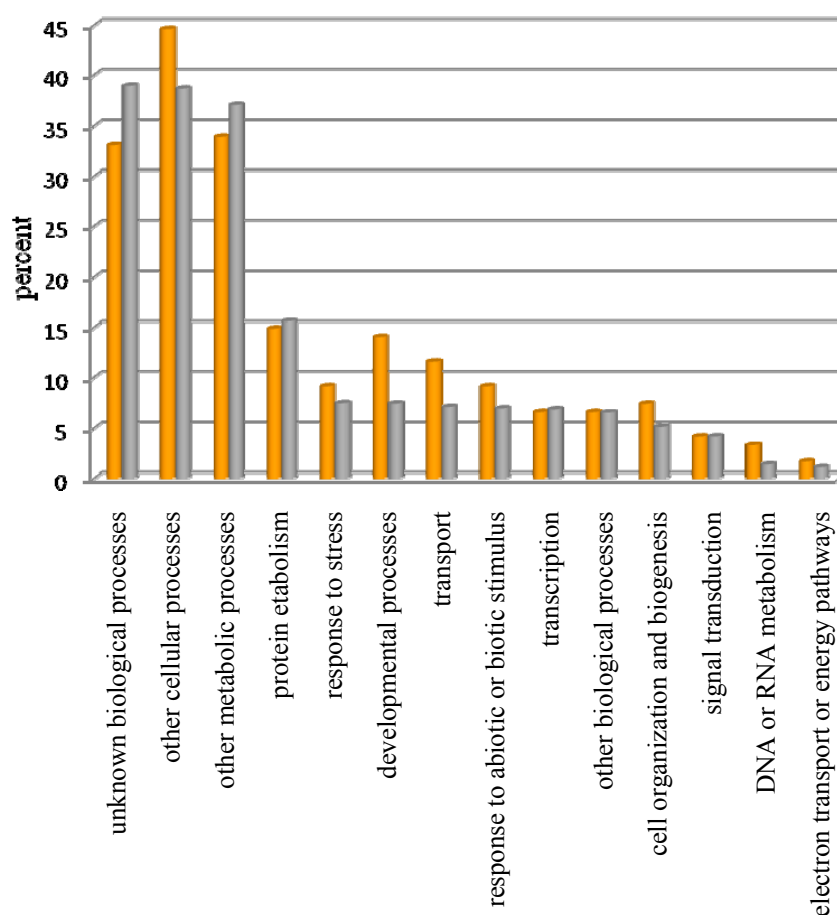


Figure 31 Ontology classification of genes localized in H3K27me3 hybridization responsive regions common to both biological replicates for Col-0 versus C24.

Orange bars indicate frequencies of ontology classes of genes in H3K27me3 hybridization responsive regions. Grey bars represent the random frequencies of genes of the same classification based on the whole genome categorization of *A. thaliana*.

4.8.7. Natural differences in the distribution of histone H3K27me3 between *A. thaliana* accessions Col-0 and Cvi

Out of 717,235 total tiles covering the whole genome, 72,268 tiles (10.08%) in 1BR and 30,009 tiles (4.18%) in 2BR were identified to show differences in histone H3K27me3 after exclusion of CGH polymorphic tiles in ChIP on chip experiments comparing *A. thaliana* accessions Col-0 and Cvi. Of the total tiles 6.08% in 1BR and 2.23% in 2BR covered sequences with a higher level of H3K27me3 in Cvi. Of these tiles, 12,598 in 1BR and 5,901 of in 2BR were assigned to singletons, while the remainder categorized into contiguous domains ranging from two to 217 tiles for 1BR and two to 67 tiles for 2BR.

The average size of regions with a higher level of histone H3K27me3 in Cvi was 2.7kb in 1BR and 2 kb in 2BR. Of regions with a higher level of H3K27me3 90%.in Cvi were less than 5 kb long in 1BR and less than 3.5 kb long in 2BR. However, larger domains of up to 35 kb in 1BR and 11 kb in 2BR could also be determined (Fig. 32). Genes which are localized in domains with a higher level of histone H3K27me3 in Cvi are summarized in Excel File 10 and Excel File 11.

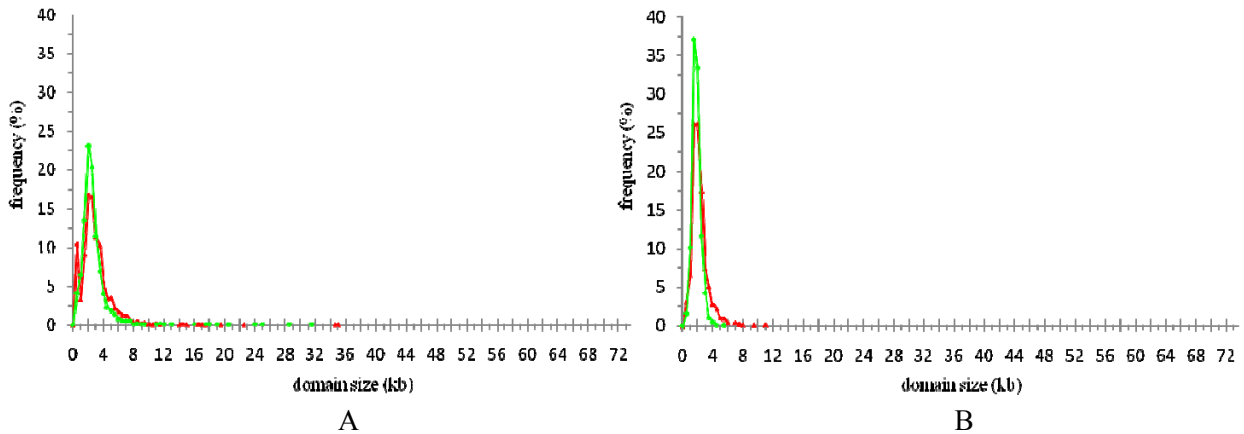


Figure 32 Sizes of regions with higher (in red) or lower levels (in green) of histone H3K27me3 in Cvi versus Col-0.

Sizes of domains were calculated by subtracting start and end positions of the first and last contiguous tile on the respective chromosome sequence. A: 1BR. B: 2BR.

Of the total tiles 3.99% in 1BR and 1.95% in 2BR covered sequences with a lower level of H3K27me3 in Cvi compared with Col-0. Classifying them into domains with contiguous tiles

resulted in 8,367 singletons in 1BR and 6,926 tiles in 2BR. In addition, domains with at least two to 172 tiles in 1BR and two to 33 in 2BR could be identified. The average size of regions with lower levels of H3K27me3 in Cvi was 2.4 kb in 1BR and 1.6 kb in 2BR. Of domains 90% were less than 4 kb long in 1BR and less than 2.5 kb long in 2BR. Larger domains of up to 31.5 kb in 1BR and up to 5.5 kb in 2BR could also be detected (Fig. 32). Genes which are localized in domains with a lower level of H3K27me3 in Cvi are summarised in Excel File 10 and Excel File 11.

To address the question whether histone H3K27me3 level polymorphisms were randomly distributed across sequences of different functions, sequences covered by tiles with different H3K27me2 levels between Col-0 and Cvi were classified according to their sequence annotations. Randomly selected tiles served as controls, as described in part 4.7.1.2. Results revealed that histone H3K27me3 level differences between Col-0 and Cvi mainly affect genic regions (Fig. 33).

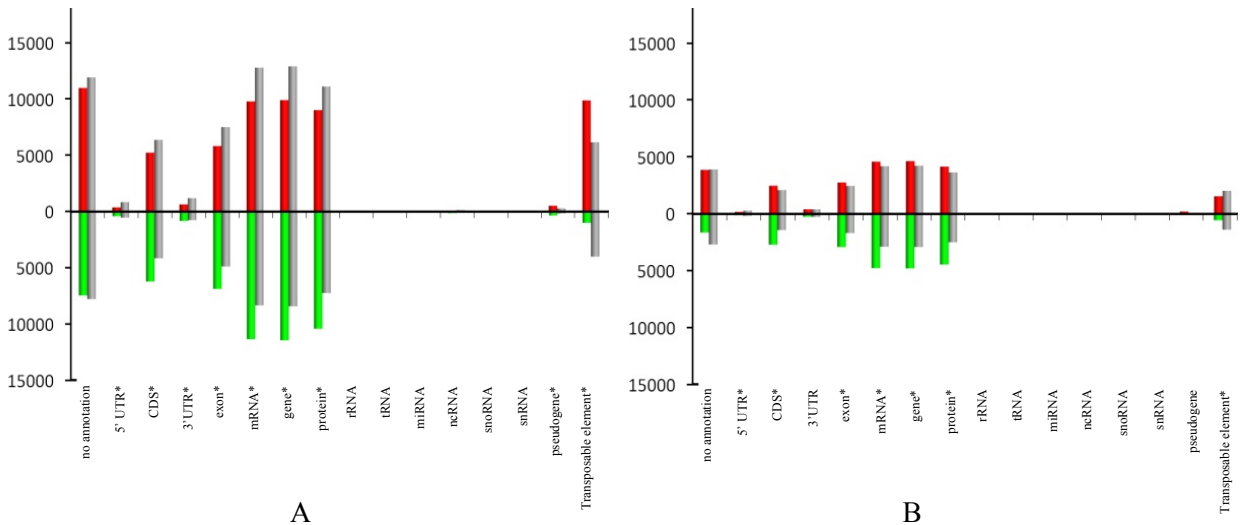


Figure 33 Classifications of tiles covering sequences with differences in H3K27me3 levels between Col-0 and Cvi.

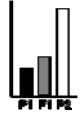
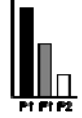






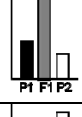
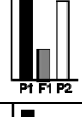
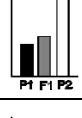
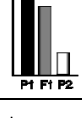


Tiles were classified based on the annotation of their underlying sequences. Red and green bars represent number of tiles covering sequences with a higher or lower level of H3K27me3 in Cvi relative to Col-0, respectively. Grey bars represent controls that were generated by random selection of tiles. Differences between experimental results (green or red bars) and random controls (grey bars) in classes marked with asterisks were significant with a p-value of 0.05. A: ChIP on chip experiments for 1BRs (Cvi-Cy5_vs_Col-Cy3). B: ChIP on chip experiments for 2BR (Cvi-Cy3_vs_Col-Cy5). Note that the scales of graphs differ from ones of the graphs for the classification of tiles showing differences for H3K4me2 (Fig. 26).

4.8.8. Histone H3K27me3 shows dynamics in response to intraspecific hybridization between Col-0 and Cvi

To investigate whether a new pattern of histone H3K27me3 arose in hybrid offspring after intraspecific hybridization between Col-0 and Cvi, an intersection analysis of tiles differing in H3K27me3 levels among Cvi versus Col-0, Col-0xCvi versus Col-0, and Col-0xCvi versus Cvi was performed (Fig. 34).

The analysis revealed that in Col-0xCvi hybrids the level of histone H3K27me3 of the majority of genomic regions hardly showed a change in response to intraspecific hybridization. The vast majority of tiles displayed H3K27me3 levels close to mid-parent values. Only 9,241 tiles in 1BR and 4,006 tiles in 2BR (1.29% and 0.55% of total tiles in 1BR and 2BR, respectively) in the hybrid progeny showed histone H3K27me3 levels higher than both parents. 2,877 tiles in 1BR and 5,028 in 2BR (0.4% and 0.7% of total tiles in 1BR and 2BR, respectively) in the hybrid progeny showed histone H3K27me3 modification levels lower than both parents (Fig. 34 part E). H3K27me3 hybridization responsive tiles overlapped between biological replicates 1BR and 2BR in 171 and 105 tiles showing higher or lower H3K27me3 levels in Col-0xCvi compared with the parents, respectively. These tiles could be categorized into 18 domains with two to 10 contiguous hybridization responsive tiles with a lower level of H3K27me3 and 32 domains with two to 10 contiguous hybridization responsive tiles with a higher level of H3K27me3 in hybrids compared with the parents. The domains were dispersed along the genome and colocalized mainly with genes of diverse functions (Excel File 12).

Ontology categorization of these genes associated with H3K27me3 hybridization responsive regions in Col-0xCvi common between both biological replicates indicates that they preferentially involve genes with a function in protein metabolism, developmental processes, transport, and DNA or RNA metabolism and show less representation of genes involved in transcription (Fig. 35).

category	higher (+) than Col-0	lower (-) than Col-0	description	number of H3K27me3 polymorphic tiles
A			Col-0 (P1) and Cvi (P2) are different	+34605,-24979, 1BR; +11665,-9871, 2BR; +3930, - 1691 , combine
B			Col-0 and Col-0xCvi (F1) are different	+17796 ,-14674, 1BR; +6562 ,-15983 ,2BR; +1439 ,- 1903 , combine
C			Cvi and Col-0xCvi are different	+26504, -25071,1BR; +15997, -13428, 2BR; +2246 ,- 1797, combine
D			Col-0 and Cvi are different; Col-0 and Col-0xCvi are different	+8203, -3025, 1BR; +4193, -3915, 2BR; +2159 ,- 134, combine
E			Col-0 and Col-0xCvi are different; Cvi and Col-0xCvi are different	+9241, -2877, 1BR; +4006, -5028, 2BR; +171, -105, combine
F			Col-0 and Cvi are different; Cvi and Col-0xCvi are different	+66, -143, 1BR; +7, -5, 2BR; 0, 0, combine
G			Col-0 and Cvi are different; Col-0 and Col-0xCvi are different; Cvi and Col-0xCvi are different	+66, +136, 1BR; +24, +38, 2BR; 0, 0, combine

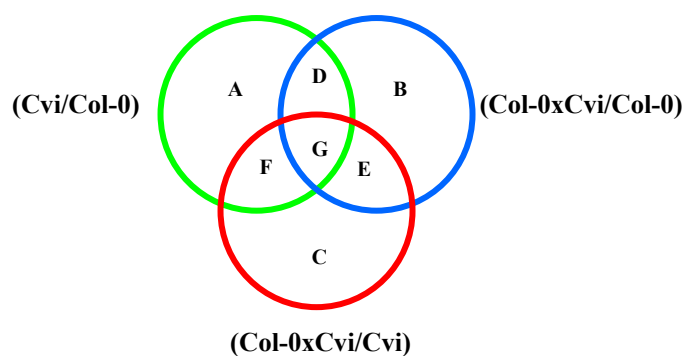


Figure 34 Intersection analysis of tiles showing differences in H3K27me3 in direct comparison between Cvi / Col-0, Col-0xCvi / Col-0, and Col-0xCvi / Cvi.

The green circle containing sections A, D, F and G represents tiles showing differences in H3K27me3 ChIP on chip experiments for Cvi versus Col-0 combination. Blue (sections B, D,

E, G) and red circles (sections C, E, F, G) represent tiles showing differences in H3K27me3 ChIP on chip experiments for Col-0xCvi versus Col-0 and Col-0xCvi versus Cvi, respectively. Section A defines tiles covering sequences with different levels of histone H3K27me3 between Cvi and Col-0. The level of histone H3K27me3 in hybrid progeny does not differ from any of both parents. Section B defines tiles covering sequences for which histone H3K27me3 levels in both parents are not different, but the level in hybrid progeny is different from Col-0, but not from Cvi. Section C defines tiles covering sequences for which the histone H3K27me3 levels in both parents are not different, but the histone H3K27me3 level in hybrid progeny is different from Cvi, but not from Col-0. Section D indicates tiles covering sequences for which the histone H3K27me3 levels differ in both parents, but the level in hybrid progeny is different from Col-0. This suggests that the histone H3K27me3 level for these regions in hybrid progeny is more similar to the histone H3K27me3 level of equivalent sequences in Cvi. Section E defines tiles covering sequences for which the histone H3K27me3 levels are similar in both parents, but differ in hybrid progeny from the parents. Of these tiles 5233 in 1BR and 7220 in 2BR showed histone H3K27me3 levels higher in the hybrid than in the parents, and 8754 in 1BR and 7374 in 2BR showed histone H3K27me3 level in parents higher than in the hybrid. Part F defines tiles covering sequences with different histone H3K27me3 levels in both parents and a level in the hybrid that is similar to Col-0, but different from Cvi. Part G represents tiles with unrelated histone H3K27me3 levels in each of both parents and the hybrid progeny.

4.8.9. H3K27me3 shows higher dynamics in response to intraspecific hybridization than H3K4me2

In contrast to H3K4me2 hybridization responsive tiles of hybrid Col-0xCvi, where no overlap was found between biological replicates, domains of contiguous hybridization responsive tiles that were common between biological replicates were found for histone H3K27me3. Almost all of these common H3K27me3 hybridization responsive domains were covering genes. An equivalent direct comparison for hybrid Col-0xC24 was not possible because the analysis of H3K4me2 and H3K7me3 was performed with different microarray systems. Nevertheless, CGH results for the comparison of Col-0 and C24 from both microarray types were rather comparable (Fig. 17). ChIP-quantitative PCR analysis confirmed the stability of H3K4me2 (Fig 24) found with chromosome 4 tiling microarrays for hybrid Col-0xC24, while *Arabidopsis* whole genome NibleGen array analysis detected the same order of magnitude of H3K27me3 responsive tiles in hybrid Col-0xC24 as in Col-0xCvi. In summary, it seems feasible to conclude that also in crosses between Col-0 and C24, H3K27me3 is more dynamic than H3K4me2. Therefore, different histone modifications seem to show different responsiveness to intraspecific hybridization.

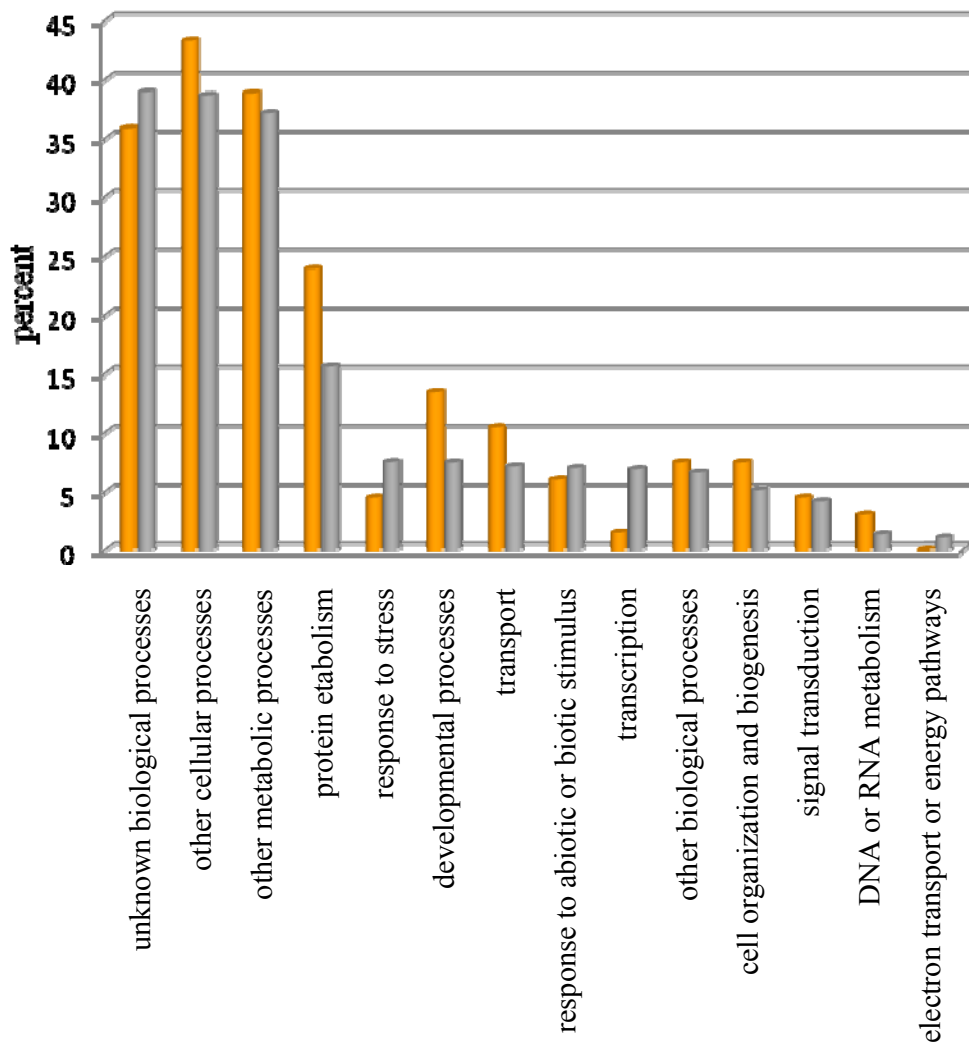


Figure 35 Ontology classification of genes localized in H3K27me3 hybridization responsive regions common to both biological replicates for Col-0 versus Cvi.

Orange bars indicate frequencies of ontology classes of genes in H3K27me3 hybridization responsive regions. Grey bars represent the random frequencies of genes in the same classification based on the whole genome categorization of *A. thaliana*.

Although we found a high level of variability between biological replicates in ChIP on chip experiments, but the magnitude of this variability was not the same for H3K27me3 polymorphic tiles and H3K27me3 hybridization responsive tiles in intersection analysis (Fig. 30 and 34 categories A and E, respectively). The overlap of H3K27me3 polymorphic tiles between biological replicates was around 16% for Col-0 and C24 and 7.5 % for Col-0 and Cvi (Fig. 30 and 34, category A). However, the overlap between H3K27me3 hybridization

responsive tiles between biological replicates was approximately 1.9 % for C24 and Col-0 and 1.3 % for Cvi and Col-0 (Fig. 30 and 34, category E). This suggests that in addition to the possible variations caused by technical issues, the observed variation between biological replicates is at least partly caused by highly dynamic behaviour of H3K27me3 in response to hybridization. This is in contrast to H3K4me2 where overlapping of H3K4me2 polymorphic tiles between biological replicates was 10.3 %, but no dynamic behaviour beyond technical noise was observed (Fig. 27).

5. Discussion

5.1. Intraspecific hybridization in *A. thaliana* does not change the endopolyploidy level

In *A. thaliana* somatic cells, endopolyploidy levels vary from 4C to 32C (Sugimoto-Shirasu and Roberts, 2003). This range of endopolyploidization was also observed in inbred accessions and their hybrid offspring (Appendix Fig.4). Comparisons between endopolyploidy levels at different time points showed increasing proportions of nuclei with higher ploidy levels. Notably, cotyledons of 10 DAS seedlings had a higher ploidy level than secondary leaves of 15 DAS plantlets (Fig. 4). This is consistent with reports that in *A. thaliana* the level of endopolyploidization is regulated during development (Galbraith et al. 1991) and rather increases with the age of the tissue analyzed. In general our data did not show any significant difference in endopolyploidy levels between inbred parental lines and offspring from reciprocal hybrids offspring. Therefore, an increased vigour and accelerated growth rate as has been reported for hybrid offsprings of crosses between different accessions of *A. thaliana* (Meyer et al. 2004) is not likely to be correlated with a higher level of endopolyploidization.

5.2. Nucleolus size is not altered by intraspecific hybridization

It is well documented that the growth rate of organisms is correlated with the amount of cellular ribosomal RNA (rRNA). A high level of rRNA should support fast protein synthesis, which in turn should favour a higher growth rate of cells. Increased copy number of rRNA genes through several rounds of DNA replication without cell division or an increase in the transcription rate per rDNA copy may account for such increased cellular rRNA levels (Elser et al., 2000). Evidence suggests that changes in ribosomal RNA gene transcription levels are achieved primarily by changing the number of genes that are actively transcribed, rather than changing the number of transcripts per gene (Pikaard, 2002). It has been shown that the size of a nucleolus is proportional to the amount of rRNA synthesized (Caspersson, 1950) and the area of silver-stained nucleolar regions is closely correlated with both nucleolar size and RNA polymerase I transcriptional activity (Derenzini et al., 1992, 2000).

The size of nucleoli in 4C nuclei of *A. thaliana* was generally bigger than in 2C nuclei in each genotype tested. This indicates that the method used for size determination is reliable. However, the variation in nucleolus area measurements within each genotype in our analysis was high. One possible explanation for this could be that whole seedlings were used as material to isolate nuclei. Seedlings contain various cell types that might differ considerably in the activity of NORs. Also, there is evidence that the size and shape of nucleoli vary widely among different cell types, and even for the same cell type, as a function of the proliferation status of the cells (Jacquet et al., 2001).

Variation of rRNA gene copy number and DNA methylation levels between *A. thaliana* accessions have been reported by Riddle and Richards (2002). Therefore, the subtle difference in nucleolus area size might reflect such differences between Col-0 and C24 and might indeed indicate slightly different rRNA synthesis rates. Nucleolus area measurements did not detect a consistent increase in hybrids versus inbred parents. Hence, the nucleolus area does not indicate an increased rRNA gene activity in hybrid offspring, which could help to explain the increased growth vigour observed by Meyer et al. (2004). This conclusion is supported by our endopolyploidy and DNA methylation data. The level of endopolyploidization was not significantly different between hybrid offspring and inbred parents (Fig. 4). Hence, it is less likely that the copy number of rRNA genes has changed in hybrid offspring. On the other hand, data from MSAP analyzes did not show any decrease in DNA methylation in hybrid offspring (Fig. 7). Therefore, it is less probable that rRNA genes, which are silent in parents are reactivated in hybrid offspring.

5.3. Global distribution of chromatin modifications is not altered by intraspecific hybridization

Indirect immunostaining using antibodies against 5mC and three histone modification marks revealed that DNA methylation and H3K9me2 co-localized with heterochromatic chromocenters, while H3K4me2 and H3K27me3 were distributed across euchromatin (Fig. 6). Fuchs et al. (2006) previously reported similar patterns, in which DAPI-stained heterochromatic chromocenters co-localized with H3K9me2- and 5mC-specific signals, while signals for

H3K4me2 and H3K27me3 were dispersed in the euchromatin of interphase nuclei and were excluded from chromocenters.

No major differences in intensity and distribution of DNA and histone methylation marks could be detected by microscopic analysis after intraspecific hybridization (Fig. 6). Thus, no major reorganisation of DNA and histone methylation marks occurs after intraspecific hybridization. This observation agrees with our MSAP and ChIP on chip analysis data in which only limited changes in DNA methylation and histone H3K27 trimethylation were observed in hybrid offspring compared to inbred parents.

5.4. Intraspecific hybrid formation is accompanied by a slight increase of DNA methylation

Genome-wide DNA methylation analysis by MSAP detected a change of methylation in about 3% of the covered *MspI* / *HpaII* recognition sites in Col-0x C24 and C24xCol-0 hybrids compared with Col-0 and C24 inbred parental lines. These hybridization-related changes were always associated with a gain of methylation at the inner cytosine of the CCGG restriction site. A similar level of differentially methylated sites (approximately 1% gain or loss) was previously reported for intraspecific hybrids of two crop plants, rice (Xiong et al., 1999) and cotton (Zhao et al., 2008). In contrast, 6.8% and 8.3% of sequences of resynthesized *Brassica napus* (Xu et al., 2009) and *Arabidopsis* allopolyploids (Madlung et al., 2002), respectively, showed DNA methylation changes compared with their diploid progenitors. In independently formed hybrids *Spartina x townsendii* and *Spartina x neyrautii*, even 30% of the parental methylation sites were altered in the hybrids (Salmon et al., 2005). Hence it is likely, that intraspecific hybridization in general induces fewer DNA methylation changes than interspecific hybridization. However, in contrast to crop plants, there was no instance of a loss of cytosine methylation in *A. thaliana* hybrids compared with their parents. These observations suggest that an intraspecific hybridization-driven loss of cytosine methylation is species dependent. A similar species-dependence has also been reported for rapid genome reorganization events after allopolyploidization (Ma and Gustafson, 2005).

A second conclusion can be drawn from comparing DNA methylation patterns in seedlings at 6 DAS and 21 DAS. The number of scorable MSAP signals declined during this time period from 368 to 291. As both, *MspI* and *HpaII* are sensitive to DNA methylation, this indicates a progressive increase of DNA methylation over time and is consistent with the observations of Ruiz-Garcia et al. (2005) that overall DNA methylation in *A. thaliana* increases with progressing developmental stages. This effect was independent of the genotype. Both inbred parents and their intraspecific hybrid offspring displayed increased methylation throughout development.

In agreement with other reports (Xiong et al., 1999; Zhao et al., 2008) methylation of the internal cytosine in CCGG sites in the genome of inbred parental lines and their hybrid offspring was detected more frequently than methylation of the external cytosine.

DNA methylation as an important epigenetic mark has two important roles in plants. First, it acts as a genome defence by silencing repeats and transposable elements and second, it is involved in gene regulation (Saze et al., 2008). Barbara McClintock pointed out that unification of two parental genomes with different composition of repetitive elements will cause a “genomic shock” which may initiate mobilization of these elements (Comai, 2000). DNA methylation at CHG sites and dimethylation of histone H3 at position lysine 9 in *A. thaliana* are tightly correlated. Both modifications occur frequently at transcriptionally silenced heterochromatic loci, including transposons in the euchromatic arms of the chromosomes (Bernatavichute et al., 2008). Silencing of transcription is associated with a decreasing level of histone H3K4 dimethylation and an increase of DNA and H3K9 methylation (Fuks, 2005). The limited changes of DNA methylation patterns in hybrid offspring in response to intraspecific hybridization together with the observed relative stability in the levels of H3K4me2 and H3K27me3 may indicate that intraspecific hybridization is a mild “genomic shock” which probably does not trigger the reactivation of many transposable elements. This idea is supported by several studies in which the inheritance of gene expression patterns in intraspecific hybridization was found to be mainly additive (Guo et al., 2006; Swanson-Wagner et al., 2006). In contrast, genome-wide non-additive gene expression, which was not found in *Arabidopsis* isogenic autotetraploids was observed in *Arabidopsis* allotetraploids (Wang et al., 2006). In accordance with our results, profiling of DNA methylation patterns in *A. thaliana* accessions of

Col-0 and Vancouver (Van) and their reciprocal hybrids at the whole genome level revealed the existence of CCGG methylation polymorphisms among these accessions (Zhang et al., 2008). Our study showed substantial methylation polymorphism between analyzed accessions (Fig. 7 A3, A4, B4 and C1). Furthermore, we also showed that inheritance of methylation polymorphisms in reciprocal hybrids is predominantly additive.

5.5. DNA methylation and transcript levels of single genes were largely unaffected by intraspecific hybrid formation

No change of the methylation status was observed in the reciprocal hybrids compared with the inbred parental lines for all seven candidate genes that were analyzed. This observation supports the rather constant DNA methylation levels between hybrid offspring and parents using MSAP analysis. The same applies to transcript levels, which were always close to calculated mid-parent values in the intraspecific hybrids. The additivity of allelic transcript levels in hybrids suggests that observed differential expression is mainly caused by *cis*-regulatory divergence between parental lines (Guo et al., 2006). This idea is also supported by our CGH data in which we could find sequence polymorphisms among inbred parental lines. In this context, it is important to note that the relative transcript levels of parental inbred lines and their reciprocal hybrids determined by quantitative RT-PCR (Fig. 9) and semi-quantitative RT-PCR (Fig. 10) were largely consistent. This confirms the results of the allele-specific semi-quantitative transcript analysis.

Relative transcript levels determined by microarray analysis are often only partially confirmed by subsequent quantitative RT-PCR. Similar observations were made in several studies on heterosis-related expression changes in maize (Uzarowska et al., 2007, Meyer et al., 2007, Hoecker et al., 2008). No clear correlation was detected between DNA methylation in the 5' regulatory region (Fig. 8) and the allele-specific transcript levels in *A. thaliana* (Fig. 10). For example, for gene At4g29200, DNA methylation was detected in the Col-0- and the C24-allele at similar levels, but transcripts were detected only from the C24 and not the Col-0 allele. This observation confirms the finding that gene expression levels are not necessarily correlated with DNA methylation levels of 5' regulatory regions of the respective genes (Zhang et al., 2006).

5.6. Sequence polymorphisms exist among different accessions of *A. thaliana*

5.6.1. Comparative genomic hybridization is a powerful technique for detecting sequence polymorphisms

Microarray-based comparative genomic hybridization (array-CGH) is a method by which variations in copy numbers of particular sequences between two genomes can be readily detected and simultaneously localized (Mantripragada et al., 2004). Currently, array-CGH employing high-density oligonucleotide arrays is widely used to study sequence polymorphisms in tumor cells (Jonsson et al., 2005), as well as in other fields of medical diagnosis (Vissers et al., 2005). In plants, sequence polymorphisms between different accessions of *A. thaliana* were mapped using high-density microarrays based on genomic (Clark et al., 2007) and transcribed sequences (Borevitz et al., 2007). According to Mitchell-Olds and Schmitt, (2006), these sequence polymorphisms in *A. thaliana* could be a valuable source for quantitative trait loci (QTL).

In our study, genomic DNA of inbred parental lines Col-0, Cvi and C24 was compared by array CGH mainly with the aim of excluding genomic regions showing sequence polymorphisms from further ChIP on chip data interpretation. Both types of tiling arrays that were used in our analysis were based on the genomic sequence of *A. thaliana* accession Col-0. Therefore, we could not detect sequences that only exist in C24 or Cvi. This explains why we found fewer CGH polymorphic tiles covering sequences with higher copy number in C24 and Cvi compared with those covering sequences with a lower copy number in C24 and Cvi (Fig. 11 and 15). Also, we could not determine possible transpositions of sequences in C24 or Cvi from their corresponding chromosomal position in Col-0.

Within the resolution of the tiling arrays used in our experiments, we were not able to detect single nucleotide polymorphism (SNP) between compared genotypes. However, increased and decreased copy number of sequences in C24 and Cvi compared with Col-0 were detectable. We estimate the minimum length of InDels that we could detect at around 165 base pair in the case of the *Arabidopsis* whole genome NimbleGen array and around 1 kb for the chromosome 4 tiling array. Approximately 20% dissimilarity between compared sequences is needed to produce differential signal ratios in microarray hybridization (Vincent Colot, personal

communication). In summary, these arguments make it likely that our CGH results rather underestimate the level of sequence polymorphism existing between different accessions of *A. thaliana*.

Patterns of sequence polymorphisms on chromosome 4 determined by chromosome 4 tiling microarray and by the whole genome *Arabidopsis* NimbleGen array were rather similar (Fig. 19). For instance, an approximately 59 kb long deletion on chromosome 4 in C24 was detected by both types of tiling arrays. Nevertheless, because of differences in design and size of tiles and microarrays, the detected polymorphic regions were not completely identical.

5.6.2. Genes related to disease resistance and transposable elements show a high level of sequence polymorphisms between different accessions of *A. thaliana*

Not surprisingly, comparison of C24 and Cvi sequences with Col-0 revealed that sequence polymorphisms exist between different accessions of *A. thaliana*. These polymorphisms are not evenly distributed across the genome. The pericentromeric heterochromatin and heterochromatic knob regions showed a higher level of sequence polymorphisms than other regions in comparison of C24 with Col-0 (Fig. 11). This finding is in agreement with a previous cytogenetic study in which authors showed that the heterochromatic knob region present in Col-0 is absent from C24 (Fransz et al., 2000). Also, sequence polymorphisms involved transposable elements more frequently than other sequence classes. A more extensive comparative analysis of the genomic sequences of *A. thaliana* accessions Col-0 and *Landsberg erecta* (*Ler*) revealed insertions and deletions (InDels) ranging between 2 bp and 38 kilobp. It was shown that a high proportion of the polymorphic regions with a length greater than 250 bp were the result of transposon insertions or excisions (The *Arabidopsis* Genome Initiative, 2000). It is also well documented in other studies that transposable elements are polymorphic between different *A. thaliana* accessions (Frank et al., 1998; Vaughn et al., 2007). Similarly, comparison of two inbred lines of maize revealed a high level of polymorphisms in transposable elements (Fu and Dooner, 2002). Our results agree with the previous observation that transposable elements evolve more rapidly than genes and gene-rich regions (Kazazian, 2004). Transposable elements can change the expression level of genes in their vicinity by *cis*-regulatory effects (Kashkush et al., 2003). The same might be true for chromatin modification

marks. Therefore, part of the observed gene expression differences between Col-0 and C24 and chromatin modification differences between Col-0, C24 and Cvi could be attributed to the presence or absence of different transposable elements.

Nevertheless, the polymorphisms that we identified by array-CGH between Col-0 and C24 or Cvi were not restricted to transposable elements. This confirms previous reports of polymorphisms in sequences other than transposable elements like for example variation in copy number of genes in different accessions of *A. thaliana* (Vaughn et al., 2007). Copy number variation in genes has been reported in humans and suggested to be a source of genetic diversity as well as spontaneous genetic disease (Sebat et al., 2004).

We also found that sequence polymorphisms affect members of some gene families - such as nucleotide-binding leucine-rich repeat (NB-LRR) genes and plant receptor-like kinase (RLK) genes- more often than members of other gene families, like transcription factors. Classification of genes which coincided with CGH polymorphic tiles based on biological processes indicated that genes involved in stress responses are over-represented. Copy number variations between different accessions of *A. thaliana* have also been reported for loci involved in plant disease resistance (Noel et al., 1999). There are around 150 genes annotated in the *A. thaliana* genome that encode nucleotide-NB-LRR type proteins. They belong to the plant disease resistance (R) genes and can recognize effectors of pathogens from diverse kingdoms in order to activate defence responses (Meyers et al., 2003). The above mentioned RLKs play a role in various processes, including development, disease resistance, and self-incompatibility (Shiu and Bleecker, 2001). Furthermore, consistent with our findings, sequence comparison of 20 different accessions of *A. thaliana* revealed that 60% of nucleotide-binding leucine-rich repeat (NB-LRR) genes and 15% of receptor-like kinase (RLK) genes contain at least one InDel event compared with the reference accession Col-0 (Clark et al., 2007). The enhanced rate of polymorphisms in genes responsible for stress response and disease resistance demonstrates that different gene families evolve at different rates.

5.6.3. *A. thaliana* accession Cvi shows slightly more sequence polymorphism than C24 compared with Col-0

The number of CGH polymorphic tiles detected with *Arabidopsis* whole genome NimbleGene microarrays in a comparison between Col-0 and C24 was 5.54% of total tiles while it was 6% (6.01% for first technical replicate, 6.35% in second biological replicate) in a comparison of Cvi and Col-0 sequences. Considering the fact that 97.43% of CGH polymorphic tiles were common between first and second technical replicates, our CGH result indicated that the number of CGH polymorphic tiles for Cvi is slightly higher than for C24 compared with Col-0. This includes an increased copy number of mitochondrial-like sequences integrated into the genome of Cvi and C24 compared with Col-0. The longest insertion in Cvi had a length about 7 kb, whereas there was only one in C24, which was less than 1 kb. Our observations support those of Schmid et al. (2003) in that SNP-based comparison of the average sequence divergence between Col-0 and different *A. thaliana* accessions (including C24) indicated that Cvi-0 is more divergent from Col-0 than other accessions. The same authors suggested that this increased genetic distance is probably due to a few highly divergent loci. DNA methylation analysis of different *A. thaliana* accessions also showed that Col-0 and C24 have very similar methylation patterns, while Cvi showed a different pattern of this DNA modification (Vaughn et al., 2007).

Nevertheless, because our tiling array-based analysis could not detect sequences that only exist in Cvi or C24 and also would not be sensitive to small InDels and SNPs, it is not surprising that our CGH data do not completely reproduce the data obtained by other methods. In the future, whole genome sequence analysis of different *A. thaliana* accessions will allow a more detailed comparison of different accessions.

5.7. Variations in the levels of chromatin modifications marks exist among different accessions of *A. thaliana*

Our ChIP on chip analysis revealed that for similar genomic regions (within the resolution of arrays used in our experiments) chromatin modification patterns can differ between different accessions of *A. thaliana*.

The interpretation of chromatin modification patterns is not straightforward. On the one hand, certain chromatin modifications can be seen as determinants of gene activity. This relationship is well-documented in cases where transposable elements are inserted next to functional genes and are thought to hamper gene expression by attracting inhibitory chromatin modifications. This situation has been demonstrated for the different expression level of *FLOWERING LOCUS C (FLC)* in different accessions of *A. thaliana*. The insertion of a transposable element in the intron of *FLC* in accession *Ler* caused less transcription of this locus and consequently earlier flowering of *Ler* in compared with *Col-0* (Liu et al., 2004). Another case was demonstrated for expression level changes of the *Agouti* locus which regulates coat colour in mice (Morgan et al., 1999). Depending on the presence and the epigenetic state of a transposon upstream of *Agouti*, the expression level of *Agouti* was variable and caused the coat colour of mice to range from solid yellow to a pigment pattern that was similar to the wild type. No epigenetic variation and consequently colour coat variation occurred when the transposable element in the 5' region of the *Agouti* gene was absent. The effect was dependent on the *Agouti* locus structure. Nevertheless, among individuals with the identical structure of the *Agouti* locus, considerable variation of *Agouti* expression could occur. Therefore, the variation in phenotype could not be solely explained by the genotype. Similarly, the differences in the level of histone H3 methylation between *Col-0*, *Cvi* and *C24* could be attributed to sequences differences as copy number polymorphisms for genes and transposable elements were observed in our CGH experiments. Nevertheless, some contribution of a possible heritability of particular chromatin states to the maintenance of the differences can not be excluded.

One the other hand, changes in chromatin modifications can be seen as consequences of changes in gene activity. Indeed, we observed different expression levels for seven candidate genes between *Col-0* and *C24* (Fig. 9). Hence, differences in the transcriptional activity of genes, transposable elements and intergenic regions between *A. thaliana* accessions *Col-0* and *C24* or *Cvi* could lead to the establishment of different patterns of H3K4me2 and H3K27me3 across the genomes.

Transposable elements were over-represented in sequences covered by H3K4me2 polymorphic tiles (Fig. 26). H3K4me2 is probably associated more with active chromatin than with heterochromatin (Jenuwein and Allis, 2001). Therefore, the high representation of TEs could

reflect high variability of transcriptional activity of transposable elements between genotypes. In contrast, genic regions were overrepresented in sequences covered by H3K27me3 polymorphic tiles (Fig. 29 and 33). This observation confirms the finding of Zhang et al. (2007) that many genes (approximately 4400) of *A. thaliana* is targeted by H3K27me3. In contrast to animals, in plants H3K27me3 is mainly found in transcribed regions of single genes. The maintenance of this modification mark seems to be independent of other epigenetic marks.

Small sequence polymorphisms, like point mutations or small InDels, which were not detectable in our CGH experiments, could also have some influence in the gene activity. For instance, in one accession a gene could have become a pseudogene and hence inactive and therefore would show a different set of histone modification marks.

The amount of variation for histones H3K27me3 and H3K4me2 between Col-0 and Cvi was different with polymorphisms being consistently more often found for H3K27me3 than for H3K4me2 - 10.08% of total tiles in 1BR and 4.18% of total tiles in 2BR for H3K27me3 compared to 2.85% of total tiles in 1BR and 0.72% of total tiles in 2BR for H3K4me2. Unfortunately, no full dataset is available for H3K4me2 between Col-0 and C24. For H3K27me3, the numbers of polymorphic tiles between Cvi and Col-0 and those polymorphic between C24 and Col-0 could be compared. The number for Cvi and Col-0 was lower than for C24 and Col-0 (10.08% and 4.18% of total tiles in 1BR and 2BR, respectively for Col-0 and Cvi and 10.38% and 15.7% of total tiles in 1BR and 2BR, respectively for Col-0 and C24). This difference is probably because the genomic sequences of the accessions differ and consequently target sites for H3K27me3 differ too. Nevertheless, the slightly higher number of CGH polymorphic tiles between Cvi and Col-0 (6.0%) compared with C24 and Col-0 (5.5%) may not support this interpretation.

In conclusion, the observed accession-dependent differences in chromatin modification variation suggest that another layer of complexity exists, which possibly plays a role in diversifying different accessions of *A. thaliana*.

5.8. Histone H3K4me2 is a stable histone modification mark in response to intraspecific hybridization

Comparison of histone H3K4me2 levels in hybrid offspring of Col-0xC24 with its inbred parental lines by ChIP on chip analysis using chromosome 4 tiling microarray identified 12 sporadic intraspecific hybridization responsive tiles, which covered about 12 kb of chromosome 4. However, ChIP-quantitative PCR did not confirm the identified H3K4me2 intraspecific responsive tiles. This suggests that the identified tiles were probably false positives reflecting the noise of the Chip on chip method.

Genome wide ChIP on chip analysis of H3K4me2 levels in Col-0xCvi compared with its inbred parental lines revealed some H3K4me2 hybridization responsive regions across the genome. The total length of these regions was approximately 88 kb (525 tiles total) and about 12 kb (a total of 69 tiles) in the first and second biological replicate, respectively. Extrapolating the data of false positive polymorphic tiles (12 kb) obtained from the above ChIP on chip determination on chromosome 4 to the entire genome would result in an expectation of polymorphic regions of a total length of 60 kb. This agrees well with the length of H3K4me2 hybridization responsive regions observed in Col-0xCvi for first and second biological replicates. Moreover, the sporadic H3K4me2 intraspecific hybridization responsive regions in Col-0xCvi were not identical for the first and second biological replicate. These suggest that the observed H3K4me2 intraspecific responsive regions are likely to be false positives and should be regarded as artefacts.

The observed stable behaviour of H3K4me2 is consistent with our MSAP-based DNA methylation data. MSAP analysis demonstrated almost no change of DNA methylation occurs in response to intraspecific hybridization. It has been reported that DNA methylation can prevent methylation of histone H3K4me2 (Hashimshony et al., 2003; Lande-Diner et al., 2007). In conclusion, H3K4me2 is a stable chromatin modification in response to intraspecific hybridization.

5.9. Histone H3K27me3 is more dynamic than H3K4me2 in response to intraspecific hybridization

In contrast to H3K4me2, distribution of H3K27me3 showed a higher variation between different accessions of *A. thaliana*. In addition, H3K27me3 showed higher dynamics in response to intraspecific hybridization. We found hundreds of H3K27me3 intraspecific hybridization responsive tiles in Col-0xC24 and Col-0xCvi.

H3K27me3 is known to be largely restricted to transcribed regions and might play a critical role for regulating developmentally important genes (Zhang et al., 2007). The observed deviation of histone H3K27me3 levels in hybrid offspring could potentially change gene expression levels. Consistent observations were made for intraspecific and interspecific *Arabidopsis* hybrids where differential epigenetic regulation of circadian clock genes and their reciprocal regulators was correlated with increased production of chlorophyll and starch in hybrid offspring under identical environmental conditions (Ni et al., 2009). This findings show how even small epigenetic changes in hybrids could result in profound phenotypic effects. Therefore, epigenetic variations between inbred parental lines and newly arisen epigenetic modification patterns in hybrid offspring in response to the intraspecific hybridization could explain at least partly the observed heterosis for biomass in Col-0xC24 intraspecific hybrids of *A. thaliana* (Meyer et al., 2004).

Histone H3K27me3 showed a higher frequency of deviation from the expected mid-parental value in hybrid offspring relative to H3K4me2. This could be attributed to a higher sensitivity of the establishment of H3K27me3 to the “genomic shock” exerted by intraspecific hybridization, conceivably through *cis*- and *trans*-acting effects. This may explain also the lack of overlap between H3K27me3 intraspecific hybridization responsive tiles in Col-0xC24 and Col-0xCvi. The sequence differences present between Cvi and C24 could generate distinct interaction networks of *cis* and *trans* elements when combined with Col-0 sequences that could cause alteration of H3K27me3 in distinct genomic regions.

In conclusion, our results suggest that only minor changes of chromatin properties occur in response to intraspecific hybridization in *A. thaliana*. It will be interesting to be test whether this observation will also hold true for crop species showing heterosis.

5.10. Significance of findings with regard to models explaining the mechanism of heterosis

On the one hand, the observed sequence polymorphisms between different accessions of *A. thaliana* favour the complementary dominance theory of heterosis partly to explain the reported heterosis for biomass production in crosses between different accessions of *A. thaliana* (Meyer et al., 2004). On the other hand, the new patterns of epigenetic modifications in hybrid offspring are consistent with the epistasis theory of heterosis which postulates that different genetic loci should interact with each other in order to generate new levels of gene expression in hybrid offspring. Since H3K27me3 is involved in the regulation of gene expression, one could imagine that new patterns of H3K27me3 might also cause overdominance gene expression patterns in hybrid offspring. In conclusion, heterosis remains a very complex phenomenon possibly involving different mechanisms. In addition to the sequence polymorphisms between inbred parental lines which could influence gene expression in hybrid offspring through *cis* and or *trans* acting effects, our results revealed chromatin modification variations between inbred parental lines and newly arisen patterns of chromatin modifications in hybrid offspring. Hybrids could use these variations for altering the amplitude of gene expression resulting in favourable phenotypes.

6. Outlook

6.1. Confirmation of histone H3K27me3 level changes detected by ChIP on chip analysis in hybrid offspring by ChIP-quantitative PCR

We intend to confirm the results of our ChIP on chip experiments for some of the H3K27me3 hybridization responsive tiles by ChIP-quantitative PCR. Specific primer pairs with similar efficiency will be designed and used for amplification from the same IP DNA preparations that were used for ChIP on chip experiments. Progeny from reciprocal crosses will be included to determine whether the dynamics of the histone H3K27me3 response to intraspecific hybridization depend on the direction of the cross.

6.2. Expression analysis of genes colocalizing with H3K27me3 hybridization responsive tiles

As H3K27me3 is a histone modification involved in the silencing of developmentally important genes, we plan to analyze the expression level of genes which are located in regions with altered levels of H3K27me3 in hybrid offspring. Quantitative RT-PCR will be used to determine the transcript levels of these genes.

6.3. Analysis of the heritability of H3K27me3 level changes induced by intraspecific hybridization

To investigate whether intraspecific hybridization induced H3K27me3 changes are heritable, the levels of H3K27me3 at these regions will be investigated via ChIP-quantitative PCR in the F₂ generation.

6.4. Analysis of the activity of transposable elements in inbred parental lines and their hybrid offspring

Results of ChIP on chip for H3K4me2 suggests that possibly the activity of some transposable elements could vary between different accessions of *A. thaliana* which could possibly cause the observed variation in chromatin modifications. Therefore, the expression level of some selected transposable elements with different levels of histone H3K4me2 in accessions of *A. thaliana* will be determined by quantitative RT-PCR. To test whether silent transposable elements in one parent will remain suppressed in hybrid offspring, the cleavage amplified polymorphic sequence (CAPS) approach will be used for determining the allele expression level of transposable elements.

6.5. Systematically analyze selected regions of the *A. thaliana* genome to determine the changes in the level of different histone modifications and their effects on gene expression levels in response to intraspecific hybridization

To investigate whether other histone modifications respond to intraspecific hybridization, ChIP-quantitative PCR for additional histone modification marks will determine the level of these histone modification marks in inbred parental lines and their hybrid offspring. Chromatin immuno-precipitation will be done using different specific antibodies for methylation, acetylation and phosphorylation of different residues of histones H3 and H4. The levels of these histone modifications will be quantified for selected regions of the *A. thaliana* genome (which contain genes, transposable elements, siRNA and others). The effect of possible changes in chromatin modification levels in response to intraspecific hybridization on gene expression will be tested by quantitative RT-PCR.

7. Summary

Endopolyploidy Levels were measured for inbred parental lines of Col-0 and C24 and their hybrid offspring. The proportion of nuclei with a higher ploidy level increased in tissues with advanced developmental stages. However, no differences in endopolyploidization were obvious between Col-0, C24 and their reciprocal hybrids. This suggests that intraspecific hybridization is not associated with an altered level of endopolyploidy.

The nucleolus area was determined in 2C and 4C nuclei in inbred parental accessions of Col-0 and C24 and their hybrids to quantify and compare the relative rRNA gene activity. The average nucleolus area was highly variable within all samples tested and the nucleolus area of intraspecific hybrids was similar to that of their parental nuclei and does not , therefore, indicate an increase in rRNA gene activity in hybrid offspring.

Indirect immunostaining was performed on nuclei with a defined ploidy level (4C) to determine the distribution pattern of DNA and histone methylation in inbred parental lines of Col-0, C24 and Cvi and their hybrid offspring. Localization of immunofluorescence signals with DAPI-stained heterochromatic chromocenters was observed for 5mC and histone H3K9me₂, while signals for H3K4me₂ and H3K27me₃ were dispersed in euchromatin of interphase nuclei and were excluded from chromocenters. Regardless of which chromatin modification was analyzed, there were no obvious differences in the distribution and intensity of immunofluorescence signals between Col-0, C24 and Cvi and their reciprocal hybrids. Hence, no major reorganization of DNA and histone methylation marks occurs after intraspecific hybridization.

Overall distribution of methylated DNA in intraspecific hybrid offspring of Col-0xC24 and C24xCol-0 was compared with their inbred parents using methylation sensitive amplified polymorphism (MSAP) analysis. Most of the MSAP signals analyzed showed patterns that did

not indicate any methylation changes in hybrid offspring compared with their inbred parents. Only about 3% of the signals showed increased methylation in hybrids. Regardless of genotype the overall DNA methylation in *A. thaliana* increased during the course of plant development. This suggests that DNA methylation is a largely stable chromatin modification in response to intraspecific hybridization.

Methylation levels at the 5' regulatory elements of seven selected genes in inbred parental lines of Col-0 and C24 and their alleles in hybrid offspring were determined by methylation-sensitive restriction cleavage followed by PCR. No change of methylation status was observed in the reciprocal hybrids compared with the inbred parental lines for all genes.

The relative transcript levels of the seven selected genes were determined by quantitative RT-PCR in inbred parental lines of Col-0 and C24 and their hybrid offspring. No change in the allele-specific transcript status was observed in the reciprocal hybrids compared with the inbred parental lines. The transcript levels of genes were always close to the calculated mid-parent values in the intraspecific hybrids.

As a necessary initial experiment for interpreting ChIP on chip data, genomic sequences of C24 and Cvi were compared with sequences of Col-0. Comparative genomic hybridization (CGH) showed that natural sequence polymorphisms exist between different accessions of *A. thaliana*. Sequence polymorphisms were not equally distributed across the genome. Polymorphisms for transposable elements and some gene families responsible for disease resistance and stress response were overrepresented. In contrast, other genic regions and their regulatory elements in 5' and 3' UTRs were underrepresented. Based on our CGH experiments Cvi was slightly more divergent than C24 compared with the genome of Col-0.

The distribution of histone H3K4me2 and H3K27me3 in inbred parental lines of Col-0, C24 and Cvi was compared to their hybrid offspring using ChIP on chip. Analysis revealed that as

well as sequence polymorphisms, chromatin modification variations exist among different accessions of *A. thaliana*. The range of these variations for histone H3K27me3 was higher than for H3K4me2. Among inbred parental lines transposable elements were overrepresented in regions with variation for histone H3K4me2, whereas genes were overrepresented for regions with variations for histone H3K27me3 between inbred parental lines.

In response to intraspecific hybridization, histone H3K4me2 behaved as a stable mark with additive inheritance in hybrid offspring. In contrast, limited genome regions showed altered levels of histone H3K27me3 deviating from the expected mid-parental value in hybrid offspring. Hence, histone H3K27me3 shows higher dynamics in response to intraspecific hybridization.

Based on our CGH and ChIP on chip data, all postulated major mechanisms for heterosis could have contributed to the reported hybrid vigour after intraspecific hybridization in *A. thaliana*. Sequence and chromatin modification level differences among accessions agree with the complementary dominance theory. Newly arisen modification levels in hybrid offspring indicate epistatic interactions between different *trans*-acting elements in hybrid offspring. Since H3K27me3 is involved in gene regulation, deviation from expected mid-parental values as overdominance is possible.

8. Zusammenfassung

Der Grad der Endopolyploidisierung wurde für die *A. thaliana* Linien Col-0 und C24, und deren intraspezifischen Hybriden bestimmt. Der Grad der Endopolyploidy korrelierte mit zunehmendem Entwicklungsstand der Pflanzen, wobei sich die untersuchten Genotypen nicht unterschieden. Damit kann gefolgert werden, dass die intraspezifische Hybridisierung sich nicht auf den Grad der Endopolyploidisierung auswirkt.

Die Nukleolusgröße von 2C und 4C Zellkernen wurde vermessen, um indirekt die relative rRNA Genaktivität der Linien Col-0 und C24, und deren intraspezifischen Hybriden miteinander vergleichen zu können. Unabhängig vom Genotyp wurde eine große Variabilität der Nukleolusgröße festgestellt. Die rRNA Gene intraspezifischer Hybriden zeigten keine erhöhte Aktivität.

Die Verteilungsmuster methylierter DNA und methylierter Histone wurde zwischen den Linien Col-0, C24 und Cvi, und deren intraspezifischen Hybriden mit Hilfe der indirekten Immunfluoreszenz an 4C Zellkernen miteinander verglichen. 5mC- und Histon H3K9me2-spezifische Signale korrelierten mit der Position DAPI-positiver Chromozentren. H3K4me2- und H3K27me3-spezifische Signale dagegen überlappten mit euchromatischen Bereichen. Für keine der untersuchten Chromatinmodifikationen wurden Genotyp-spezifische Verteilungsmuster identifiziert. Es kann gefolgert werden, dass keine mikroskopisch detektierbaren Verlagerungen dieser Modifikationen im Zuge der intraspezifischen Hybridisierung stattfinden.

Mit Hilfe der methylierungssensitiven DNA-Polymorphismen Amplifikationsmethode (MSAP) wurde die Verteilung methylierter DNA zwischen den elterlichen Linien Col-0 und C24, und deren intraspezifischen Hybriden (Col-0xC24 und C24xCol-0) verglichen. Die meisten MSAP-Signale waren zwischen Eltern und Hybriden identisch. Nur 3% der Signale zeigten eine verstärkte Methylierung in Hybriden. Unabhängig vom Genotyp wurde mit zunehmendem Entwicklungsstand der Pflanzen eine Methylierungszunahme nachgewiesen. Es wird gefolgert, dass die DNA Methylierung sich im Prozess der intraspezifischen Hybridisierung nur unwesentlich verändert.

Der DNA-Methylierungszustand des 5'-regulierenden Bereichs von sieben vorselektierten Genen wurde zwischen den Linien Col-0 und C24, und deren reziproken Hybriden, mit Hilfe einer Spaltung mit methylierungssensitiven Restriktionsenzymen und anschließender PCR verglichen. Im Zuge der intraspezifischen Hybridisierung wurde keine Methylierungsveränderung nachgewiesen.

Die relative Transkriptmenge von sieben vorselektierten Genen wurde für die Linien Col-0 und C24, und deren Hybriden bestimmt. Der Allel-spezifische Transkriptionszustand reziproker Hybriden und deren elterlicher Linien war identisch. Der für intraspezifische Hybriden gemessene Transkriptionsniveau entsprach dem rechnerischen Durchschnitt der Transkriptionsniveaus der Eltern.

Voraussetzung für die Interpretation der „ChIP on chip“ Ergebnisse ist der vorherige Sequenzvergleich der Linien C24 und Cvi mit der genomischen Sequenz der Referenzlinie Col-0. Mittels vergleichender genomischer Hybridisierung (CGH) wurden zwischen den zu untersuchenden Linien Sequenzunterschiede identifiziert. Es zeigten sich Sequenztyp-abhängige Unterschiede in der Zusammensetzung der genomischen DNA. Polymorphismen wurden besonders für transposable Elemente und für einige Genfamilien, welche für Krankheitsresistenzen und Stressantwort spezifisch sind, identifiziert. Im Unterschied dazu waren andere Gene und deren 5'- und 3'-regulierenden Regionen unterrepräsentiert. Auf den CGH Daten basierend kann gefolgert werden, dass sich Cvi von Col-0 stärker unterscheidet als C24 von Col-0.

Die Verteilung der Histonmodifikationen H3K4me2 und H3K27me3 wurde mit Hilfe der „ChIP on chip“ Methode zwischen den elterlichen Linien Col-0, C24 und Cvi, und deren intraspezifischen Hybriden verglichen. Die Analyse zeigte, dass Unterschiede in der Verteilung dieser Chromatinmodifikationen zwischen den elterlichen Linien existieren. H3K27me3 variierte stärker als H3K4me2. Abweichungen in der Verteilung von H3K4me2 waren für transposable Elemente elterlicher Linien besonders ausgeprägt. Dagegen zeigten Gene eine höhere Variabilität für H3K27me3.

H3K4me2 zeigte eine additive Vererbung und verhielt sich in Reaktion auf eine intraspezifische Hybridisierung stabil. Im Unterschied dazu zeigten begrenzte Genombereiche eine vom erwarteten elterlichen Durchschnitt abweichende H3K27me3-Verteilung in Hybriden.

Es kann gefolgert werden, dass H3K27me3 dynamischer auf den Prozess der intraspezifischen Hybridisierung reagiert.

Basierend auf den CGH und "ChIP on chip" Ergebnissen kann postuliert werden, dass unterschiedliche Mechanismen bei der Ausprägung heterotischer Hybridpflanzen eine Rolle spielen können. Die beobachteten Unterschiede in der genomischen Sequenz und in der Verteilung von Chromatinmodifikationen zwischen den unterschiedlichen Genotypen stimmen mit der Komplementärdominanz-Theorie überein. Die beobachteten Chromatinveränderungen suggerieren epistatische Interaktionen zwischen *trans*-aktiven Elementen in intraspezifischen Hybriden. Da H3K27me3 bei der Regulation der Genaktivität involviert ist, können Abweichungen vom erwarteten Durchschnitt der Elternwerte in Hybriden als eine Form der Überdominanz gesehen werden.

9. Literature

- Axelsson, T., Bowman, C.M., Sharpe, A.G., Lydiate, D.J., and Lagercrantz, U.** (2000). Amphidiploid *Brassica juncea* contains conserved progenitor genomes. *Genome* **43**, 679-688.
- Barow, M., and Meister, A.** (2003). Endopolyploidy in seed plants is differently correlated to systematics, organ, life strategy and genome size. *Plant, Cell and Environment* **26**, 571-584.
- Baumel, A., Ainouche, M., Kalendar, R., and Schulman, A.H.** (2002). Retrotransposons and genomic stability in populations of the young allopolyploid species *Spartina anglica* C.E. Hubbard (Poaceae). *Molecular biology and evolution* **19**, 1218-1227.
- Berger, S.L.** (2007). The complex language of chromatin regulation during transcription. *Nature* **447**, 407-412.
- Bernatavichute, Y.V., Zhang, X., Cokus, S., Pellegrini, M., and Jacobsen, S.E.** (2008). Genome-wide association of histone H3 lysine nine methylation with CHG DNA methylation in *Arabidopsis thaliana*. *PLoS ONE* **3**, e3156.
- Bernstein, B.E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D.K., Huebert, D.J., McMahon, S., Karlsson, E.K., Kulbokas, E.J., 3rd, Gingeras, T.R., Schreiber, S.L., and Lander, E.S.** (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* **120**, 169-181.
- Birchler, J.A., Auger, D.L., and Riddle, N.C.** (2003). In search of the molecular basis of heterosis. *Plant Cell* **15**, 2236-2239.
- Bohlander, S.K., Espinosa, R., Lebeau, M.M., Rowley, J.D., and Diaz, M.O.** (1992). A Method for the rapid sequence-Independent amplification of microdissected chromosomal material. *Genomics* **13**, 1322-1324.
- Borevitz, J.O., Hazen, S.P., Michael, T.P., Morris, G.P., Baxter, I.R., Hu, T.T., Chen, H., Werner, J.D., Nordborg, M., Salt, D.E., Kay, S.A., Chory, J., Weigel, D., Jones, J.D., and Ecker, J.R.** (2007). Genome-wide patterns of single-feature polymorphism in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **104**, 12057-12062.
- Brubaker, C., Brown, A., Stewart, J., Kilby, M., and Grace, J.** (1999). Production of fertile hybrid germplasm with diploid Australian *Gossypium* species for cotton improvement. *Euphytica* **108**, 199-213.
- Caspersson, T.** (1950). Cell Growth and Cell Function, a Cytochemical Study. *Growth* **14**, 353-356.
- Castiglione, M.R., Cremonini, R., and Frediani, M.** (2002). DNA methylation patterns on plant chromosomes. *Caryologia* **55**, 275-282.
- Chan, S.W., Henderson, I.R., and Jacobsen, S.E.** (2005). Gardening the genome: DNA methylation in *Arabidopsis thaliana*. *Nature reviews* **6**, 351-360.
- Chen, L., and Chen, J.** (2008). Changes of cytosine methylation induced by wide hybridization and allopolyploidy in Cucumis. *Genome* **51**, 789-799.
- Chen, Z.J., and Pikaard, C.S.** (1997). Transcriptional analysis of nucleolar dominance in polyploid plants: biased expression/silencing of progenitor rRNA genes is developmentally regulated in *Brassica*. *Proc. Natl. Acad. Sci. USA* **94**, 3442-3447.
- Chen, Z.J., and Ni, Z.** (2006). Mechanisms of genomic rearrangements and gene expression changes in plant polyploids. *Bioessays* **28**, 240-252.

- Chen, Z.J., Comai, L., and 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. *Proc. Natl. Acad. Sci. USA* **95**, 14891-14896.
- Chomczynski, P., and Mackey, K.** (1995). Modification of the TRI reagent procedure for isolation of RNA from polysaccharide- and proteoglycan-rich sources. *Biotechniques* **19**, 942-945.
- Clark, R.M., Schweikert, G., Toomajian, C., Ossowski, S., Zeller, G., Shinn, P., Warthmann, N., Hu, T.T., Fu, G., Hinds, D.A., Chen, H., Frazer, K.A., Huson, D.H., Scholkopf, B., Nordborg, M., Ratsch, G., Ecker, J.R., and Weigel, D.** (2007). Common sequence polymorphisms shaping genetic diversity in *Arabidopsis thaliana*. *Science* **317**, 338-342.
- Cockerham, C.C., and Zeng, Z.B.** (1996). Design III with marker loci. *Genetics* **143**, 1437-1456.
- Cokus, S.J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C.D., Pradhan, S., Nelson, S.F., Pellegrini, M., and Jacobsen, S.E.** (2008). Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature* **452**, 215-219.
- Comai, L.** (2000). Genetic and epigenetic interactions in allopolyploid plants. *Plant molecular biology* **43**, 387-399.
- Comai, L.** (2005). The advantages and disadvantages of being polyploid. *Nature reviews* **6**, 836-846.
- Comai, L., Tyagi, A.P., Winter, K., Holmes-Davis, R., Reynolds, S.H., Stevens, Y., and Byers, B.** (2000). Phenotypic instability and rapid gene silencing in newly formed *Arabidopsis* allotetraploids. *Plant Cell* **12**, 1551-1568.
- Comings, D.E., and MacMurray, J.P.** (2000). Molecular heterosis: a review. *Molecular genetics and metabolism* **71**, 19-31.
- de Vienne, D., Bost, B., Fievet, J., Zivy, M., and Dillmann, C.** (2001). Genetic variability of proteome expression and metabolic control. *Plant Physiology and Biochemistry* **39**, 271-283.
- DeAngelis, J.T., Farrington, W.J., and Tollefsbol, T.O.** (2008). An overview of epigenetic assays. *Molecular biotechnology* **38**, 179-183.
- Derenzini, M., Farabegoli, F., and Trere, D.** (1992). Relationship between interphase AgNOR distribution and nucleolar size in cancer cells. *The Histochemical journal* **24**, 951-956.
- Derenzini, M., Trere, D., Pession, A., Govoni, M., Sirri, V., and Chieco, P.** (2000). Nucleolar size indicates the rapidity of cell proliferation in cancer tissues. *The Journal of pathology* **191**, 181-186.
- Duvick, D.N.** (1999). Hazard identification of agricultural biotechnology. *Science* **286**, 418-419.
- Duvick, D.N.** (2001). Biotechnology in the 1930s: the development of hybrid maize. *Nature reviews* **2**, 69-74.
- Earley, K., Lawrence, R.J., Pontes, O., Reuther, R., Enciso, A.J., Silva, M., Neves, N., Gross, M., Viegas, W., and Pikaard, C.S.** (2006). Erasure of histone acetylation by *Arabidopsis* HDA6 mediates large-scale gene silencing in nucleolar dominance. *Genes & development* **20**, 1283-1293.

- Elser, J.J., Sterner, R.W., Gorokhova, E., Fagan, W.F., Markow, T.A., Cotner, J.B., Harrison, J.F., Hobbie, S.E., Odell, G.M., and Weider, L.J.** (2000). Biological stoichiometry from genes to ecosystems. *Ecology Letters* **3**, 540-550.
- Feldman, M., Liu, B., Segal, G., Abbo, S., Levy, A.A., and Vega, J.M.** (1997). Rapid elimination of low-copy DNA sequences in polyploid wheat: a possible mechanism for differentiation of homoeologous chromosomes. *Genetics* **147**, 1381-1387.
- Feschotte, C., Jiang, N., and Wessler, S.R.** (2002). Plant transposable elements: where genetics meets genomics. *Nature reviews* **3**, 329-341.
- Fischle, W., Wang, Y., and Allis, C.D.** (2003). Histone and chromatin cross-talk. *Current opinion in cell biology* **15**, 172-183.
- Frank, M.J., Preuss, D., Mack, A., Kuhlmann, T.C., and Crawford, N.M.** (1998). The *Arabidopsis* transposable element Tag1 is widely distributed among *Arabidopsis* ecotypes. *Mol Gen Genet* **257**, 478-484.
- Fransz, P.F., Armstrong, S., de Jong, J.H., Parnell, L.D., van Drunen, C., Dean, C., Zabel, P., Bisseling, T., and Jones, G.H.** (2000). Integrated cytogenetic map of chromosome arm 4S of *A. thaliana*: structural organization of heterochromatic knob and centromere region. *Cell* **100**, 367-376.
- Fu, H., and Dooner, H.K.** (2002). Intraspecific violation of genetic colinearity and its implications in maize. *Proc. Natl. Acad. Sci. USA* **99**, 9573-9578.
- Fuchs, J., Demidov, D., Houben, A., and Schubert, I.** (2006). Chromosomal histone modification patterns--from conservation to diversity. *Trends in plant science* **11**, 199-208.
- Fuks, F.** (2005). DNA methylation and histone modifications: teaming up to silence genes. *Current opinion in genetics & development* **15**, 490-495.
- Galbraith, D.W., Harkins, K.R., and Knapp, S.** (1991). Systemic Endopolyploidy in *Arabidopsis thaliana*. *Plant physiology* **96**, 985-989.
- Galili, G., and Feldman, M.** (1984). Intergenomic suppression of endosperm protein genes in common wheat. *Canadian Journal of Genetics and Cytology* **26**, 651-656.
- Gehring, M. and Henikoff, S.** (2008) DNA Methylation and Demethylation in *Arabidopsis*: July 9, 2008. *The Arabidopsis Book*. Rockville, MD: American Society of Plant Biologists doi: 10.1199/tab.0102 <http://www.aspb.org/publications/arabidopsis/>
- Gendrel, A.V., Lippman, Z., Martienssen, R., and Colot, V.** (2005). Profiling histone modification patterns in plants using genomic tiling microarrays. *Nature methods* **2**, 213-218.
- Godde, J.S., and Ura, K.** (2008). Cracking the enigmatic linker histone code. *Journal of biochemistry* **143**, 287-293.
- Grummt, I., and Pikaard, C.S.** (2003). Epigenetic silencing of RNA polymerase I transcription. *Nat Rev Mol Cell Biol* **4**, 641-649.
- Guo, M., Rupe, M.A., Yang, X., Crasta, O., Zinselmeier, C., Smith, O.S., and Bowen, B.** (2006). Genome-wide transcript analysis of maize hybrids: allelic additive gene expression and yield heterosis. *TAG*.
- Hashimshony, T., Zhang, J., Keshet, I., Bustin, M., and Cedar, H.** (2003). The role of DNA methylation in setting up chromatin structure during development. *Nature genetics* **34**, 187-192.
- Hegarty, M.J., and Hiscock, S.J.** (2008). Genomic clues to the evolutionary success of polyploid plants. *Curr Biol* **18**, R435-444.

- Hegarty, M.J., Barker, G.L., Wilson, I.D., Abbott, R.J., Edwards, K.J., and Hiscock, S.J.** (2006). Transcriptome shock after interspecific hybridization in senecio is ameliorated by genome duplication. *Curr Biol* **16**, 1652-1659.
- Henderson, I.R., and Jacobsen, S.E.** (2007). Epigenetic inheritance in plants. *Nature* **447**, 418-424.
- Hirochika, H., Okamoto, H., and Kakutani, T.** (2000). Silencing of retrotransposons in arabidopsis and reactivation by the *ddm1* mutation. *The Plant Cell* **12**, 357-369.
- Hizume, M., Sato, S., and Tanaka, A.** (1980). A highly reproducible method of nucleolus organizing regions staining in plants. *Stain technology* **55**, 87-90.
- Hoecker, N., Keller, B., Muthreich, N., Chollet, D., Descombes, P., Piepho, H.P., and Hochholdinger, F.** (2008). Comparison of Maize (*Zea mays* L.) F1-Hybrid and Parental Inbred Line Primary Root Transcriptomes Suggests Organ-Specific Patterns of Nonadditive Gene Expression and Conserved Expression Trends. *Genetics* **179**, 1275-1283.
- Houben, A., Demidov, D., Gernand, D., Meister, A., Leach, C.R., and Schubert, I.** (2003). Methylation of histone H3 in euchromatin of plant chromosomes depends on basic nuclear DNA content. *Plant J* **33**, 967-973.
- Hsieh, T.F., Hakim, O., Ohad, N., and Fischer, R.L.** (2003). From flour to flower: how Polycomb group proteins influence multiple aspects of plant development. *Trends in plant science* **8**, 439-445.
- Hubbell, H.R.** (1985). Silver staining as an indicator of active ribosomal genes. *Stain technology* **60**, 285-294.
- Jackson, J.P., Lindroth, A.M., Cao, X., and Jacobsen, S.E.** (2002). Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* **416**, 556-560.
- Jackson, J.P., Johnson, L., Jasencakova, Z., Zhang, X., PerezBurgos, L., Singh, P.B., Cheng, X., Schubert, I., Jenuwein, T., and Jacobsen, S.E.** (2004). Dimethylation of histone H3 lysine 9 is a critical mark for DNA methylation and gene silencing in *Arabidopsis thaliana*. *Chromosoma* **112**, 308-315.
- Jacobs, S.A., Taverna, S.D., Zhang, Y., Briggs, S.D., Li, J., Eissenberg, J.C., Allis, C.D., and Khorasanizadeh, S.** (2001). Specificity of the HP1 chromo domain for the methylated N-terminus of histone H3. *The EMBO journal* **20**, 5232-5241.
- Jacquet, B., Canet, V., Giroud, F., Montmasson, M.P., and Brugal, G.** (2001). Quantitation of AgNORs by flow versus image cytometry. *J Histochem Cytochem* **49**, 433-438.
- Jenuwein, T., and Allis, C.D.** (2001). Translating the histone code. *Science (New York, N.Y)* **293**, 1074-1080.
- Jonsson, G., Naylor, T.L., Vallon-Christersson, J., Staaf, J., Huang, J., Ward, M.R., Greshock, J.D., Luts, L., Olsson, H., Rahman, N., Stratton, M., Ringner, M., Borg, A., and Weber, B.L.** (2005). Distinct genomic profiles in hereditary breast tumors identified by array-based comparative genomic hybridization. *Cancer research* **65**, 7612-7621.
- Kashkush, K., Feldman, M., and Levy, A.A.** (2002). Gene loss, silencing and activation in a newly synthesized wheat allotetraploid. *Genetics* **160**, 1651-1659.
- Kashkush, K., Feldman, M., and Levy, A.A.** (2003). Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. *Nature genetics* **33**, 102-106.
- Kazazian, H.H., Jr.** (2004). Mobile elements: drivers of genome evolution. *Science* **303**, 1626-1632.

- Kinoshita, T., Yadegari, R., Harada, J.J., Goldberg, R.B., and Fischer, R.L.** (1999). Imprinting of the MEDEA polycomb gene in the *Arabidopsis* endosperm. *The Plant cell* **11**, 1945-1952.
- Kouzarides, T.** (2007). Chromatin modifications and their function. *Cell* **128**, 693-705.
- Lachner, M., O'Sullivan, R.J., and Jenuwein, T.** (2003). An epigenetic road map for histone lysine methylation. *Journal of cell science* **116**, 2117-2124.
- Lande-Diner, L., Zhang, J., Ben-Porath, I., Amariglio, N., Keshet, I., Hecht, M., Azuara, V., Fisher, A.G., Rechavi, G., and Cedar, H.** (2007). Role of DNA methylation in stable gene repression. *The Journal of biological chemistry* **282**, 12194-12200.
- Lawrence, R.J., Earley, K., Pontes, O., Silva, M., Chen, Z.J., Neves, N., Viegas, W., and Pikaard, C.S.** (2004). A concerted DNA methylation/histone methylation switch regulates rRNA gene dosage control and nucleolar dominance. *Molecular cell* **13**, 599-609.
- Li, Z.K., Luo, L.J., Mei, H.W., Wang, D.L., Shu, Q.Y., Tabien, R., Zhong, D.B., Ying, C.S., Stansel, J.W., Khush, G.S., and Paterson, A.H.** (2001). Overdominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in rice. I. Biomass and grain yield. *Genetics* **158**, 1737-1753.
- Lippman, Z., Gendrel, A.V., Colot, V., and Martienssen, R.** (2005). Profiling DNA methylation patterns using genomic tiling microarrays. *Nature methods* **2**, 219-224.
- Lippman, Z., Gendrel, A.V., Black, M., Vaughn, M.W., Dedhia, N., McCombie, W.R., Lavine, K., Mittal, V., May, B., Kasschau, K.D., Carrington, J.C., Doerge, R.W., Colot, V., and Martienssen, R.** (2004). Role of transposable elements in heterochromatin and epigenetic control. *Nature* **430**, 471-476.
- Lippman, Z.B., and Zamir, D.** (2007). Heterosis: revisiting the magic. *Trends Genet* **23**, 60-66.
- Lister, R., O'Malley, R.C., Tonti-Filippini, J., Gregory, B.D., Berry, C.C., Millar, A.H., and Ecker, J.R.** (2008). Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. *Cell* **133**, 523-536.
- Liu, B., Brubaker, C.L., Mergeai, G., Cronn, R.C., and Wendel, J.F.** (2001). Polyploid formation in cotton is not accompanied by rapid genomic changes. *Genome* **44**, 321-330.
- Liu, J., He, Y., Amasino, R., and Chen, X.** (2004). siRNAs targeting an intronic transposon in the regulation of natural flowering behavior in *Arabidopsis*. *Genes & development* **18**, 2873-2878.
- Livak, K.J., and Schmittgen, T.D.** (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-Delta Delta C(T)} Method. *Methods (San Diego, Calif)* **25**, 402-408.
- Ma, X.F., and Gustafson, J.P.** (2005). Genome evolution of allopolyploids: a process of cytological and genetic diploidization. *Cytogenetic and genome research* **109**, 236-249.
- Madlung, A., Masuelli, R.W., Watson, B., Reynolds, S.H., Davison, J., and Comai, L.** (2002). Remodeling of DNA methylation and phenotypic and transcriptional changes in synthetic *Arabidopsis* allotetraploids. *Plant physiology* **129**, 733-746.
- 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. *Plant J* **41**, 221-230.
- Mantripragada, K.K., Buckley, P.G., de Stahl, T.D., and Dumanski, J.P.** (2004). Genomic microarrays in the spotlight. *Trends Genet* **20**, 87-94.

- Martienssen, R.A., Doerge, R.W., and Colot, V.** (2005). Epigenomic mapping in *Arabidopsis* using tiling microarrays. *Chromosome Research* **13**, 299-308.
- Martin, C., and Zhang, Y.** (2005). The diverse functions of histone lysine methylation. *Nat Rev Mol Cell Biol* **6**, 838-849.
- Mathieu, O., Reinders, J., Caikovski, M., Smathajitt, C., and Paszkowski, J.** (2007). Transgenerational stability of the *Arabidopsis* epigenome is coordinated by CG methylation. *Cell* **130**, 851-862.
- McClintock, B.** (1951). Chromosome organization and genic expression. Cold Spring Harbor symposia on quantitative biology **16**, 13-47.
- McClintock, B.** (1984). The significance of responses of the genome to challenge. *Science* **226**, 792-801.
- Meyer, R.C., Torjek, O., Becher, M., and Altmann, T.** (2004). Heterosis of biomass production in *Arabidopsis*. Establishment during early development. *Plant physiology* **134**, 1813-1823.
- Meyer, S., Pospisil, H., and Scholten, S.** (2007). Heterosis associated gene expression in maize embryos 6 days after fertilization exhibits additive, dominant and overdominant pattern. *Plant molecular biology* **63**, 381-391.
- Meyer, R.C., Kusterer, B., Lisek, J., Steinfath, M., Becher, M., Scharr, H., Melchinger, A.E., Selbig, J., Schurr, U., Willmitzer, L., and Altmann, T.** (2009). QTL analysis of early stage heterosis for biomass in *Arabidopsis*. TAG.
- Meyers, B.C., Kozik, A., Griego, A., Kuang, H., and Michelmore, R.W.** (2003). Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. *Plant Cell* **15**, 809-834.
- Mitchell-Olds, T., and Schmitt, J.** (2006). Genetic mechanisms and evolutionary significance of natural variation in *Arabidopsis*. *Nature* **441**, 947-952.
- Morgan, H.D., Sutherland, H.G., Martin, D.I., and Whitelaw, E.** (1999). Epigenetic inheritance at the agouti locus in the mouse. *Nature genetics* **23**, 314-318.
- Navashin, M.** (1934) Chromosomal alterations caused by hybridization and their bearing upon certain general genetic problems. *Cytologia* **5**, 169-203
- Ni, Z., Kim, E.D., Ha, M., Lackey, E., Liu, J., Zhang, Y., Sun, Q., and Chen, Z.J.** (2009). Altered circadian rhythms regulate growth vigour in hybrids and allopolyploids. *Nature* **457**, 327-331.
- Noel, L., Moores, T.L., van Der Biezen, E.A., Parniske, M., Daniels, M.J., Parker, J.E., and Jones, J.D.** (1999). Pronounced intraspecific haplotype divergence at the RPP5 complex disease resistance locus of *Arabidopsis*. *The Plant cell* **11**, 2099-2112.
- O'Neill, R.J., O'Neill, M.J., and Graves, J.A.** (1998). Undermethylation associated with retroelement activation and chromosome remodelling in an interspecific mammalian hybrid. *Nature* **393**, 68-72.
- Osborn, T.C., Pires, J.C., Birchler, J.A., Auger, D.L., Chen, Z.J., Lee, H.S., Comai, L., Madlung, A., Doerge, R.W., Colot, V., and Martienssen, R.A.** (2003). Understanding mechanisms of novel gene expression in polyploids. *Trends Genet* **19**, 141-147.
- Pecinka, A., Schubert, V., Meister, A., Kreth, G., Klatt, M., Lysak, M.A., Fuchs, J., and Schubert, I.** (2004). Chromosome territory arrangement and homologous pairing in nuclei of *Arabidopsis thaliana* are predominantly random except for NOR-bearing chromosomes. *Chromosoma* **113**, 258-269.
- Penterman, J., Zilberman, D., Huh, J.H., Ballinger, T., Henikoff, S., and Fischer, R.L.** (2007). DNA demethylation in the *Arabidopsis* genome. *Proc. Natl. Acad. Sci. USA* **104**, 6752-6757.

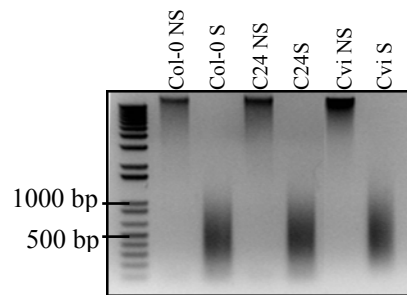
- Pfluger, J., and Wagner, D.** (2007). Histone modifications and dynamic regulation of genome accessibility in plants. *Current opinion in plant biology* **10**, 645-652.
- Pien, S., and Grossniklaus, U.** (2007). Polycomb group and trithorax group proteins in *Arabidopsis*. *Biochimica et biophysica acta* **1769**, 375-382.
- Pikaard, C.S.** (2000). Nucleolar dominance: uniparental gene silencing on a multi-megabase scale in genetic hybrids. *Plant molecular biology* **43**, 163-177.
- Pikaard, C.S., and Lawrence, R.J.** (2002). Uniting the paths to gene silencing. *Nature genetics* **32**, 340-341.
- Pokholok, D.K., Harbison, C.T., Levine, S., Cole, M., Hannett, N.M., Lee, T.I., Bell, G.W., Walker, K., Rolfe, P.A., Herbolsheimer, E., Zeitlinger, J., Lewitter, F., Gifford, D.K., and Young, R.A.** (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* **122**, 517-527.
- Pontes, O., Neves, N., Silva, M., Lewis, M., Madlung, A., Comai, L., Viegas, W., and Pikaard, C.** (2004). Chromosomal locus rearrangements are a rapid response to formation of the allotetraploid *Arabidopsis suecica* genome. *Proc. Natl. Acad. Sci. USA* **101**, 18240-18245.
- Probst, A.V., Dunleavy, E., and Almouzni, G.** (2009). Epigenetic inheritance during the cell cycle. *Nat Rev Mol Cell Biol* **10**, 192-206.
- Rivin, C.J., Cullis, C.A., and Walbot, V.** (1986). Evaluating quantitative variation in the genome of *Zea mays*. *Genetics* **113**, 1009-1019.
- Riddle, N.C., and Richards, E.J.** (2002). The control of natural variation in cytosine methylation in *Arabidopsis*. *Genetics* **162**, 355-363.
- Rocha, W., and Verreault, A.** (2008). Clothing up DNA for all seasons: Histone chaperones and nucleosome assembly pathways. *FEBS letters* **582**, 1938-1949.
- Rufas, J.S., Iturra, P., Desouza, W., and Esponda, P.** (1982). Simple silver staining procedures for the location of nucleolus and nucleolar organizer under Light and electron-microscopy. *Archives of Biology* **93**, 267-274.
- Ruiz-Garcia, L., Cervera, M.T., and Martinez-Zapater, J.M.** (2005). DNA methylation increases throughout *Arabidopsis* development. *Planta* **222**, 301-306.
- Salmon, A., Ainouche, M.L., and Wendel, J.F.** (2005). Genetic and epigenetic consequences of recent hybridization and polyploidy in *Spartina* (Poaceae). *Molecular ecology* **14**, 1163-1175.
- Sambrook, J., Russell, D. W.** (2001) *Molecular cloning: a laboratory manual, the third edition*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York,
- Saze, H., Sasaki, T., and Kakutani, T.** (2008). Negative regulation of DNA methylation in plants. *Epigenetics* **3**, 122-124.
- Scheer, U., and Weisenberger, D.** (1994). The nucleolus. *Current opinion in cell biology* **6**, 354-359.
- Schmid, K.J., Sorensen, T.R., Stracke, R., Torjek, O., Altmann, T., Mitchell-Olds, T., and Weisshaar, B.** (2003). Large-scale identification and analysis of genome-wide single-nucleotide polymorphisms for mapping in *Arabidopsis thaliana*. *Genome research* **13**, 1250-1257.
- Schubert, D., Clarenz, O., and Goodrich, J.** (2005). Epigenetic control of plant development by Polycomb-group proteins. *Current opinion in plant biology* **8**, 553-561.
- Schwenk, K., Brede, N., and Streit, B.** (2008). Introduction. extent, processes and evolutionary impact of interspecific hybridization in animals. *Philosophical Transactions of the Royal Society B-Biological Sciences* **363**, 2805-2811.

- Sebat, J., Lakshmi, B., Troge, J., Alexander, J., Young, J., Lundin, P., Maner, S., Massa, H., Walker, M., Chi, M., Navin, N., Lucito, R., Healy, J., Hicks, J., Ye, K., Reiner, A., Gilliam, T.C., Trask, B., Patterson, N., Zetterberg, A., and Wigler, M. (2004).** Large-scale copy number polymorphism in the human genome. *Science* **305**, 525-528.
- Seifert, M., Banaei, A., Keilwagen, J., Mette, M.F., Houben, A., Roudier, F., Colot, V., Grosse, I. and Strickert, M. (2009)** Array-based genome comparison of *Arabidopsis* ecotypes using hidden Markov models, Proceedings of Biosignals , Second International Conference on Bio-inspired Systems and Signal Processing, ISBN: 978-989-8111-65-4, p. 3-11.
- Shaw, P.J., and Jordan, E.G. (1995).** The nucleolus. *Annual review of cell and developmental biology* **11**, 93-121.
- Shilatifard, A. (2008).** Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation. *Current opinion in cell biology* **20**, 341-348.
- Shiu, S.H., and Bleecker, A.B. (2001).** Plant receptor-like kinase gene family: diversity, function, and signaling. *Sci STKE* **2001**, RE22.
- Shull, G.H. (1948).** What Is "Heterosis"? *Genetics* **33**, 439-446.
- Song, K., Lu, P., Tang, K., and Osborn, T.C. (1995).** Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. *Proc. Natl. Acad. Sci. USA* **92**, 7719-7723.
- Soppe, W.J., Jasencakova, Z., Houben, A., Kakutani, T., Meister, A., Huang, M.S., Jacobsen, S.E., Schubert, I., and Fransz, P.F. (2002).** DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in *Arabidopsis*. *The EMBO journal* **21**, 6549-6559.
- Stefanovsky, V.Y., Pelletier, G., Hannan, R., Gagnon-Kugler, T., Rothblum, L.I., and Moss, T. (2001).** An immediate response of ribosomal transcription to growth factor stimulation in mammals is mediated by ERK phosphorylation of UBF. *Molecular cell* **8**, 1063-1073.
- Strahl, B.D., and Allis, C.D. (2000).** The language of covalent histone modifications. *Nature* **403**, 41-45.
- Sugimoto-Shirasu, K., and Roberts, K. (2003).** "Big it up": endoreduplication and cell-size control in plants. *Current opinion in plant biology* **6**, 544-553.
- Swanson-Wagner, R.A., Jia, Y., Decook, R., Borsuk, L.A., Nettleton, D., and Schnable, P.S. (2006).** All possible modes of gene action are observed in a global comparison of gene expression in a maize F1 hybrid and its inbred parents. *Proc. Natl. Acad. Sci. USA* **103**, 6805-6810.
- Tariq, M., and Paszkowski, J. (2004).** DNA and histone methylation in plants. *Trends Genet* **20**, 244-251.
- The *Arabidopsis* Genome Initiative. (2000).** Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796-815.
- Tsaftaris, S.A. (1995).** Molecular Aspects of Heterosis in Plants. *Physiologia Plantarum* **94**, 362-370.
- Tsaftaris, AS., Kafka, M., Polidoros, A., Tani, W. (1997)** Epigenetic changes in maize DNA and heterosis. Abstracts of the international Symposium on "The genetics and exploitation of heterosis in crops", Mexico City, pp 112-113.
- Turck, F., Roudier, F., Farrona, S., Martin-Magniette, M.L., Guillaume, E., Buisine, N., Gagnot, S., Martienssen, R.A., Coupland, G., and Colot, V. (2007).** *Arabidopsis*

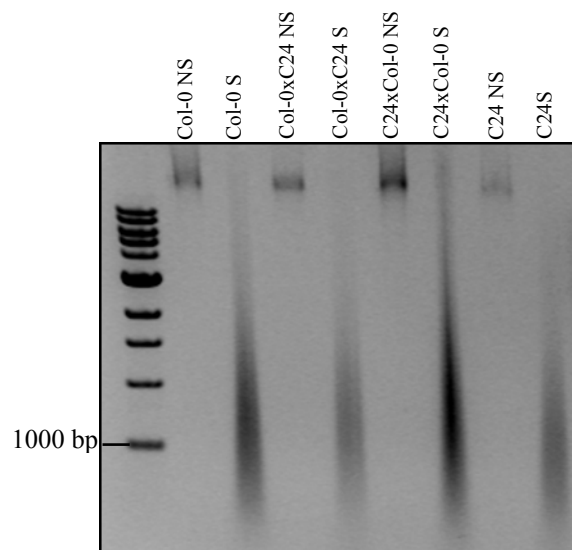
- TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27. *PLoS genetics* **3**, e86.
- Ungerer, M.C., Strakosh, S.C., and Zhen, Y.** (2006). Genome expansion in three hybrid sunflower species is associated with retrotransposon proliferation. *Curr Biol* **16**, R872-873.
- Uzarowska, A., Keller, B., Piepho, H., Schwarz, G., Ingvaridsen, C., Wenzel, G., and Lubberstedt, T.** (2007). Comparative expression profiling in meristems of inbred-hybrid triplets of maize based on morphological investigations of heterosis for plant height. *Plant molecular biology* **63**, 21-34.
- Van Speybroeck, L.** (2002). From epigenesis to epigenetics: the case of C. H. Waddington. *ANN. NY. ACAD. SCI.* **981**, 61-81.
- Vaughn, M.W., Tanurd Ic, M., Lippman, Z., Jiang, H., Carrasquillo, R., Rabinowicz, P.D., Dedhia, N., McCombie, W.R., Agier, N., Bulski, A., Colot, V., Doerge, R.W., and Martienssen, R.A.** (2007). Epigenetic Natural Variation in *Arabidopsis thaliana*. *PLoS biology* **5**, e174.
- Vissers, L.E., Veltman, J.A., van Kessel, A.G., and Brunner, H.G.** (2005). Identification of disease genes by whole genome CGH arrays. *Human molecular genetics* **14 Spec No. 2**, R215-223.
- Wang, J., Tian, L., Madlung, A., Lee, H.S., Chen, M., Lee, J.J., Watson, B., Kagochi, T., Comai, L., and Chen, Z.J.** (2004). Stochastic and epigenetic changes of gene expression in *Arabidopsis* polyploids. *Genetics* **167**, 1961-1973.
- Wang, J., Tian, L., Lee, H.S., Wei, N.E., Jiang, H., Watson, B., Madlung, A., Osborn, T.C., Doerge, R.W., Comai, L., and Chen, Z.J.** (2006). Genomewide nonadditive gene regulation in *Arabidopsis* allotetraploids. *Genetics* **172**, 507-517.
- Warner, J.R.** (1999). The economics of ribosome biosynthesis in yeast. *Trends in biochemical sciences* **24**, 437-440.
- Workman, J.L., and Abmayr, S.M.** (2004). Histone H3 variants and modifications on transcribed genes. *Proc. Natl. Acad. Sci. USA* **101**, 1429-1430.
- Wu, C., and Morris, J.R.** (2001). Genes, genetics, and epigenetics: a correspondence. *Science* **293**, 1103-1105.
- Xiong, L.Z., Xu, C.G., Saghai Maroof, M.A., and Zhang, Q.** (1999). Patterns of cytosine methylation in an elite rice hybrid and its parental lines, detected by a methylation-sensitive amplification polymorphism technique. *Mol Gen Genet* **261**, 439-446.
- Xu, Y., Zhong, L., Wu, X., Fang, X., and Wang, J.** (2009). Rapid alterations of gene expression and cytosine methylation in newly synthesized *Brassica napus* allopolyploids. *Planta* **229**, 471-483.
- Zhang, X., Shiu, S., Cal, A., and Borevitz, J.O.** (2008). Global analysis of genetic, epigenetic and transcriptional polymorphisms in *Arabidopsis thaliana* using whole genome tiling arrays. *PLoS genetics* **4**, e1000032.
- Zhang, X., Clarenz, O., Cokus, S., Bernatavichute, Y.V., Pellegrini, M., Goodrich, J., and Jacobsen, S.E.** (2007). Whole-genome analysis of histone H3 lysine 27 trimethylation in *Arabidopsis*. *PLoS biology* **5**, e129.
- Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, S.W., Chen, H., Henderson, I.R., Shinn, P., Pellegrini, M., Jacobsen, S.E., and Ecker, J.R.** (2006). Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. *Cell* **126**, 1189-1201.

- Zhang, Y., and Reinberg, D.** (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes & development* **15**, 2343-2360.
- Zhao, Y., Yu, S., Xing, C., Fan, S., and Song, M.** (2008). Analysis of DNA methylation in cotton hybrids and their parents. *Molecular Biology* **42**, 169-178.
- Zilberman, D., and Henikoff, S.** (2005). Epigenetic inheritance in *Arabidopsis*: selective silence. *Current opinion in genetics & development* **15**, 557-562.
- Zilberman, D., Gehring, M., Tran, R.K., Ballinger, T., and Henikoff, S.** (2007). Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nature genetics* **39**, 61-69.

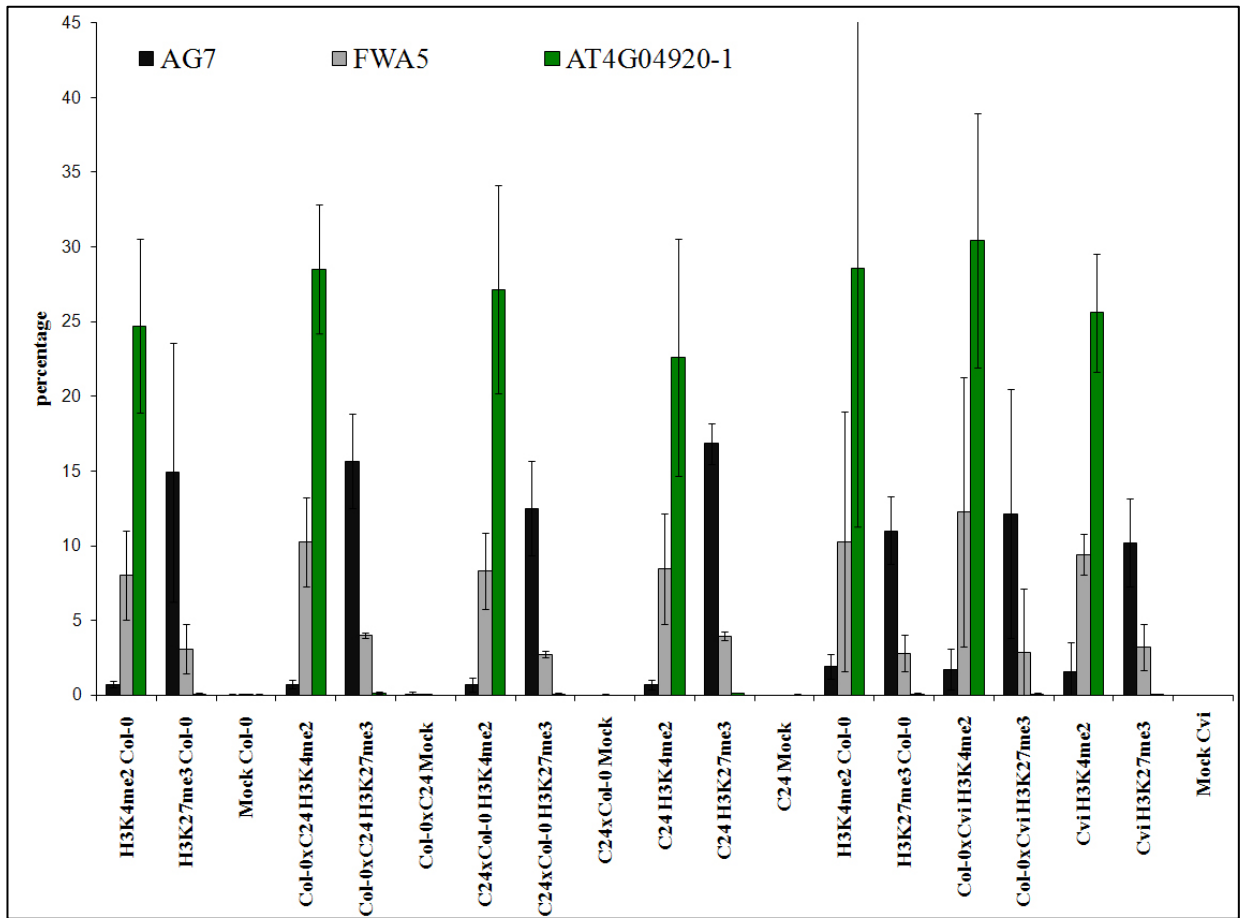
Appendix



Appendix Figure 1 Fragmentation of genomic DNA by sonication. Genomic DNA of Col-0, C24 and Cvi was sheared to an average size of 800 bp by sonication. S: sonicated DNA, NS: non-sonicated DNA. Inverse images of fluorescence signals after agarose gel electrophoresis and ethidium bromide staining are shown.

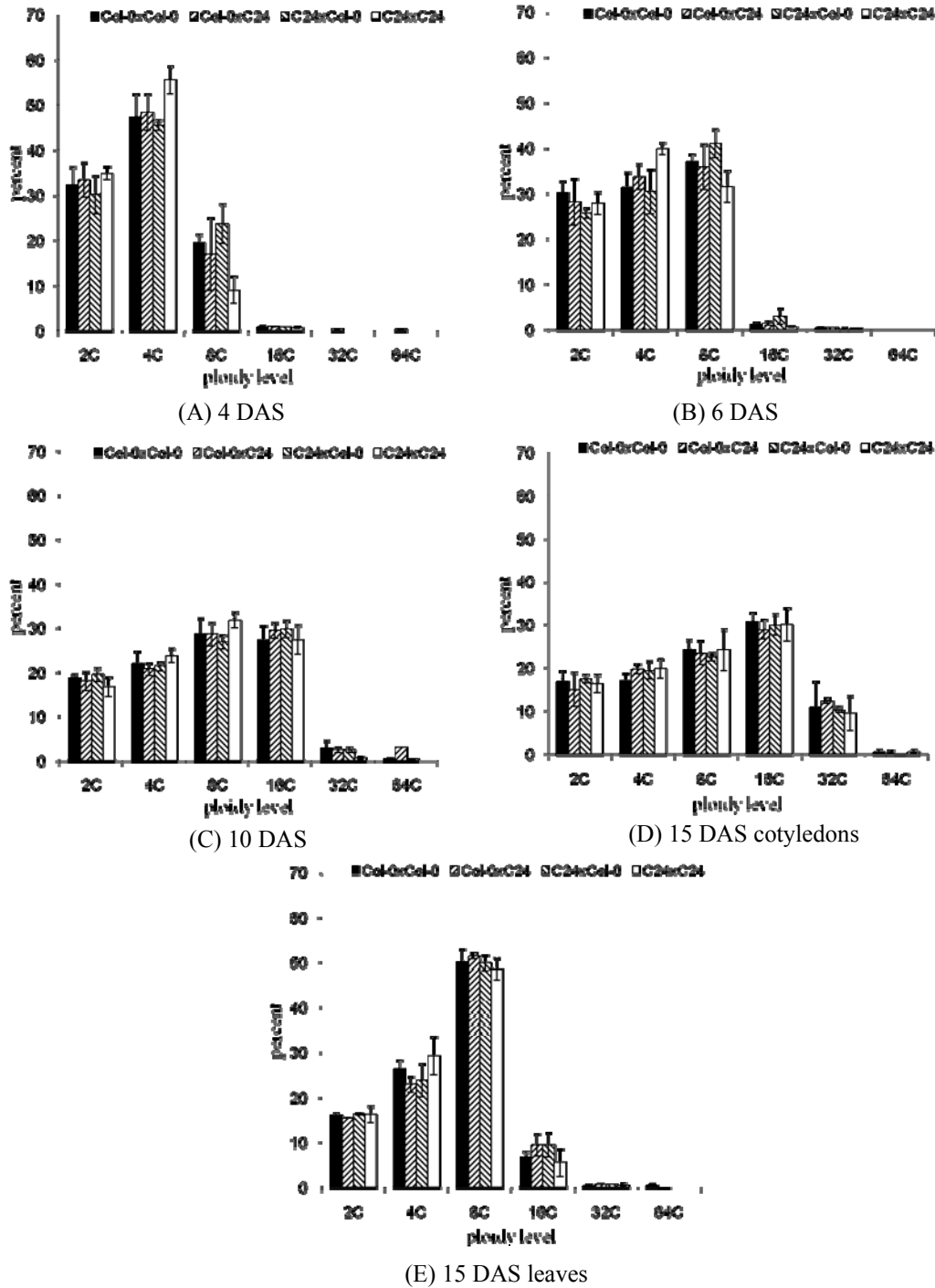


Appendix Figure 2 Fragmentation of chromatin by sonication. Chromatin was extracted from seedlings of genotypes Col-0, Col-0xC24, C24xCol-0, C24, Cvi, Col-0xCvi and Cvi x Col-0. Chromatins preparations were sheared to an average size of 800 by sonication. Representative sonicated and non sonicated chromatin extracted from first biological replicate of Col-0, Col-0xC24, C24xCol-0 and C24 is shown. S: DNA from sonicated chromatin, NS: DNA from non-sonicated chromatin. Inverse images of fluorescence signals after agarose gel electrophoresis and ethidium bromide staining are shown.



Appendix Figure 3 Efficiency of IP experiments for H3K4me2 and H3K27me3 on chromatin from Col-0, Col-0xC24, C24xCol-0, C24, Col-0xCvi and Cvi.

ChIP quantitative PCR was performed with primer pairs for reference regions known to be enriched in H3K4me2 (AG7, At4g49920-1) or H3K27me3 (FWA5) and obtained results were calibrated by the results for the input controls for the respective chromatin preparations. Similar enrichment of these reference sequences in IP samples for a specific modification mark was taken as indicating equal IP efficiencies. Analysis was performed for two biological replicates in two technical replicates, each. Error bars indicate the absolute deviation among measurements.



Appendix Figure 4 Ploidy levels at different developmental stages in inbred parental lines and their interspecific hybrid offsprings.

Nuclei isolated at different developmental stages (4, 6, 10 and 15 DAS) of Col-0 (black bars) and C24 (white bars) and their reciprocal hybrid offspring (hatched bars). A) Nuclei from cotyledons of seedling at 4 DAS. B) Nuclei from cotyledons of seedling at 6 DAS. C) Nuclei from cotyledons of seedling at 10 DAS. D) Nuclei from cotyledons of plantlets at 15 DAS. E) Nuclei from rosette leaves of plantlets at 15 DA. Error bars indicating standard deviation among three biological replicates.

Appendix Table 1 Number of nuclei of defined ploidy for each genotype at different developmental stages.

genotype		Col-0						Col-0xC24					
developmental stage	ploidy level	2C	4C	8C	16C	32C	64C	2C	4C	8C	16C	32C	64C
4 DAS	Cotyledons	9682	14208	5828	276	0	0	10021	14506	5057	265	108	108
6 DAS	Cotyledons	9026	9376	11084	353	154	0	8431	10093	10756	435.9	120	0
10 DAS	Cotyledons	5643	6577	8606	8155	870	141	5391	6187	8562	8808	753	864
15 DAS	Cotyledons	3371	3391	4831	6126	2162	114	3004	3943	4695	5785	2483	101
15 DAS	leaves	3209	5271	10025	1339	83	100	3073	4577	10294	1883	138	60
genotype		C24xCol-0						C24					
developmental stage	ploidy level	2C	4C	8C	16C	32C	64C	2C	4C	8C	16C	32C	64C
4 DAS	Cotyledons	9039	13643	7084	230	0	0	10449	16654	2692	201	0	0
6 DAS	Cotyledons	7705	9132.3	12293	877	114	0	8340	11954	9446	177	105	0
10 DAS	Cotyledons	5831	6391	8055	8929	718	103	4987	7127	9516	8168	195	0
15 DAS	Cotyledons	3485	3894	4507	5991	2071	68	3253	3953	4832	5994	1881	121
15 DAS	leaves	3249	4757	9983	1876	129	0	3228	5838	9715	1098	116	0

Appendix Table 2 Number of nuclei for which nucleolus areas were determined.

genotype	ploidy level			
	replicate 1		replicate 2	
	2C	4C	2C	4C
Col-0	63	119	162	149
Col-0xC24	125	147	118	151
C24xCol-0	115	296	149	101
C24	131	117	202	100

Appendix Table 3 Candidate tails covering sequences with significantly higher or lower level of H3K4me2 in hybrid offspring compared to their inbred parents based on chromosome 4 tiling microarray.

Name	Function	level of H3K4me2 in offspring compared with both parents
AT4G00970	protein kinase	lower
AT4G02480	ATPase family	higher
AT4G04480	unknown protein	lower
AT4G08160	hydrolase	lower
AT4G26130	unknown protein	higher
AT4G08830	RNA-directed DNA polymerase	higher
intergenic region between AT4G13260 and AT4G13270	intergenic region	higher
AT4G19310	cysteine-type peptidase	higher
AT4G21380	kinase	higher
AT4G34360	methyltransferase	lower
AT4G35720 & AT4G35725	unknown proteins	higher
intergenic region between AT4G39760 and AT4G39770	intergenic region	higher

Publications related to the submitted thesis

Seifert, M., Banaei, A., Keilwagen, J., Mette, M.F., Houben, A., Roudier, F., Colot, V., Grosse, I. and Strickert, M. Array-based genome comparison of *Arabidopsis* ecotypes using hidden Markov models, Proceedings of Biosignals 2009, Second International Conference on Bio-inspired Systems and Signal Processing, ISBN: 978-989-8111-65-4, p. 3-11.

Banaei Moghaddam, A.M., Fuchs, J., Czauderna, T., Houben, A., and Mette, M.F. (2010). Intraspecific hybrids of *Arabidopsis thaliana* revealed no gross alterations in endopolyploidy, DNA methylation, histone modifications and transcript levels. *Theoretical and Applied Genetics* **120**, 215-226.

Herr Ali mohammad Banaei Moghaddam
Corrensstraße 3
06466 Gatersleben

Eidesstattliche Erklärung

Hiermit erkläre ich, dass diese Arbeit von mir bisher weder der Mathematisch-Naturwissenschaftlich-Technischen Fakultät der Martin-Luther-Universität Halle-Wittenberg noch einer anderen wissenschaftlichen Einrichtung zum Zweck der Promotion eingereicht wurde.

Ich erkläre ferner, dass ich diese Arbeit selbständig und nur unter Zuhilfenahme der angegebenen Hilfsmittel und Literatur angefertigt habe.

Gatersleben, den

Ali Mohammad Banaei Moghaddam

CURRICULUM VITAE

Name Ali Mohammad Banaei Moghaddam
Birth 24 / 08 / 1979 in Yazd, Iran
Nationality Iranian
Marital Status Married
Contact information Research Group Chromosome Structure and Function
Leibniz Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstraße 3
06466 Gatersleben, Germany
phone: 0049 (0) 39482 5477
FAX: 0049 (0) 39482 5137
e-mail: banaei@ipk-gatersleben.de

ACADEMIC QUALIFICATIONS

2001-2004 **M.Sc. in Biochemistry**, Tehran University, Iran.
1997-2001 **B.Sc. in Cellular and Molecular Biology**, Tehran University, Iran
1993–1997 **Diploma** of theoretical secondary education studies, Yazd, Iran

EMPLOYMENT HISTORY / RESEARCH EXPERIENCE

Since June 2005 PhD. Student, Chromosome Structure and Function group, Department of Cytogenetic, Institute of Plant Genetics and Crop Plant Research (IPK), Germany, Supervisor Dr. A. Houben

2004 –2005 Junior Scientist, Laboratory of Physiology and Proteomics, ABRII, Iran. Research topic: “Study of B-chromosome effects on rye proteome profile”. Supervisor Dr. Gh. Hosseini Salekdeh.

2002 – 2004 M. Sc. research work in the Lab of cytogenetics, Institute of Biochemistry and Biophysics (IBB), , Iran. Research topic: “Effect of B-chromosomes on proteome pattern of *Acanthophyllum Laxiusculum*”. Supervisor Dr. Mahmud Ghaffari.