Akademie věd
České republiky

Teze disertace
k získání vědeckého titulu "doktor věd"
ve skupině věd biologicko-ekologických

INSIGHTS INTO GENOME EVOLUTION OF THE MUSTARD FAMILY (BRASSICACEAE)

Martin A. Lysák

Středoevropský technologický institut (CEITEC), Masarykova univerzita

Brno, 2016

Teze disertace
k získání vědeckého titulu "doktor věd"
ve skupině věd biologicko-ekologických

Insights into genome evolution of the mustard family (Brassicaceae)

název disertace

Komise pro obhajoby doktorských disertací v oboru botanika a fyziologie rostlin

Jméno uchazeče: doc. Mgr. Martin A. Lysák, Ph.D.

Pracoviště uchazeče: Středoevropský technologický institut (CEITEC), Masarykova univerzita, Kamenice 753/5, 625 00 Brno

V Brně, 24.1.2016

Summary

This dissertation thesis summarizes fifteen years of comparative cytogenomics and phylogenomics of the mustard family (crucifers, Brassicaceae) since the sequencing of the thale cress (*Arabidopsis thaliana*) genome in 2000. The thesis provides a brief overview of our current understanding of trends and principles driving crucifer genome evolution with a particular focus on scientific achivements of the applicant. Besides presenting insights into genome evolution of the mustard family, the dissertation is also a historical guide compiling major technological and conceptual advances in comparative genomics of crucifers during the last fifteen years.

The mustard family includes a wide range of important vegetable, condiment, forage and oil seed crops such as, for example, broccoli, cabbage, cauliflower or rapeseed, as well as ornamental plants, and several model organisms. The Brassicaceae comprises c. 320 genera and 3,666 species grouped into 49 tribes. A recently published phylogenetic study resolved at least six major clades (A to F) within the Brassicaceae. Comparative linkage mapping, comparative cytogenetics (particularly comparative chromosome painting), genome sequencing and molecular phylogenetic studies were instrumental for our understanding of karyotype and genome evolution in crucifers. These studies resulted in a theoretical concept of the Ancestral Crucifer Karyotype (ACK) with eight chromosomes (n = 8) and 22 conserved genomic blocks. The ACK genome has been proven to be important for radiation of the larger Clade A (e.g. genus *Arabidopsis*). During the diversification of the mustard family, ACK has evolved following four basic patterns: (i) genome stasis without major chromosomal reshuffling (n = 8), (ii) prevailing genome stasis with a few structural rearrangements (n = 8), (iii) substantial genome reshuffling but ancestral chromosome number stasis (n = 8), and (iv) genome reshuffling and descending dysploidy (chromosome number decrease) from n = 8 (n = 7, 6, and 5). From ancestral genomes derived from the ACK via descending dysploidy, the formation of the proto-Calepineae Karyotype (PCK, n = 7) preceded diversification of the species-rich Clade B (e.g. genus *Brassica*). Genomes of crucifer plants were also shaped by several tribe- and genus-specific whole-genome duplications (polyploidizations). These whole-genome duplications (tetraploidy) and triplications (hexaploidy) are of different age and usually were followed by genome-wide diploidization. Post-polyploidy rediploidization is associated with major repatterning of genomic blocks and descending dysploidy (up to n = 4 chromosomes). Cyclic rounds of polyploidization and rediploidization seem to be important drivers of genetic, physiological and ecological diversification, and ultimately of speciation. By contrast, some important clade diversifications in the mustard family were not associated Detailed analyses of crucifer genomes helped to (re)discover major mechanisms underlying chromosomal and karyotype evolution in the Brassicaceae and other angiosperm families. These mechanisms include inversions, reciprocal chromosome translocations, centromere invactivation/loss, centromere repositioning and centric fissions.

Abbreviations used

ACK	ancestral crucifer karyotype
CP	comparative painting
CCP	comparative chromosome painting
CR	centromere repositioning
BAC	bacterial artificial chromosome
FISH	fluorescence *in situ* hybridization
MCR	major chromosomal rearrangement
mFISH	multicolour fluorescence *in situ* hybridization
mya	million years ago
NGS	next-generation sequencing
PAC	bacteriophage P1-derived artificial chromosome
PCK	proto-Calepineae karyotype
SNP	single nucleotide polymorphism
SSR	simple sequence repeats
TAC	transformation-competent artificial chromosome
TCT	terminal chromosome translocation
tPCK	translocation proto-Calepineae karyotype
WGD	whole-genome duplication
WGT	whole-genome triplication
YAC	yeast artificial chromosome

Contents

1. Introduction

After the genome sequence of *Arabidopsis thaliana* (Arabidopsis or thale cress) has been released sixteen years ago (AGI 2000), plant sciences changed forever. The development of Arabidopsis genomic resources and tools was crucial in establishing Arabidopsis as the prime cytogenetic model, and soon after Arabidopsis resources became instrumental in developing the field of comparative cytogenomics in the mustard family (crucifers or Brassicaceae). More than fifteen years later, this field is at the forefront of plant comparative genomics with some of the most important findings being inferred from cytogenetic and genome sequencing studies in crucifers.

But why to study plants mostly with no significant agricultural value? First, it has to be acknowledged that our understanding of principles governing the structure and organisation of plant genomes is still fragmentary and therefore more observations are needed to establish the overarching patterns and models. Second, observations from the model species and families can be used when analyzing and improving plants that we may want to eat. Recent technological and conceptual advances in comparative genomics, phylogenomics and paleogenomics, illustrated here by case studies from the mustard family, promise to provide new unique findings for model as well as non-model plant species and families.

In this dissertation, I attempted to briefly summarize exciting fifteen years of my professional life spent by studying chromosomes and genomes of crucifer plants. My thesis aims to provide the reader with an overview of our current understanding of trends and mechanisms driving crucifer genome evolution, with a particular focus on my scientific contributions to the field. Due to space constraints, I had to refrain from providing more exhaustive introduction and discussion of all the interesting aspects covered.

2. Crucifers (Brassicaceae)

The mustard family (Brassicaceae) includes a wide range of important vegetable, condiment, forage and oil seed crops such as broccoli, cabbage, cauliflower, kohl rabi (all cultivars of *Brassica oleracea*), radish (*Raphanus sativus*), mustard (*Sinapis alba*), rapeseed (*B. napus*), as well as ornamentals (e.g. *Arabis*, *Aubrieta*, *Erysimum*, *Iberis*, *Lunaria*, *Matthiola* and others), and several model organisms (e.g. *Arabidopsis thaliana*, *A. suecica*, *Arabis alpina*, *Cardamine hirsuta*). Although the Brassicaceae shows a worldwide distribution on all continents except Antarctica, the species diversity is not distributed equally and the most important diversification centres are found in the Irano-Turanian and the Mediterranean region, respectively (Hedge 1976; Appel and Al-Shehbaz 2003; Franzke et al. 2008). This distribution pattern and several recent studies (e.g. Karl and Koch 2013; Özüdoğru et al. 2015) are strongly suggesting that the family originated and diversified in the eastern Mediterranean (Asia Minor or present-day Turkey) and adjacent regions.

The Brassicaceae is a large plant family with approximately 320 genera and 3,666 species (Al-Shehbaz 2012). Phylogenetically, crucifers are sister to the family Cleomaceae, and both families belong to the order Brassicales (Hall et al. 2002, 2004). The high diversity of the Brassicaceae called upon to introduce an intrafamiliar tribal classification system. First tribal systems, mostly based on morphological characters, were artificial and did not reflect the true phylogenetic relationships (e.g. Koch et al. 2003; Al-Shehbaz et al. 2006). The

present-day system of 49 *bona fide* monophyletic tribes (Al-Shehbaz 2012) is based on numerous phylogenetic studies published throughout the last decade. Among the 49 tribes, the tribe Aethionemeae was repeatedly identified as sister to all the remaining tribes of the so-called crown group (the core Brassicaceae taxa). The crown-group tribes were initially clustered into three major phylogenetic lineages (Lineage I-III, Fig. 1A), but leaving several tribes and genera without assignment to a lineage or tribe, respectively (Al-Shehbaz et al. 2006, Bailey et al. 2006, Beilstein et al. 2006, 2008; Franzke et al. 2011). Very recently, a new phylogenetic study based on more than 100 orthologous nuclear markers retrieved from sequenced transcriptomes of 55 species resulted in definition of five major clades (A to E, Fig. 1B) within the crown group (Huang et al. 2015). In the feature, this multi-gene extended for a larger species sample should pave the way for a robust and reliable family tree.

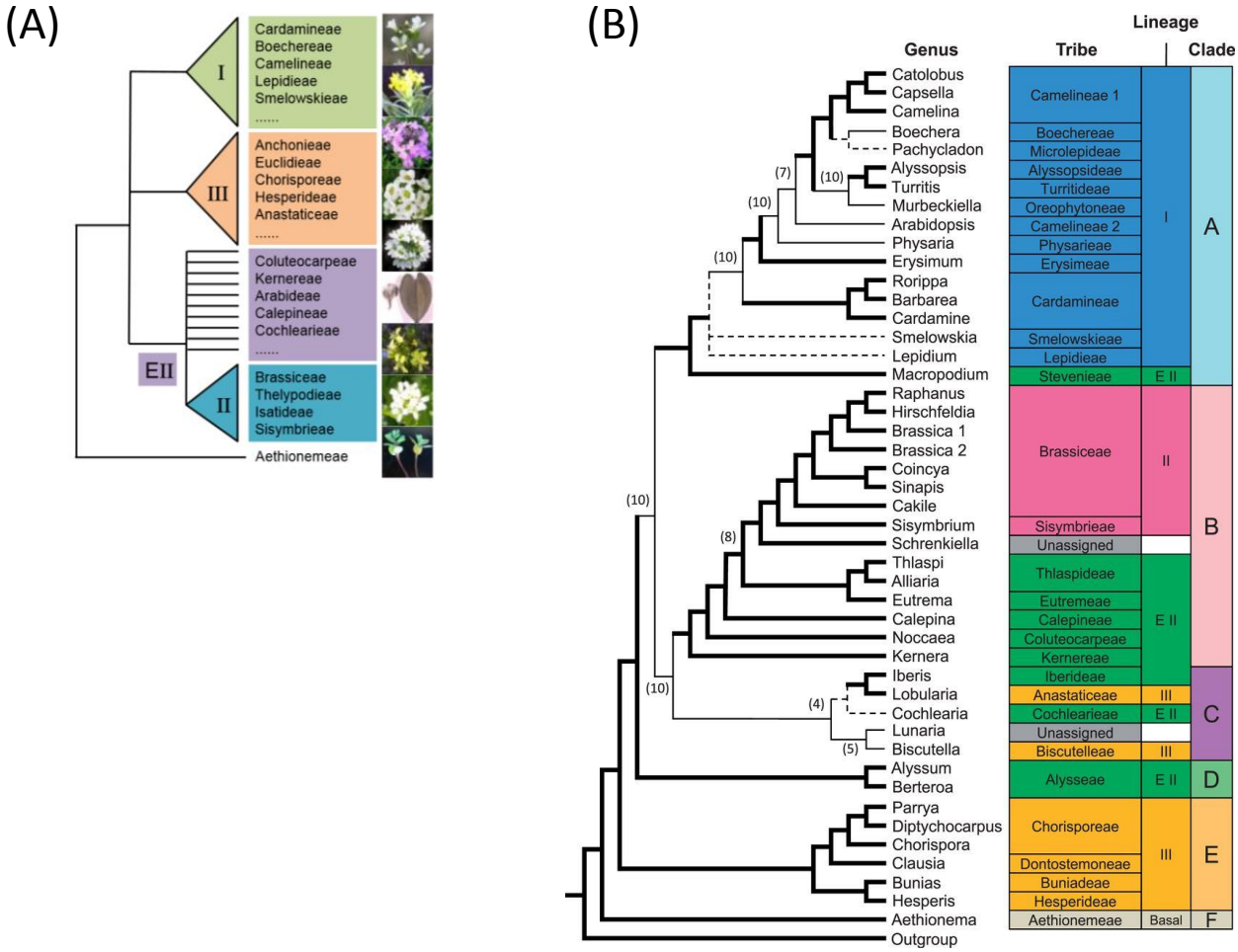

Figure 1. Phylogenetic relationships in the mustard family. (A) A schematic family-wide phylogeny proposed by Franzke et al. (2011). Tribes were clustered into three major lineages – Lineage I, II and III (EII refers to expanded Lineage II). (B) A schematic family-wide phylogeny published by Huang et al. (2015). The Clade F (Aethionemeae) is sister to the crown group comprising five clades (Clade A, B, C, D and E).

3. Comparative Cytogenomics Approaches

3.1 Genetic Mapping

Comparative genetic maps are an invaluable source of information on the extent of inter-specific genome and chromosomal colinearity as well as on the incidence of major chromosomal rearrangements (MCRs) differentiating two or more genomes compared. In crucifer genomics, comparative genetic mapping studies allowed us to i) infer inter-genomic collinearity, ii) initiate the very first large-scale comparative cytogenetic analyses (see 3.2), and iii) navigate whole-genome sequence assemblies (see 3.3). The Arabidopsis genome sequence (AGI 2000) along with the wealth of genetic markers have been crucial for placing this genome as a central reference in comparative genomics of crucifers. With more crucifer genomes being sequenced and due to deeper understanding of genome evolution in the Brassicaceae, the role of the evolutionary derived Arabidopsis genome has become less prominent.

Early mapping studies were focused on revealing the origin and genome relationships of three „diploid" *Brassica* species (*B. nigra, B. oleracea* and *B. rapa*), and between Arabidopsis and *B. nigra* (Lagercrantz and Lydiate 1996; Lagercrantz 1998). These comparisons indicated that *Brassica* genomes comprise three copies of an ancestral genome(s) as a consequence of a presumed ancient hexaploidy event. The ancient whole-genome triplication (WGT) was convincingly demonstrated by Parking et al. (2005). The mapping of a large RFLP marker set in the allotetraploid *B. napus* and Arabidopsis gained strongly evidence that "diploid" *Brassica* species have undergone an ancient WGT (see also 5.2). The identification of large orthologous genomic regions shared between *Brassica* and Arabidopsis was instrumental in proposing the system of 24 conserved genomic blocks building crucifer genomes (Schranz et al. 2006).

Equally influential and important as the study of Parkin et al., were comparative genetic maps between Arabidopsis and two n = 8 Camelineae species - *Arabidopsis lyrata* subsp. *petraea* (Kuittinen et al. 2004) and *Capsella rubella* (Boivin et al. 2004). Later, both maps were compared with the results of genetic mapping between Arabidopsis and *A. lyrata* subsp. *lyrata* (Yogeeswaran et al. 2005). In contrast to the Arabidopsis-*Brassica* maps (see above), eight linkage groups of *A. lyrata* and *C. rubella* showed a strikingly high extent of colinearity with the five Arabidopsis chromosomes. The three-way (*A. thaliana* - *A. lyrata* - *C. rubella*) comparison concluded that genomes of *A. lyrata* and *C. rubella* are in fact identical and differ from the Arabidopsis karyotype by three translocation-mediated "chromosome fusions". The overall similarity of the *Capsella* and *A. lyrata* karyotypes indicated that they may resemble an ancestral n = 8 karyotype shared by Camelineae species, later named the Ancestral Crucifer Karyotype (**Fig. 2**; Lysak et al. 2006a; Schranz et al. 2006). These seminal studies accelerated construction of several other comparative genetic maps (e.g. Schranz et al. 2007; Burell et al. 2011) and served as a basis for the first comparative chromosome painting studies in the Brassicaceae (see 3.2).

Genetic maps played a pivotal role in whole-genome sequencing projects. Suquence scaffolds were anchored using genetic maps in the genome sequencing of *A. lyrata* (Hu et al. 2011), *C. rubella* (Slotte et al. 2013) and *B. rapa* (Wang et al. 2011). Next-generation sequencing (NGS) techniques allow to generate millions of genetic markers, such as simple sequence repeats (SSR) and single nucleotide polymorphism (SNP) markers. The sequenced genome of *B. oleracea* has been assembled using a genetic map based on SSR and SNP

markers (Wang et al. 2012; Liu et al. 2014) and a similar strategy was followed to assemble the neoallotetraploid genome of *B. napus* (Delourme et al. 2013; Chalhoub et al. 2014). In the last sequenced crucifer species, the hexaploid false flax (*Camelina sativa*), the sequences of its 20 chromosomes were assembled using a genetic map with more than 3,500 polymorphic loci (SNPs, SSRs and insertion/deletion polymorphisms (Kagale et al. 2014).

3.2 Comparative Chromosome Painting

Sequencing of the Arabidopsis genome has been accomplished through the use of genomic DNA libraries of high-capacity vectors including cosmids, YACs, BACs, bacteriophage P1-derived artificial chromosomes (PACs, P1 clones), and transformation-competent artificial chromosomes (TACs) (AGI 2000). The sequencing project led to the assembly of chromosome-arm-specific contigs of BAC clones from two Arabidopsis BAC libraries (IGF library, Mozo et al. 1998; TAMU library, Choi et al. 1995). The available minimum tiling paths of BACs covering the ten chromosome arms raised the question whether these clones can be used as specific DNA probes to visualize or „paint" the non-repetitive chromosome regions or even entire chromosomes of Arabidopsis. Several studies (e.g. Lysak et al. 2001, 2003; Fransz et al. 2002; Pecinka et al. 2004) demonstrated the feasibility of this approach. The terms BAC FISH and chromosome painting (CP) were coined to specify the use of BAC clones and BAC contigs (i.e. a continuous series of BACs) as probes for multicolour fluorescence in situ hybridization (mFISH). Whereas BAC FISH refers to the application of a single or few BAC clones, CP in the context of Arabidopsis and crucifer cytogenetics, describes the use of BAC contigs comprising several up to several hundred clones covering large chromosomal regions or whole chromosomes. When chromosome-specific BAC contigs (or alternative painting probes) are applied as probes to chromosomes of other species than the species of origin, we speak about comparative chromosome painting (CCP).

The advent of BAC-based chromosome painting, sparkerd off a revolution in Arabidopsis cytogenetics and comparative cytogenomics of crucifers. Large-scale CP with BAC contigs covering both arms of Arabidopsis chromosome At4 enabled visualization of the chromosome during mitosis, meiosis as well as within interphase nuclei (Lysak et al. 2001). This was historically the first report on painting of an entire chromosome in an euploid plant. Painting of other Arabidopsis chromosomes (Lysak et al. 2003) facilitated the simultaneous visualization of all chromosomes by multicolour CP using BAC contigs covering all ten chromosome arms (Pecinka et al. 2004). The technique has served particularly as a tool to study chromosomal rearrangements and spatial organization of interphase chromosomes (e.g. Fransz et al. 2002; Pecinka et al. 2004; Berr and Schubert 2007; Schubert et al. 2012, 2014; Tan et al. 2015).

In the context of crucifer comparative genomics, CCP refers to the identification of large homeologous chromosomal regions and/or whole chromosomes shared by two or more crucifer species, based on the use of Arabidopsis BAC contigs (Lysak et al. 2003). The extent of shared chromosomal homeology reflects inter-species relatedness and the character of chromosomal rearrangements generating the extant structural variation. Only Arabidopsis BACs are being used as painting probes for CCP as chromosome-specific BAC libraries are not available for other crucifer species. However, CCP using other than Arabidopsis BAC contigs is feasible, as shown for BAC clones of *A. halleri* used together with Arabidopsis painting probes for CCP (Lysak et al. 2010). As all Arabidopsis BACs possess

exact chromosome coordinates, the clones may be combined and differently labelled according to a required experimental scheme. CP and CCP in Arabidopsis and crucifer species, respectively, are based on shared coding sequences and therefore BACs containing repeats cannot be used as painting probes. It follows that heterochromatic and repeat-rich pericentromere regions are not covered by Arabidopsis BAC contigs, and thus, homeologous centromeres cannot be directly identified by CCP. As tiny and condensed mitotic chromosomes, spanning only a few micrometers in size, and similarly sized meiotic chromosomes at late prophase, metaphase and anaphase I were found to be unsuitable for (C)CP, extended pachytene chromosomes have been adopted for detailed cytogenetic analyses in crucifer species.

The untapped potential of CCP for comparative genomics of Brassicaceae was for the first time manifested by two studies of Arabidopsis chromosome At4 in nine species assigned to tribes Arabideae and Camelineae (Comai et al. 2003; Lysak et al. 2003). CCP analysis revealed grossly conserved cross-species chromosomal homeology shared by all the analyzed species as well as unique chromosomal reshuffling in some species, and proven that the comparative painting approach is feasible in crucifer taxa. The two seminal studies were then followed by a number of CCP studies exploring evolutionary patterns of chromosomal and genome evolution across the Brassicaceae (e.g. Geiser et al. 2016; Hay et al. 2014; Lysak et al. 2005, 2006; Mandáková and Lysak 2008; Mandáková et al. 2010a, 2010b, 2012, 2013, 2015a, 2015b, 2016). These results are detailed in section 5.

3.3 Genome Sequencing

Whole-genome sequencing of crucifer taxa has accelerated due to advances in high-throughput NGS methods. These high-throughput, efficient and cost-effective approaches started a new era in comparative genomics of Brassicaceae (e.g. Sharma et al. 2014; Koenig and Weigel 2015). To date, genomes of thirteen species have been completely or partly sequenced (Table 1), and more crucifer genomes are being sequenced under the umbrella of the BMAP sequencing project (http://www.brassica.info/resource/sequencing/bmap.php).

Although *de novo* sequencing is producing millions of reads, placing these sequences together on correct chromosomes represents a serious challenge (see Fierst 2015 for a review). Sequence contigs and scaffolds can be anchored to individual chromosomes using available genetic linkage maps. As the development of genetic markers, establishing sufficiently large mapping populations and their genotyping can be technically difficult, time-consuming and expensive, some whole-genome sequencing initiatives produced only several hundred of scaffolds not anchored to chromosomes (see Table 1). For genome assemblies of diploid Brassicaceae species, comparative cytogenetic maps are a cost and time effective alternative to genetic linkage maps. The sequenced genomes of *Schrenkiella parvula* (Dassanayake et al. 2011) and *Arabis alpina* (Willing et al. 2015) have been assembled using comparative cytomolecular maps constructed through comparative chromosome painting. The same approach was followed for assembling the *Boechera stricta* genome (Mitchell-Olds, Schranz, Mandáková and Lysak; in preparation) and some more crucifer genome sequencing projects will benefit from the available comparative cytogenetic maps.

Table 1. Summary of genome sequences completed for the Brassicaceae species (Sharma et al. 2014, Wang et al. 2015 and author's compilation).

Species	Tribe	n	Genome size (Mb)[a]	Status	Reference
Aethionema arabicum	Aethionemeae	11	240	scaffold	Haudry et al. 2013
Arabidopsis thaliana	Camelineae	5	157	chromosome	AGI 2000
Arabidopsis lyrata	Camelineae	8	245	chromosome	Hu et al. 2011
Arabis alpina	Arabideae	8	375	chromosome	Willing et al. 2015
Brassica oleracea	Brassiceae	9	630	chromosome	Yu et al. 2013
Brassica napus	Brassiceae	19	1130	chromosome	Chalhoub et al. 2014
Brassic rapa	Brassiceae	10	530	chromosome	Wang et al. 2011
Camelina sativa	Camelineae	20	750	chromosome	Kagale et al. 2014
Capsella rubella	Camelineae	8	216	chromosome	Slotte et al. 2013
Eutrema salsugineum	Eutremeae	7	314	chromosome	Yang et al. 2013
Leavenworthia alabamica	Cardamineae	11	316	scaffold	Haudry et al. 2013
Schrenkiella parvula	unknown	7	140	chromosome	Dassanayake et al. 2011
Sisymbrium irio	Sisymbrieae	7	262	scaffold	Haudry et al. 2013

[a] genome size reference from the original references or adapted from Lysak et al. (2009).

4. Ancestral Genomes

Comparative genomics or paleogenomics is based on parsimony reconstructions of ancestral genomes. Only after ancestral genomes are inferred, the evolutionary processes shaping the extant genomic structures can be deciphered. As ancestral genomes are inferred based on a few analyzed extant genomes, information on phylogenetic relationships among the modern species is essential. By creating a „Jurassic database" of ancestral genomes (Ouzounis 2005), we are capable to distinguish between ancestral and derived genome states.

The very first consideration of an ancestral Brassicaceae genome was based on x = 8 to be the most common base number across the Brassicaceae (Warwick and Al-Shehbaz 2006). The notion that the ancestral genome should have eight chromosomes was further suppported by the similarity between the n = 8 genomes of *A. lyrata* and *C. rubella* (see 3.1). Lysak et al. (2006) aimed to prove that a plausible ancestral genome had eight chromosomes and structurally resembled genomes of *A. lyrata* and *C. rubella* by reconstructing karyotype evolution in Arabidopsis (n = 5) and four closely related species with varying chromosome numbers (n = 6 - 8). By CCP analysis they could show that the origin of the Arabidopsis karyotype (n = 5) has been marked by the reduction of the ancestral chromosome number (n = 8). In *Neslia paniculata* (n = 7, Camelineae), *Turritis glabra* (n = 6, Turritideae) and *Hornungia alpina* (n = 6, Descurainieae), CCP revealed largely preserved chromosomal colinearity between six homeologous chromosomes in *Neslia*, and four in *Hornungia* and *Turritis*, and the corresponding chromosome types in *A. lyrata* and *C. rubella*. In summary, the study has shown that the origin of modern crucifer genomes with chromosome numbers ranging from n = 5 to n = 7 can be explained by a series of MCRs independently reducing the ancestral eight chromosomes. These conclusions along with the earlier comparative genetic data (*A. lyrata*, *C. rubella*) led to a hypothetical concept of Ancestral Crucifer Karyotype (ACK) with eight ancestral chromosomes shared among Arabidopsis and closely related clades (Lysak et al. 2006).

The ACK concept was further expanded by defining apparently conserved genomic blocks which make up the eight ancestral chromosomes (Schranz et al. 2006). The conserved ancestral blocks were revealed through comparative genetic mapping between Arabidopsis and three other crucifer species, namely *B. napus* (Parkin et al. 2005), *A. lyrata* and *C. rubella* (e.g. Koch and Kiefer 2005; Lysak and Lexer 2006), and by comparative cytogenetics (Lysak et al. 2006). A synthesis of these complementary data resulted in establishing a set of 24 (A - X) conserved genomic blocks together building up the eight chromosomes of the ACK (**Fig. 2**; Schranz et al. 2006). Recently, in the light of newly available structural and sequence data, the ACK has been revised and the number of genomic blocks reduced to 22 (**Fig. 3** and **Table 2**, Lysak et al. 2016).

Since its definition a decade ago (Schranz et al. 2006), the ACK system has become the most important reference genome in genomic studies across the Brassicaceae (see Franzke et al 2011; Lysak and Koch 2011; Lysak et al. 2016 for reviews; Fig. 3). Besides the evolutionary significance of genomic blocks, the ACK concept unified and simplified any cross-species comparison between crucifer genomes. Genomic blocks were also utilized to describe the structure of newly sequenced and assembled crucifer genomes (Dassanayake et al. 2011; Wang et al. 2011; Kagale et al. 2014; Willing et al. 2015).

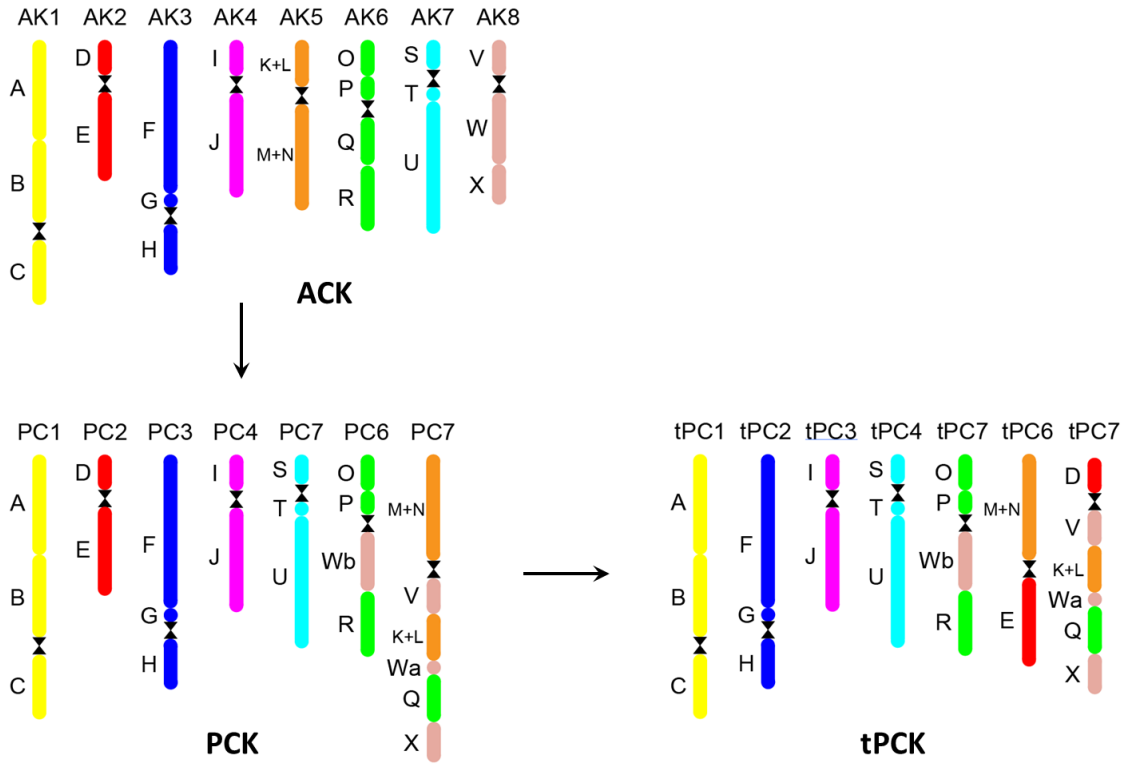

Figure 2. Three ancestral crucifer karyotypes inferred from comparative genetic, cytogenetic and genome sequencing studies. ACK: Ancestral Crucifer Karyotype, PCK: Proto-Calepineae Karyotype and tPCK: translocation PCK. The 22 genomic blocks are indicated by uppercase letters (A to X) and colored according to their position on chromosomes AK1 to AK8 of the ACK (Schranz et al. 2006; Lysak et al. 2016).

Any ancestral genome is a relative concept applicable only to a monophyletic clade at certain time point, meaning that any ancestral genome is preceded and succeeded yet by another genome(s). When analyzing a number of species, all with genomes based on seven chromosomes (x = 7), belonging to six tribes of the crucifer Clade B (= Lineage II) or being closely affiliated with this group (Fig. 1), an ancestral Proto-Calepineae Karyotype (PCK; n = 7; Fig. 2) was inferred (Mandáková and Lysak 2008). As we could show that the PCK genome shares five chromosomes and conserved associations of genomic blocks with the ACK, we proposed that the PCK is an evolutionary younger descendant of the ACK. This descending dysploidy was an important alteration of the eight ancestral chromosomes, presumably associated with the diversification of the species-rich Clade B (Fig. 1). In some Lineage II genera and tribes, the PCK has been altered to the translocation PCK (tPCK; Fig. 2) due to an additional reciprocal chromosome translocation (Mandáková and Lysak 2008; Cheng et al. 2013).

ACK and PCK are the most frequently inferred ancestral genomes for the hitherto analyzed Brassicaceae species. As ancestral genomes for many crucifer taxa remain rather uncertain, more species from less known crucifer clades have to be analyzed to corroborate the exitsting evolutionary patterns or to unveil new ancestral genomes.

5. Karyotype and Genome Evolution

5.1 Karyotype Evolution

Although the first decade of modern comparative cytogenomics of crucifers was marked by several important discoveries, information on the karyotype structure is available for only a handful of species (Table 2, Lysak et al. 2016). Except for technological and financial limitations, a more focused targetting of key genera and species is hindered by unresolved relationships at the base of the Brassicaceae family tree (but see Huang et al. 2015 and Stockenhuber et al. 2015 for possible solutions). Despite these impediments some preliminary conclusions on karyotype evolution have been drawn from comparative genetic and cytogenetic mapping studies, and genome sequencing projects.

As discussed earlier (see section 4.), the ACK genome with eight chromosomes and 22 genomic blocks was widely accepted as an ancestral genome presumably predating the diversification of the Brassicaceae crown group. It can be hypothesized that ACK actually represents a diploidized tetraploid paleogenome formed by the At-α paleopolyploidization and that all crown-group genomes have descended from ACK. Since its formation, ACK has evolved following four basic patterns: (a) genome stasis without major chromosomal reshuffling (e.g. *A. lyrata*, *C. rubella*), (b) prevailing genome stasis with a few structural rearrangements (e.g. diploid *Cardamine* species), (c) the ancestral chromosome number but substantial genome reshuffling (e.g. *Arabis alpina*), and (d) descending dysploidy from n = 8 to n = 7, 6, and 5 (e.g. *A. thaliana*, PCK).

Figure 3 outlines our current understanding of crucifer karyotype evolution based on comparative genetic and cytogenomic data of the last decade put in phylogenetic context of the recently resolved five crown-group clades (Huang et al. 2015). In **Clade A** (= Lineage I), ACK can be identified as the ancestral genome of this lineage. ACK was shown to be an ancestral progenitor of tribes Camelineae, Boechereae, Cardamineae, Crucihimalayeae, Descurainieae, Halimolobeae, Microlepidieae, Smelowskieae, and Turritideae (Table 2). In

five tribes, ACK has undergone independent chromosome number reductions to seven or six chromosomes. Descending dysploidy from n = 8 to n = 7 preceded the radiation in tribes Boechereae and Descurainieae. Smelowskieae and both Turritideae species have six chromosome pairs. Some modern Camelineae genomes (*Arabidopsis* n = 8 species, *Capsella,* Halimolobeae) still retain the conserved ancestral structure of ACK or the eight ancestral chromosomes were reshuffled without descending dysploidy (Cardamineae and Crucihimalayeae). The origin of the Microlepidieae, endemic to Australia and New Zealand, was associated probably with two hybridization events. *Pachycladon* originated through a merger of two ACK-like genomes (Mandáková et al. 2010b), whereas the Australian Microlepidieae were formed upon hybridization between ACK and ACK-derived n = 7 genome similar to that of Descurainieae and Smelowskieae (Mandáková et al. 2010a and unpublished data).

Clade B (= Lineage II and expanded Lineage II) contains a number of important model and crop species whose genomes were sequenced (e.g. *Brassica* species, *Schrenkiella parvula*). Clade B seems to be a monophyletic group defined by a shared ancestral n = 7 genome, namely PCK (see 4. Ancestral Genomes). The origin of PCK from the older ACK genome via descending dysploidy occurred prior the diversification of the whole clade into its approx. 10 tribes (Fig. 1). PCK genome is remarkably stable throughout all tribes lacking later tribe-specific whole-genome duplications (see 4.9 below), i.e. in Calepineae, Conringieae and Coluteocarpeae (N.B. Coluteocarpeae genomes were secondary altered by inversions, Mandáková et al. 2015). The only major rearrangement PCK has undergone was a whole-arm translocation leading to the origin of an evolutionary younger genome variant – tPCK, found in the ancestry of Brassiceae, Eutremeae, Isatideae, Sisymbrieae, and in *Schrenkiella parvula*. Although data are not available for all Clade B tribes, it can be concluded that diversification and radiations in Clade B were not associated with major chromosomal repatterning and that MCRs did not played an important role in speciation in this crucifer lineage.

Karyotype evolution in the three remaining clades of the Brassicaceae crown group remains to be surrounded by question marks. As no genetic or cytogenomic data are available for the tribes of **Clade E** (= Lineage III) and **Clade D** (Alysseae), it remains hard to speculate on ancestral genomes and genome evolution in these clades. Some first insight into evolution of Clade C offers the recently reconstructed genome of *Biscutella laevigata* (see 4.3 and Fig. X). The tetraploid proto-*Biscutella* genome seemed to descend from a so-called ancPCK genome (n = 8), presumably representing a transition between two ancestral genomes, namely between ACK (n = 8) and PCK (n = 7). However, more data from other *Biscutella* species and Clade C tribes are needed to reinforce this conclusion. The recently sequenced *A. alpina* genome (Arabideae, Willing et al. 2015) was not yet assigned to one of the five clades, however, the n = 8 genome has unambiguously descended from an ACK-like ancestor.

Table 2. Presence/abscence of ancestral genomic block associations in 35 crucifer species.

Species	Tribe Clade [a] (Lineage [b])	Ancestral genomic block association										Reference
		A-B	**F-G**	**K-L**	**M-N**	**O-P**	**Q-R**	**T-U**	**W-X**	**Wb-R [c]**	**V-K-L-Wa-Q-X [c]**	
ACK ancestor												
Arabidopsis lyrata	Camelineae A (I)	+	+	+	+	+	+	+	+	-	-	Lysak et al. 2006
Arabidopsis thaliana	Camelineae A (I)	+	-	-	+	+	+	+	+	-	-	Schranz et al. 2006
Ballantinia antipoda [d]	Microlepidieae A (I)	+	+	+	+	-	+	-	+	-	-	Mandáková et al. 2010a
Boechera divaricarpa	Boechereae A (I)	+	+	+	+	+	+	+	+	-	-	Mandáková et al. 2015a
Boechera stricta	Boechereae A (I)	+	+	+	+	+	+	+	+	-	-	Mandáková et al. 2015a
Camelina sativa [e]	Camelineae A (I)	+	+	+	+	+	+	+	+	-	-	Kagale et al. 2014b
Cardamine amara	Cardamineae A (I)	+	+	+	+	+	+	+	+	-	-	Mandáková et al. 2013
Cardamine flexuosa [d]	Cardamineae A (I)	+	+	+	+	+	+	+	+	-	-	Mandáková et al. 2014
Cardamine hirsuta	Cardamineae A (I)	+	+	+	+	+	+	+	+	-	-	Hay et al. 2014
Capsella rubella	Camelineae A (I)	+	+	+	+	+	+	+	+	-	-	Slotte et al. 2013
Crucihimalaya wallichii	Crucihimalayeae A (I)	+	+	+	+	+	+	+	+	-	-	Mandáková et al. 2010a
Hornungia alpina	Descurainieae A (I)	+	+	+	+	+	+	+	+	-	-	Lysak et al. 2006
Neslia paniculata	Camelineae A (I)	+	+	+	+	+	+	+	+	-	-	Lysak et al. 2006
Pachycladon exilis [d]	Microlepidieae A (I)	+	+	+	+	+	+	+	+	-	-	Mandáková et al. 2010b
Stenopetalum lineare [d]	Microlepidieae A (I)	-	+	+	+	-	-	+	+	-	-	Mandáková et al. 2010a
Stenopetalum nutans [d]	Microlepidieae A (I)	-	+	+	+	-	-	-	+	-	-	Mandáková et al. 2010a
Transberingia bursifolia	Crucihimalayeae A (I)	+	+	+	+	+	+	+	+	-	-	Mandáková et al. 2010a
Turritis glabra	Turritideae A (I)	+	+	+	+	+	+	+	+	-	-	Lysak et al. 2006
PCK ancestor												
Brassica oleracea [e]	Brassiceae B (II)	-	-	+	-	+	-	-	-	+	+	Parkin et al. 2014
Brassica rapa [e]	Brassiceae B (II)	-	-	+	-	+	-	-	-	+	+	Kim et al. 2014,
Brassica napus [d+e]	Brassiceae B (II)	-	-	+	-	+	-	-	-	+	+	Parkin et al. 2014
Calepina irregularis	Calepineae B (EII)	+	+	+	+	+	-	+	-	+	+	Mandáková and Lysak 2008
Conringia orientalis	Conringieae B (EII)	+	+	+	+	+	-	+	-	+	+	Mandáková and Lysak 2008
Caulanthus amplexicaulis [d]	Thelypodieae ? (EII)	+	-	?	+	-	-	+	-	+	+	Burrell et al. 2011
Glastaria glastifolia	Isatideae B (II)	+	+	+	+	+	-	+	-	+	+	Mandáková and Lysak 2008
Goldbachia laevigata [d]	Calepineae B (EII)	+	+	+	+	-	-	+	-	+	+	Mandáková and Lysak 2008
Myagrum perfoliatum	Isatideae B (II)	+	+	+	+	+	-	+	-	+	+	Mandáková and Lysak 2008
Noccaea caerulescens	Coluteocarpeae B (EII)	+	+	+	+	-	-	-	-	+	+	Mandáková et al. 2015b
Noccaea jankae [d]	Coluteocarpeae B (EII)	+	+	+	+	+	-	-	-	+	+	Mandáková et al. 2015b
Ochthodium aegyptiacum	Sisymbrieae B (II)	+	+	+	+	+	-	+	-	+	+	Mandáková and Lysak 2008
Raparia bulbosa [d]	Coluteocarpeae B (EII)	+	+	+	+	+	-	-	-	+	+	Mandáková et al. 2015b
Schrenkiella parvula	unassigned B (EII)	+	+	+	+	+	-	+	-	+	+	Mandáková and Lysak 2008
Thellungiella salsuginea	Eutremeae B (EII)	+	+	+	+	+	-	+	-	+	+	Wu et al. 2012
Unresolved ancestral genome												
Arabis alpina	Arabideae ? (EII)	+	+	+	+	-	+	-	+	-	-	Willing et al. 2015
Biscutella laevigata [d]	Biscutelleae C (?)	+	+	+	+	+	-	?	-	+	-	Geiser et al. 2016

[a] according to Huang et al. 2015
[b] according to Franzke et al. 2011
[c] GB associations specific for PCK
[d] species with a whole-genome duplication postdating the At-alpha WGD
[e] species with a whole-genome triplication postdating the At-alpha WG

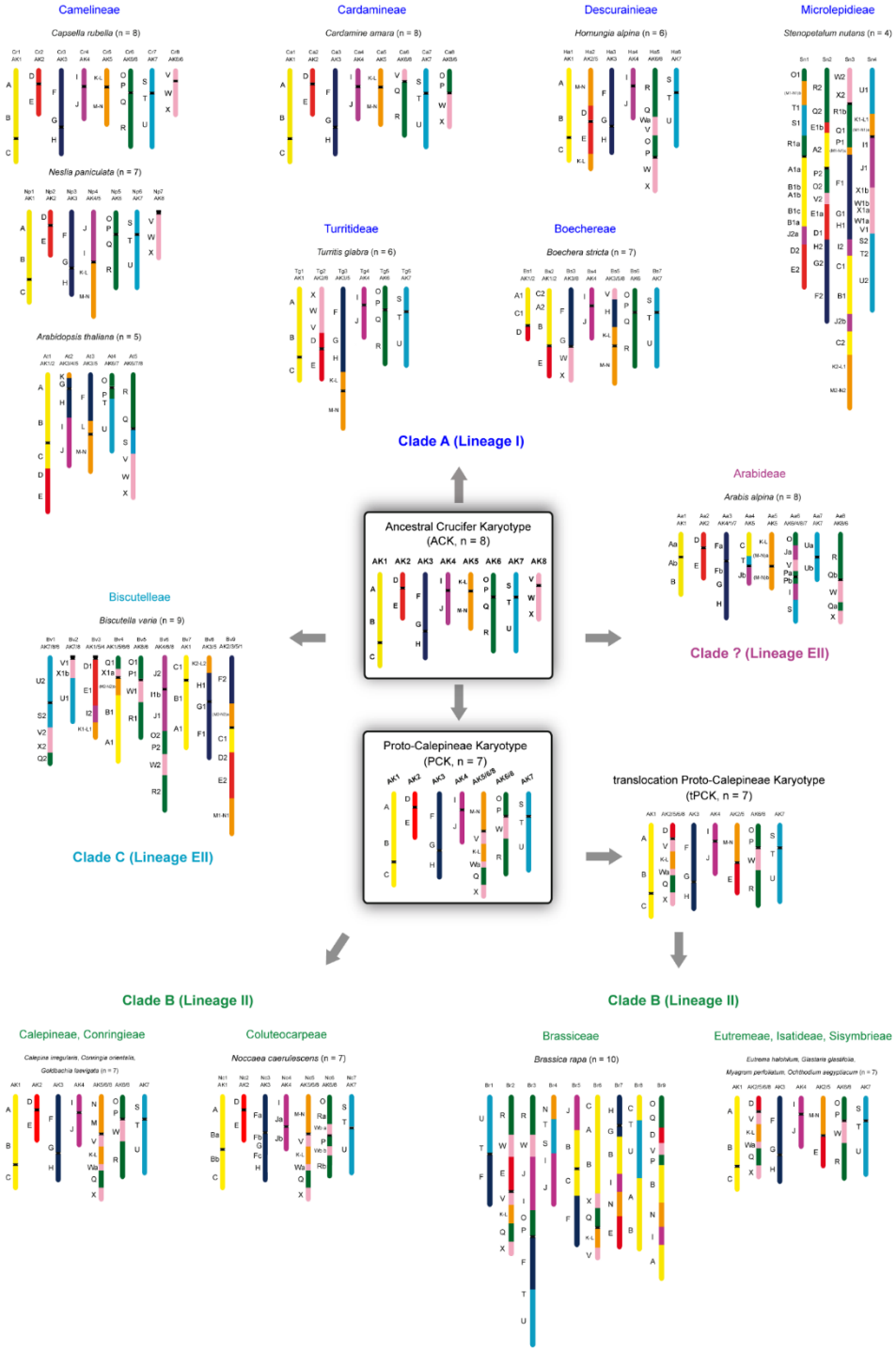

Figure 3. Schematic overview of karyotype evolution in the Brassicaceae. The tentative evolutionary scenario assumes the Ancestral Crucifer Karyotype (ACK, n = 8) being ancestral for crown-group clades resolved in Figure 1B (Huang et al. 2015). Ancestral and modern karyotypes of Clade A (Lineage I), Clade B (Lineage II and expanded Lineage II), Clade C (expanded Lineage II) and that of *Arabis alpina* have presumably descended from the ACK. The 22 genomic blocks are indicated by uppercase letters (A to X) and colored according to their position on chromosomes AK1 to AK8 of the ACK (Schranz et al. 2006, Lysak et al. 2016). Basde on data from Cheng et al. (2013), Geiser et al. (2016), Lysak et al. (2006), Mandáková and Lysak (2008), Mandáková et al. (2010a, 2013, 2015a, 2015b) and Willing et al. (2015).

5.2 Whole-Genome Duplications

Since the genome sequence of *A. thaliana* was published (AGI 2000), the ever-increasing numbers of whole-genome duplication (WGD) events have been identified in both gymnosperms and angiosperms (e.g. Soltis et al. 2009; Vanneste et al. 2014; Li et al. 2015, Soltis et al. 2015). The repeated cycles of polyploidization and subsequent diploidization most likely played a key role in the diversification of early angiosperm lineages described by Charles Darwin as an abominable mystery (De Bodt et al. 2005) and are believed to be advantageous for plant survival at the Cretaceous-Paleogene (K-Pg) mass extinction event c. 66 million years ago (mya) (Vanneste et al. 2014).

The family Brassicaceae, prior its first radiation some 32 mya (Hohmann et al. 2015), went through three WGD events. Crucifers as all dicots share an ancient triplication event (γ; Jaillon et al. 2007). A younger paleotetraploidy event (β) is specific for the core Brassicales including Brassicaceae (Ming et al. 2008; Tang et al. 2008), whereas the α (aka At-α) tetraploidization event (Blanc et al. 2003; Bowers et al. 2003) has occurred about 40 (Edger et al. 2015) to 47 mya (Kagale et al. 2014) prior the split of the Brassicaceae crown group from the Clade F (Aethionemeae).

The three WGDs older than 40 (47) mya can be classified as paleopolyploidizations. The footprints of the youngest α-paleopolyploidization were revealed as large-scale intra- and inter-chromosomal segmental duplications after sequencing of the Arabidopsis genome. However, the segmental duplications cannot be detected cytogenetically and genetically by chromosome painting and genetic mapping, respectively. In contrast, duplicated genomic regions resulting from WGDs post-dating the At-α can be detected by CCP and comparative genetic mapping (e.g. Lagercrantz 1998, Parkin et al. 2005; Lysak et al. 2005, 2007; Ziolkowski et al. 2006). Relative to the α-paleopolyploidization, younger events are called mesopolyploid WGDs (Mandáková et al. 2010), and are characterized by extensive diploidization (often diploid-like meiosis, disomic inheritance, low pseudo-diploid chromosome numbers). The very recent polyploidization events (neopolyploidy) are easily recognized by increased chromosome numbers and by the absence or onset of diploidization (e.g. Mandáková et al. 2013, 2014, 2016). The origin of several crucifer neopolyploids was dated to the Pleistocene (Marhold and Lihová 2006), 1.8 million to 11,500 years before present.

As mesopolyploid WGDs were masqueraded by the chromosomal and genetic diploidization processes involving MCRs leading to chromosome number reduction, genome size downsizing and diploid-like inheritance, these events are not easy to detect. CCP was invaluable in uncovering clade-specific mesopolyploidies. The hexaploidization event specific for *Brassica* and most probably for the whole tribe Brassiceae was the first mesopolyploidy event discovered (Lagercrantz 1998; Parkin et al. 2005; Lysak et al. 2005, 2007; Ziolkowski et al. 2006; Wang et al. 2011). To date, four tribes (Heliophileae [Mandáková et al. 2012], Microlepidieae [Mandáková et al. 2010a, 2010b], Brassiceae [Lysak et al. 2005; Wang et al. 2011], and Biscutelleae [Geiser et al. 2016]) and the genus *Laevenworthia* (Haudry et al. 2013) have been shown to be mesopolyploids, which have undergone post-At-α polyploidizations. However, the total number of mesopolyploidization events is higher. Some additional mesopolyploidy events have been postulated (*Pringlea* and *Stanleya* [Kagale et al. 2014], *Caulanthus* [Burrell et al. 2011], Physarieae [Lysak et al. 2009], Thelypodieae, and unpublished data from the Lysak lab), pending further validation.

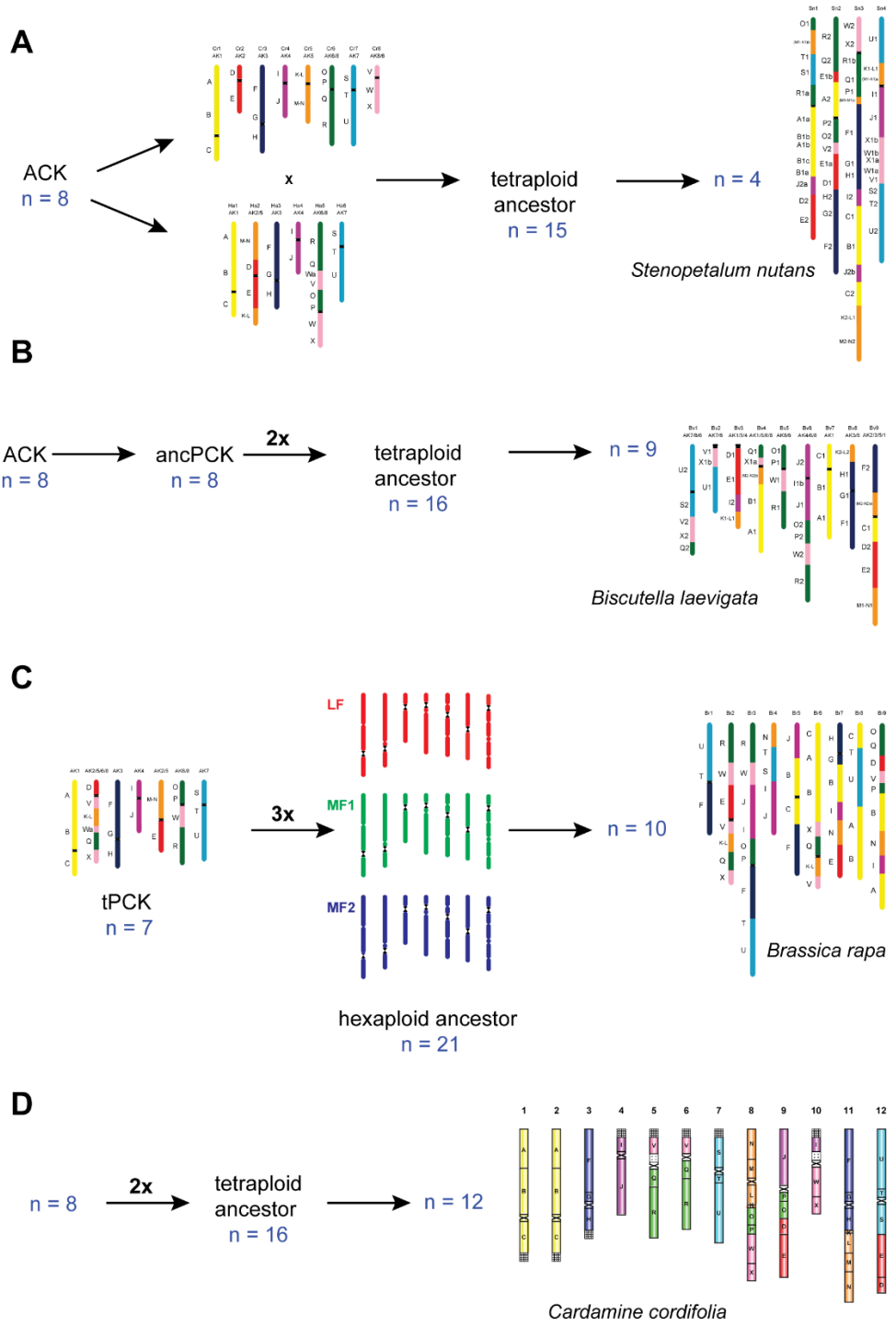

Figure 4. Examples of mesopolyploid whole-genome duplications and triplications post-dating the At-α paleopolyploidization. Clade-/taxon-specific polyploidizations are followed by genome diploidization, shuffling the duplicated genomic blocks and reducing chromosome number. The process of mesopolyploidization followed by diploidization is shown for the Microlepidieae (A), the Biscutelleae(B), *Brassica* (C) and *Cardamine cordifolia* (D). The 22/24 genomic blocks are indicated by uppercase letters (A to X) and colored according to their position on chromosomes AK1 to AK8 of the ACK (Schranz et al. 2006; Lysak et al. 2016). Based on data from Cheng et al. (2013), Geiser et al. (2016) and Mandáková et al. (2010a, 2016).

Several basic conclusions can be drawn from recently unveiled clade-specific mesopolyploid WGDs (i) genome evolution in Brassicaceae is characterized by cyclic rounds of WGDs followed by diploidization, (ii) WGDs seem to be important drivers of genetic, morphological and physiological diversification, and thus, of species radiations, (iii) chromosome number is not a realiable indicator of taxon's evolutionary past, and therefore the assessed percentage of polyploid crucifer species (c. 40%; Warwick and Al-Shehbaz 2006) is grossly underestimated.

5.3 Mechanisms of chromosomal evolution

Morphology and number of chromosomes change due to MCRs, also known as structural aberrations. Some MCRs, such pericentric inversions, large-scale duplication and deletions, impact shape and size of chromosomes, whereas other rearrangements, such as chromosome translocations, are important in reducing the number of chromosomes (descending dysploidy). Here we discuss only the most important mechanisms driving karyotype evolution in the Brassicaceae, including those (re)discovered in crucifer studies but having similar roles also in other angiosperm families (Schubert and Lysak 2011; Lysak 2014).

Inversions

In genome evolution of the Brassicaceae, pericentric inversions seemed to play a two-fold role. First, pericentric inversions can turn a (sub)metacentric or acrocentric chromosome into an acrocentric or telocentric one. Consequently, the pericentromere with adjacent repeat-rich heterochromatic arrays are relocated towards a chromosome terminus. Then a translocation between two chromosome ends may result into a merger of the two chromosomes and loss of one centromere. Thus, a pericentric inversion is a necessary pre-requisite of a subsequent terminal translocation (see below). Second, both para- and pericentromeric inversions change the position of genes and regulatory regions on a chromosome. This may or may not have a phenotypic effect and evolutionary significance (e.g. Rieseberg 2001; Lowry and Willis 2010). Due to pericentric inversions with one breakpoint close to or within the pericentromere and another one in euchromatic region, genes from within euchromatin are brought into the immediate proximity of the pericentromeric heterochromatin, and vice versa, parts of pericentromeric heterochromatin are relocated into gene-rich regions. It is well established that heterochromatin spreading into neighboring euchromatic regions can cause gene silencing or modify gene expression, whereas the relocation of genes from the vicinity of pericentromeric heterochromatin into euchromatic context may conversely enhance the expression of the relocated genes. A change in chromosomal position can also alter recombination rates, which are known to be nonrandom along chromosomes and usually suppressed in centromeric regions (e.g. Rieseberg 2001; Mercier et al. 2015). Pericentric inversions might generally contribute to recombination of low-recombining regions when they are brought to more distal, heterochromatin-free positions. Although inversions were frequent during genome evolution of crucifer species, their role in reproductive isolation and speciation is hard to evaluate. Some evidence for potential adaptive role of pericentromeric inversion was obtained in *Noccaea caerulescens* (*Thlaspi caerulescens*), an extremophile heavy metal hyperaccumulator species exhibiting metal tolerance and leaf accumulation of zinc,

cadmium, and nickel. The enhanced gene expression in *N. caerulescens*, especially that of metal homeostasis genes, was associated with overall more distal chromosome positions of genes, or the inversion-mediated movement to a more distal chromosomal position, when compared with the ancestral PCK karyotype (Mandáková et al. 2015). In the adaptation to changing environments and new ecological niches, a tight linkage of genes into a supergene can become advantageous. Inversions can be implicated as a mechanism of supergene formation. In *N. caerulescens*, two inversions led to supergene formation on chromosome NC6: Cu and organellar metal homeostasis genes clustered on the top chromosome arm, whereas Zn/Fe homeostasis genes likely related to metal hyperaccumulation and hypertolerance accumulated on the bottom arm. Thus, inversions in *N. caerulescens* could have facilitated the evolution of enhanced metal homeostasis gene expression, a known hallmark of metal hyperaccumulation (Mandáková et al. 2015).

Terminal chromosome translocations

Chromosome translocation events may lead to exchanges of large chromosome regions between two or more non-homologous chromosomes. Some chromosome translocation may result in evolutionary fixed reduction of chromosome number, i.e. descending dysploidy. These terminal chromosome translocations (TCTs, Lysak 2014) occur through translocation events involving terminal regions of two chromosomes – usually with breakpoints close to the centromere of the long arm of a telo- or acrocentric chromosome and one end of any type of chromosome (telo- or acrocentrics are frequently results of pericentric inversions, see above). TCTs result in a large translocation (incorrectly "fusion") chromosome comprising most parts of the original chromosomes and a small centromere-containing minichromosome. The minichromosome product, comprising dispensable sequences, is meiotically unstable due to its inability to form chiasmata and segregate regularly and/or because of insufficient sister chromatid cohesion during the first meiotic division, and as such is eliminated. TCTs include so-called Robertsonian translocations, that is, whole-arm translocations (sometimes confusingly called centric fusions), that transform two telocentric or acrocentric (rod-shaped) chromosomes into one V-shaped (sub)metacentric chromosome (e.g. Robertson 1916; Darlington 1937; Tobgy 1943; Schubert and Lysak 2011; Stimpson et al. 2012). Some Robertsonian translocations may yield a dicentric translocation chromosome and an acentric fragment. If the two centromeres are not too distant, the dicentric chromosome can be stably inherited (Stimpson et al. 2012, Lysak 2014). TCTs between two metacentrics or between an acro- or telocentric and metacentric chromosome with breakpoints close to chromosome termini (sometimes incorrectly called end-to-end fusions) result in the origin of a dicentric chromosome. As the centromeres on a dicentric chromosome are not in a close proximity, the translocation chromosome has to become stabilized by centromere inactivation and/or deletion (e.g, Mandáková et al. 2010a, b; Lysak 2014; Mandáková et al. 2015; Willing et al. 2015). It would be expected that centromere removal will follow centromere inactivation, however both processes might concur. Molecular mechanism(s) of centromere inactivation remain elusive and presumably include predominantly epigenetic modification of centromere-specific proteins. Centromere loss (deletion) is thought to be associated with a removal of centromere-specific sequences, for example, via misrepair of double-strand breaks on opposite sites of a centromere, or by unequal recombination between LTR retrotransposons or tandem repeats on opposite centromere borders (see Lysak 2014).

Centromere repositioning and neocentromere formation

Centromere shift or repositioning (CR) refers to an intra-chromosomal centromere relocation without alterations in chromosome collinearity. CR consists of two interlinked processes, namely 'decay' of old centromere-specific sequences and epigenetic marks, and the emergence of a new functional centromere with appropriate epigenetic characteristics. Clearly, the mechanics of gradual 'decay'of an old centromere and formation of a new centromere on the same chromosome faces the difficulty of overcoming instability by a transient acentric or a dicentric state. In crucifers, CRs were documented only through an increasing number of case studies published over the last decade. A CR was assumed to take place in *Cardamine rivularis* (Mandáková et al. 2013) and three independent CRs were inferred to occur during evolution of the *A. alpina* genome (Willing et al. 2015). Although originally thought to be absent or very rare, CR seems to be a relatively frequent phenomenon in chromosomal evolution of the Brassicaceae (unpublished data from the Lysak lab).

In plants, a newly formed centromere without centromere-specific sequences but with centromere-specific epigenetic marks was reported for a maize A-chromosome fragment (Fu et al. 2013) or B chromosome (Zhang et al. 2013). In the Brassicaceae, a neocentromere emergence can be inferred only for one of the eight chromosomes of *A. alpina* (Willing et al. 2015).

Chromosome fission

Compared to descending dysploidy, the opposite process of decreasing chromosome numbers (ascdending dysploidy) was not know from the family Brassicaceae until recently (Mandáková et al. 2015). Ascending dysploidy refers to a fixed aneuploidy-like increase of chromosome number different from whole-genome multiplications (polyploidy) or agmatoploidy in taxa with holokinetic chromosomes. In a crucifer genus *Boechera*, some species with euploid chromosome number (2n = 2x = 14) have either sexual or apomictic mode of reproduction, while some apomictic populations possess 15 chromosomes (2n = 2x+1 = 15) (e.g. Kantama et al. 2007). Our recent CCP study (Mandáková et al. 2015) revealed that in diploid apomictic plants one of the two (BS1) homologues is largely heterochromatic (= *Het* chromosome). In apomictic *B. polyantha* (2n = 15), the *Het* has undergone a centric fission within the heterochromatin-rich centromere resulting into two smaller chromosomes *Het'* and *Del* (for deletion). As the *Het* resulted from a whole-arm translocation between two different chromosomes, the *Het* centromere can be potentially prone to fission due to its 'hybrid' nature and/or breakpoint reuse. These findings demonstrate that new chromosomes can be formed by a centric fission and can be fixed in populations due to the apomictic mode of reproduction.

6. Conclusions and perspectives

Crucifers and grasses represent two families with the most advanced comparative cytogenomics and phylogenomics among all land plant. Consequently, a wealth of important findings and discoveries comes from these two families. Comparative cytogenomics of crucifers was developed based on three milestones, namely sequencing of the Arabidopsis genome, the feasibility of comparative chromosome painting, and high throughput genome

sequencing. Even the data on genome structure of Brassicaceae species are still scant, some evolutionary patterns and trends are emerging. Some crucial conclusions are listed here:

1) A working hypothesis of an ancestral genome (ACK) with eight chromosomes and 22 genomic blocks was repeatedly corroborated for the crucifer crown group (i.e. 48 tribes and some not yet assigned genera).
2) The ancestral karyotype followed three evolutionary trajectories: (a) genomic stasis (e.g. *A. lyrata*, *C. rubella*), (b) genome shuffling but stasis of the ancestral chromosome number (e.g. *A. alpina*, *Cardamine*), and (c) descending dysploidy (n = 5 – 7) associated with genome shuffling (e.g. *A. thaliana*, *N. paniculata*, PCK).
3) Some major diversifications and species radiations in the Brassicaceae were apparently not associated with chromosomal rearrangements and descending dysploidies. Some genome repatterning with or without descending dysploidy could contribute to reproductive isolation and speciation, however, compelling evidence is missing.
4) The number of tribe- and genus-specific whole-genome duplication (WGD) events is much higher than originally thought. Polyploidization is regularly followed by genome rediploidization turning the polyploid genomes into pseudodiploid ones. Some clade-specific WGDs were most likely important drivers of diversification and species radiation. Possible triggers and rates of the rediploidization process remain to be specified.

During the last fifteen years, the number of cytogenetic reconstructions and/or draft genome sequences of crucifer species increased from one (*A. thaliana*) to more than thirty-five. In the near future, the ongoing comparative cytogenomic analyses and genome sequence projects like the Brassicaceae Map Alignment Project (BMAP) will double, if not triple, the number of available crucifer genomes. Technological innovations such as long-molecule sequencing are setting the stage for complete genome assemblies including repeat-rich regions such as pericentromeres. Besides chromosome-specific BAC contigs, cytogenetic analysis is also now possible using gene-based and custom-synthesized short oligonucleotides, solving the problematic use of repeat-containing BAC clones. Other exciting developments include the use of genome mapping technologies (optical mapping) allowing for extremely long reads of DNA molecules to accurately and rapidly anchor sequence scaffolds to chromosomes. This will revolutionize our ability to analyze the highly dynamic and non-collinear regions of any Brassicaceae taxa including polyploids.

7. References

AGI (The Arabidopsis Genome Initiative, 2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408: 796-815.

Al-Shehbaz IA (2012) A generic and tribal synopsis of the Brassicaceae (Cruciferae). *Taxon* 61: 931-954.

Al-Shehbaz IA, Beilstein MA, Kellogg EA (2006) Systematics and phylogeny of the *Brassicaceae* (*Cruciferae*): an overview. *Plant Syst Evol* 259: 89-120.

Appel O, Al-Shehbaz IA (2003) Cruciferae. In: The Families and Genera of Vascular Plants, Kubitzki K, and Bayer C (eds.), pp. 75-174. Springer-Verlag, Berlin.

Bailey CD, Koch MA, Mayer M, Mummenhoff K, O'Kane SL Jr, Warwick SI, Windham MD, Al-Shehbaz IA (2006) Toward a global phylogeny of the *Brassicaceae*. *Mol Biol Evol* 23: 2142-2160.

Beilstein MA, Al-Shehbaz IA, Kellogg EA (2006) *Brassicaceae* phylogeny and trichome evolution. *Am J Bot* 93: 607-619.

Beilstein MA, Al-Shehbaz IA, Mathews S, Kellogg EA (2008) Brassicaceae phylogeny inferred from phytochrome A and ndhF sequence data: tribes and trichomes revisited. *Am J Bot* 95: 1307-1327.

Berr A, Schubert I (2007) Interphase chromosome arrangement in *Arabidopsis thaliana* is similar in differentiated and meristematic tissues and shows a transient mirror symmetry after nuclear division. *Genetics* 176: 853-863.

Blanc G, Hokamp K, Wolfe KH (2003) A recent polyploidy superimposed on older large-scale duplications in the *Arabidopsis* genome. *Genome Res* 13: 137-144.

Boivin K, Acarkan A, Mbulu R-S, Clarenz O, Schmidt R (2004) The Arabidopsis genome sequence as a tool for genome analysis in *Brassicaceae*. A comparison of the Arabidopsis and *Capsella rubella* genomes. *Plant Physiol* 135: 735-744.

Bowers JE, Chapman BA, Rong J, Paterson AH (2003) Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* 422: 433-438.

Burrell AM, Taylor KG, Williams RJ, Cantrell RT, Menz MA, Pepper AE (2011) A comparative genomic map for *Caulanthus amplexicaulis* and related species (Brassicaceae). *Mol Ecol* 20: 784-798.

Chalhoub B, Denoeud F, Liu S, Parkin IA, Tang H, Wang X, Chiquet J, Belcram H, Tong C, Samans B, Corréa M, Da Silva C, Just J, Falentin C, Koh CS, Le Clainche I, Bernard M, Bento P, Noel B, Labadie K, Alberti A, Charles M, Arnaud D, Guo H, Daviaud C, Alamery S, Jabbari K, Zhao M, Edger PP, Chelaifa H, Tack D, Lassalle G, Mestiri I, Schnel N, Le Paslier MC, Fan G, Renault V, Bayer PE, Golicz AA, Manoli S, Lee TH, Thi VH, Chalabi S, Hu Q, Fan C, Tollenaere R, Lu Y, Battail C, Shen J, Sidebottom CH, Wang X, Canaguier A, Chauveau A, Bérard A, Deniot G, Guan M, Liu Z, Sun F, Lim YP, Lyons E, Town CD, Bancroft I, Wang X, Meng J, Ma J, Pires JC, King GJ, Brunel D, Delourme R, Renard M, Aury JM, Adams KL, Batley J, Snowdon RJ, Tost J, Edwards D, Zhou Y, Hua W, Sharpe AG, Paterson AH, Guan C, Wincker P (2014) Plant genetics. Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science* 22: 950-953.

Cheng F, Mandáková T, Wu J, Xie Q, Lysak MA, Wang X (2013) Deciphering the diploid ancestral genome of the mesohexaploid *Brassica rapa*. *Plant Cell* 25: 1541-1554.

Choi S, Creelman RA, Mullet JE, Wing R (1995) Construction and characterization of a bacterial artificial chromosome library of *Arabidopsis thaliana*. *Plant Mol Biol Report* 13: 124-128.

Comai L, Tyagi AP, Lysak MA (2003) FISH analysis of meiosis in Arabidopsis allopolyploids. *Chromosome Res* 11: 217-226.

Darlington CD (1937) Recent Advances in Cytology. J. & A. Churchill, London.

Dassanayake M, Oh DH, Haas JS, Hernandez A, Hong H, Ali S, Yun DJ, Bressan RA, Zhu JK, Bohnert HJ, Cheeseman JM (2011) The genome of the extremophile crucifer *Thellungiella parvula*. *Nat Genet* 43: 913-918.

De Bodt S, Maere S, Van de Peer Y (2005) Genome duplication and the origin of angiosperms. *Trends Ecol Evol* 20: 591-597.

Delourme R, Falentin C, Fomeju BF, Boillot M, Lassalle G, André I, Duarte J, Gauthier V, Lucante N, Marty A, Pauchon M, Pichon JP, Ribière N, Trotoux G, Blanchard P, Rivière N, Martinant JP, Pauquet J (2013) High-density SNP-based genetic map development and linkage disequilibrium assessment in *Brassica napus* L. *BMC Genomics* 14: 120.

Edger PP, Heidel-Fischer HM, Bekaert M, Rota J, Glöckner G, Platts AE, Heckel DG, Der JP, Wafula EK, Tang M, Hofberger JA, Smithson A, Hall JC, Blanchette M, Bureau TE, Wright SI, dePamphilis CW, Eric Schranz M, Barker MS, Conant GC, Wahlberg N, Vogel H, Pires JC, Wheat CW (2015) The butterfly plant arms-race escalated by gene and genome duplications. *Proc Natl Acad Sci USA* 112: 8362-8366.

Fierst JL (2015) Using linkage maps to correct and scaffold de novo genome assemblies: methods, challenges, and computational tools. *Front Genet* 6: 220.

Fransz P, de Jong JH, Lysak M, Castiglione MR, Schubert I (2002) Interphase chromosomes in *Arabidopsis* are organized as well defined chromocenters from which euchromatin loops emanate. *Proc Natl Acad Sci USA* 99: 14584-14589.

Franzke A, German D, Al-Shehbaz IA, Mummenhoff K (2008) *Arabidopsis* family ties: Molecular phylogeny and age estimates in the Brassicaceae. *Taxon* 58:425-437.

Franzke A, Lysak MA, Al-Shehbaz IA, Koch MA, Mummenhoff K (2011) Cabbage family affairs: the evolutionary history of Brassicaceae. *Trends Plant Sci* 16: 108-116.

Fu S, Lv Z, Gao Z, Wu H, Pang J, Zhang B, Dong Q, Guo X, Wang XJ, Birchler JA, Han F (2013) De novo centromere formation on a chromosome fragment in maize. *Proc Natl Acad Sci USA* 110: 6033-6036.

Geiser C, Mandáková T, Arrigo N, Lysak MA, Parisod C (2016) Repeated whole-genome duplication, karyotype reshuffling and biased retention of stress-responding genes in Buckler Mustards. *Plant Cell* (in press).

Hall JC, Sytsma KJ, Iltis HH (2002) Phylogeny of *Capparaceae* and *Brassicaceae* based on chloroplast sequence data. *Am J Bot* 89: 1826-1842.

Hall JC, Iltis HH, Sytsma KJ (2004) Molecular phylogenetics of core Brassicales, placement of orphan genera *Emblingia*, *Forchhammeria*, *Tirania*, and character evolution. *Syst Bot* 29: 654-669.

Haudry A, Platts AE, Vello E, Hoen D, Leclercq M, Williamson R, Joly-Lopez Z, Forczek E, Steffen J, Hazzouri K, Dewar K, Stinchcombe JR, Schoen DJ, Wang X, Schmutz J, Town CD, Edger PP, Pires JC, Schumaker KS, Jarvis DE, Mandáková T, Lysak MA, Schranz ME, van den Bergh E, Harrison P, Moses A, Bureau TE, Wright SI, Blanchette M (2013) An atlas of over 90,000 conserved noncoding sequences provides insight into crucifer regulatory regions. *Nat Genet* 45: 891-898.

Hay AS, Pieper B, Cooke E, Mandáková T, Cartolano M, Tattersall AD, Ioio RD, McGowan SJ, Barkoulas M, Galinha C, Rast MI, Hofhuis H, Then C, Plieske J, Ganal M, Mott R, Martinez-Garcia JF, Carine MA, Scotland RW, Gan X, Filatov DA, Lysak MA, Tsiantis M. (2014) *Cardamine hirsuta*: a versatile genetic system for comparative studies. *Plant J* 78: 1-15.

Hedge IC (1976) A systematic and geographical survey of the Old World *Cruciferae*. In: Vaughn JG, Macleod AJ, Jones BMG (eds.), The Biology and Chemistry of the *Cruciferae*. pp. 1-45, Academic Press, London.

Hohmann N, Wolf EM, Lysak MA, Koch MA (2015) A time-calibrated road map of Brassicaceae species radiation and evolutionary history. *Plant Cell* 27: 2770-2784.

Hu TT, Pattyn P, Bakker EG, Cao J, Cheng JF, Clark RM, Fahlgren N, Fawcett JA, Grimwood J, Gundlach H, Haberer G, Hollister JD, Ossowski S, Ottilar RP, Salamov AA, Schneeberger K, Spannagl M, Wang X, Yang L, Nasrallah ME, Bergelson J, Carrington JC, Gaut BS, Schmutz J, Mayer KF, Van de Peer Y, Grigoriev IV, Nordborg M, Weigel D, Guo YL (2011) The *Arabidopsis lyrata* genome sequence and the basis of rapid genome size change. *Nature Genet* 43: 476-481.

Huang CH, Sun R, Hu Y, Zeng L, Zhang N, Cai L, Zhang Q, Koch MA, Al-Shehbaz I, Edger PP, Pires JC, Tan DY, Zhong Y, Ma H (2015) Resolution of Brassicaceae phylogeny using nuclear genes uncovers nested radiations and supports convergent morphological evolution. *Mol Biol Evol* pii: msv226 [Epub ahead of print]

Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C, Vezzi A, Legeai F, Hugueney P, Dasilva C, Horner D, Mica E, Jublot D, Poulain J, Bruyère C, Billault A, Segurens B, Gouyvenoux M, Ugarte E, Cattonaro F, Anthouard V, Vico V, Del Fabbro C, Alaux M, Di Gaspero G, Dumas V, Felice N, Paillard S, Juman I, Moroldo M, Scalabrin S, Canaguier A, Le Clainche I, Malacrida G, Durand E, Pesole G, Laucou V, Chatelet P, Merdinoglu D, Delledonne M, Pezzotti M, Lecharny A, Scarpelli C, Artiguenave F, Pè ME, Valle G, Morgante M, Caboche M, Adam-Blondon AF, Weissenbach J, Quétier F, Wincker P; French-Italian Public Consortium for Grapevine Genome Characterization (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* 449: 463-467.

Kagale S, Robinson SJ, Nixon J, Xiao R, Huebert T, Condie J, Kessler D, Clarke WE, Edger PP, Links MG, Sharpe AG, Parkin IA (2014a) Polyploid evolution of the Brassicaceae during the Cenozoic era. *Plant Cell* 26: 2777-2791.

Kagale S, Koh C, Nixon J, Bollina V, Clarke WE, Tuteja R, Spillane C, Robinson SJ, Links MG, Clarke C, Higgins EE, Huebert T, Sharpe AG, Parkin IA (2014b) The emerging biofuel crop *Camelina sativa* retains a highly undifferentiated hexaploid genome structure. *Nature Comm* 5: 3706.

Kantama L, Sharbel TF, Schranz ME, Mitchell-Olds T, de Vries S, de Jong H (2007) Diploid apomicts of the *Boechera holboellii* complex display large-scale chromosome substitutions and aberrant chromosomes. *Proc Natl Acad Sci USA* 104: 14026-14031.

Karl R, Koch MA (2013) A world-wide perspective on crucifer speciation and evolution: phylogeny, biogeography and trait evolution in tribe Arabideae. *Ann Bot* 112: 983-1001.

Kim J, Lee J, Choi JP, Park I, Yang K, Kim MK, Lee YH, Nou IS, Kim DS, Min SR, Park SU, Kim H (2014) Functional innovations of three chronological mesohexaploid *Brassica rapa* genomes. *BMC Genomics* 15: 606.

Koch MA, Kiefer M (2005) Genome evolution among cruciferous plants: a lecture from the comparison of the genetic maps of three diploid species - *Capsella rubella*, *Arabidopsis lyrata* subsp. *petraea*, and *A. thaliana*. *Am J Bot* 92: 761-767.

Koch M, Al-Shehbaz IA, Mummenhoff K (2003) Molecular systematics, evolution, and population biology in the mustard family (*Brassicaceae*). *Ann Missouri Bot Gard* 90: 151-171.

Koenig D, Weigel D (2015) Beyond the thale: comparative genomics and genetics of Arabidopsis relatives. *Nat Rev Genet* 16: 285-298.

Kuittinen H, de Haan AA, Vogl C, Oikarinen S, Leppala J, Koch M, Mitchell-Olds T, Langley CH, Savolainen O (2004) Comparing the linkage maps of the close relatives *Arabidopsis lyrata* and *A. thaliana*. *Genetics* 168: 1575-1584.

Lagercrantz U (1998) Comparative mapping between *Arabidopsis thaliana* and *Brassica nigra* indicates that *Brassica* genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. *Genetics* 150: 1217-1228.

Lagercrantz U, Lydiate D (1996) Comparative genome mapping in *Brassica*. *Genetics* 144: 1903-1910.

Li Z, Baniaga AE, Sessa EB, Scascitelli M, Graham SW, Rieseberg LH, Barker MS (2015) Early genome duplications in conifers and other seed plants. *Sci Adv* 1: e1501084.

Liu S, Liu Y, Yang X, Tong C, Edwards D, Parkin IA, Zhao M, Ma J, Yu J, Huang S, Wang X, Wang J, Lu K, Fang Z, Bancroft I, Yang TJ, Hu Q, Wang X, Yue Z, Li H, Yang L, Wu J, Zhou Q, Wang W, King GJ, Pires JC, Lu C, Wu Z, Sampath P, Wang Z, Guo H, Pan S, Yang L, Min J, Zhang D, Jin D, Li W, Belcram H, Tu J, Guan M, Qi C, Du D, Li J, Jiang L, Batley J, Sharpe AG, Park BS, Ruperao P, Cheng F, Waminal NE, Huang Y, Dong C, Wang L, Li J, Hu Z, Zhuang M, Huang Y, Huang J, Shi J, Mei D, Liu J, Lee TH, Wang J, Jin H, Li Z, Li X, Zhang J, Xiao L, Zhou Y, Liu Z, Liu X, Qin R, Tang X, Liu W, Wang Y, Zhang Y, Lee J, Kim HH, Denoeud F, Xu X, Liang X, Hua W, Wang X, Wang J, Chalhoub B, Paterson AH (2014) The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nat Commun* 23: 3930.

Lowry DB, Willis JH (2010) A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biol* 8: pii: e1000500.

Lysak MA (2014) Live and let die: centromere loss during evolution of plant chromosomes. *New Phytol* 203: 1082-1089.

Lysak MA, Lexer C (2006) Towards the era of comparative evolutionary genomics in *Brassicaceae*. *Pl Syst Evol* 259: 175-198.

Lysak MA,Koch MA (2011) Phylogeny, genome, and karyotype evolution of crucifers (Brassicaceae). In: Genetics and Genomics of the Brassicaceae (eds. Schmidt R & Bancroft I), Springer, New York, New York, USA, pp. 1-31.

Lysak MA, Fransz PF, Ali HBM, Schubert I (2001) Chromosome painting in *Arabidopsis thaliana*. *Plant J* 28: 689-697.

Lysak MA, Pecinka A, Schubert I (2003) Recent progress in chromosome painting of *Arabidopsis* and related species. *Chromosome Res* 11: 195-204.

Lysak MA, Koch MA, Pecinka A, Schubert I (2005) Chromosome triplication found across the tribe Brassiceae. *Genome Res* 15: 516-525.

Lysak MA, Berr A, Pecinka A, Schmidt R, McBreen K, Schubert I (2006) Mechanisms of chromosome number reduction in *Arabidopsis thaliana* and related *Brassicaceae* species. *Proc Natl Acad Sci USA* 103: 5224-5229.

Lysak MA, Cheung K, Kitschke M, Bures P (2007) Ancestral chromosomal blocks are triplicated in Brassiceae species with varying chromosome number and genome size. *Plant Physiol* 145: 402-410.

Lysak MA, Koch MA, Beaulieu JM, Meister A, Leitch IJ (2009) The dynamic ups and downs of genome size evolution in *Brassicaceae*. *Mol Biol Evol* 26: 85-98.

Lysak MA, Mandáková T, Lacombe E (2010) Reciprocal and multi-species chromosome BAC painting in crucifers (Brassicaceae). *Cytogenet Genome Res* 129: 184-189.

Lysak MA, Mandáková T, Schranz EM (2016) Comparative paleogenomics of crucifers: ancestral genomic blocks revisited. *Curr Opin Pl Biol* (in press).

Mandáková T, Lysak MA (2008) Chromosomal phylogeny and karyotype evolution in x=7 crucifer species (*Brassicaceae*). *Plant Cell* 20: 2559-2570.

Mandáková T, Joly S, Krzywinski M, Mummenhoff K, Lysak MA (2010a) Fast diploidization in close mesopolyploid relatives of *Arabidopsis*. *Plant Cell* 22: 2277-2290.

Mandáková T, Heenan PB, Lysak MA (2010b) Island species radiation and karyotypic stasis in *Pachycladon* allopolyploids. *BMC Evol Biol* 29: 367.

Mandáková T, Mummenhoff K, Al-Shehbaz IA, Mucina L, Mühlhausen A, Lysak MA (2012) Whole-genome triplication and species radiation in the southern African tribe Heliophileae (Brassicaceae). *Taxon* 61: 989-1000.

Mandáková T, Shimizu Inatsugi R, Zozomová-Lihová J, Shimizu K, Kovařík A, Marhold K, Lysak MA (2013) The more the merrier: recent hybridization and polyploidy in *Cardamine*. *Plant Cell* 25: 3280-3295.

Mandáková T, Marhold K, Lysak MA (2014) The widespread crucifer species *Cardamine flexuosa* is an allotetraploid with a conserved subgenomic structure. *New Phytol* 201: 982-992.

Mandáková T, Schranz ME, Sharbel TF, de Jong H, Lysak MA (2015a) Karyotype evolution in apomictic *Boechera* and the origin of the aberrant chromosomes. *Plant J* 82: 785-793.

Mandáková T, Singh V, Krämer U, Lysak MA (2015b) Genome structure of the heavy metal hyperaccumulator *Noccaea caerulescens* and its stability on metalliferous and nonmetalliferous soils. *Plant Physiol* 169: 674-689.

Marhold K, Lihová J (2006) Polyploidy, hybridization and reticulate evolution: lessons from the *Brassicaceae*. *Plant Syst Evol* 259: 143-174.

Mercier R, Mézard C, Jenczewski E, Macaisne N, Grelon M (2015) The molecular biology of meiosis in plants. *Annu Rev Plant Biol* 66: 297-327.

Ming R, Hou S, Feng Y, Yu Q, Dionne-Laporte A, Saw JH, Senin P, Wang W, Ly BV, Lewis KL, Salzberg SL, Feng L, Jones MR, Skelton RL, Murray JE, Chen C, Qian W, Shen J, Du P, Eustice M, Tong E, Tang H, Lyons E, Paull RE, Michael TP, Wall K, Rice DW, Albert H, Wang ML, Zhu YJ, Schatz M, Nagarajan N, Acob RA, Guan P, Blas A, Wai CM, Ackerman CM, Ren Y, Liu C, Wang J, Wang J, Na JK, Shakirov EV, Haas B, Thimmapuram J, Nelson D, Wang X, Bowers JE, Gschwend AR, Delcher AL, Singh R, Suzuki JY, Tripathi S, Neupane K, Wei H, Irikura B, Paidi M, Jiang N, Zhang W, Presting G, Windsor A, Navajas-Pérez R, Torres MJ, Feltus FA, Porter B, Li Y, Burroughs AM, Luo MC, Liu L, Christopher DA, Mount SM, Moore PH, Sugimura T, Jiang J, Schuler MA, Friedman V, Mitchell-Olds T, Shippen DE, dePamphilis CW, Palmer JD, Freeling M, Paterson AH, Gonsalves D, Wang L, Alam M (2008) The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). *Nature* 452: 991-996.

Mozo T, Fischer S, Meier-Ewert S, Lehrach H, Altmann T (1998) Use of the IGF BAC library for physical mapping of the *Arabidopsis thaliana* genome. *Plant J* 16: 377-384.

Ouzounis CA (2005) Ancestral state reconstructions for genomes. *Curr Opin Genet Dev* 15: 595-600.

Özüdoğru B, Akaydın G, Erik S; Al-Shehbaz IA, Mummenhoff K (2015) Phylogeny, diversification and biogeographic implications of the eastern Mediterranean endemic genus *Ricotia* (Brassicaceae). *Taxon* 64: 727-740.

Parkin IAP, Gulden SM, Sharpe AG, Lukens L, Trick M, Osborn TC, Lydiate DJ (2005) Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. *Genetics* 171: 765-781.

Parkin IA, Koh C, Tang H, Robinson SJ, Kagale S, Clarke WE, Town CD, Nixon J, Krishnakumar V, Bidwell SL, Denoeud F, Belcram H, Links MG, Just J, Clarke C, Bender T, Huebert T, Mason AS, Pires JC, Barker G, Moore J, Walley PG, Manoli S, Batley J, Edwards D, Nelson MN, Wang X, Paterson AH, King G, Bancroft I, Chalhoub B, Sharpe AG (2014) Transcriptome and methylome profiling reveals relics of genome dominance in the *mesopolyploid Brassica oleracea*. *Genome Biol* 15: R77.

Pecinka A, Schubert V, Meister A, Kreth G, Klatte M, Lysak MA, Fuchs J, 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.

Rieseberg LH (2001) Chromosomal rearrangements and speciation. *Trends Ecol Evol* 16: 351-358.

Robertson WRB (1916) Chromosome studies. I. Taxonomic relationships shown in the chromosomes of Tettigidae and Acrididae. V-shaped chromosomes and their significance in Acrididae, Locustidae and Gryllidae: chromosome and variation. *J Morph* 27: 179-331.

Schranz ME, Lysak MA, Mitchell-Olds T (2006) The ABC's of comparative genomics in the *Brassicaceae*: building blocks of crucifer genomes. *Trends Plant Sci* 11: 535-542.

Schranz ME, Windsor AJ, Song B, Lawton-Rauh A, Mitchell-Olds T (2007b) Comparative genetic mapping in *Boechera stricta*, a close relative of Arabidopsis. *Plant Physiol* 144: 286-298.

Schubert I, Lysak MA (2011) Interpretation of karyotype evolution should consider chromosome structural constraints. *Trends Genet* 27: 207-216.

Slotte T, Hazzouri KM, Ågren JA, Koenig D, Maumus F, Guo YL, Steige K, Platts AE, Escobar JS, Newman LK, Wang W, Mandáková T, Vello E, Smith LM, Henz SR, Steffen J, Takuno S, Brandvain Y, Coop G, Andolfatto P, Hu TT, Blanchette M, Clark RM, Quesneville H, Nordborg M, Gaut BS, Lysak MA, Jenkins J, Grimwood J, Chapman J, Prochnik S, Shu S, Rokhsar D, Schmutz J, Weigel D, Wright SI (2013) The *Capsella rubella* genome and the genomic consequences of rapid mating system evolution. *Nat Genet* 45: 831-835.

Schubert V, Berr A, Meister A (2012) Interphase chromatin organisation in Arabidopsis nuclei: constraints versus randomness. *Chromosoma* 121: 369-387.

Schubert V, Rudnik R, Schubert I (2014) Chromatin associations in Arabidopsis interphase nuclei. *Front Genet* 5: 389.

Sharma A, Li X, Lim YP (2014) Comparative genomics of Brassicaceae crops. *Breed Sci* 64: 3-13.

Sharma S, Padmaja KL, Gupta V, Paritosh K, Pradhan AK, Pental D (2014) Two plastid DNA lineages - *Rapa*/*Oleracea* and *Nigra* - within the tribe Brassiceae can be best explained by reciprocal crosses at hexaploidy: evidence from divergence times of the plastid genomes and R-block genes of the A and B genomes of *Brassica juncea*. *PLoS One* 9: e93260.

Soltis DE, Albert VA, Leebens-Mack J, Bell CD, Paterson AH, Zheng C, Sankoff D, Depamphilis CW, Wall PK, Soltis PS (2009) Polyploidy and angiosperm diversification. *Am J Bot* 96: 336-348.

Soltis PS, Marchant DB, Van de Peer Y, Soltis DE (2015) Polyploidy and genome evolution in plants. *Curr Opin Genet Dev* 35: 119-125.

Stimpson KM, Matheny JE, Sullivan BA (2012) Dicentric chromosomes: unique models to study centromere function and inactivation. *Chrom Res* 20: 595-605.

Stockenhuber R, Zoller S, Shimizu-Inatsugi R, Gugerli F, Shimizu KK, Widmer A, Fischer MC (2015) Efficient detection of novel nuclear markers for Brassicaceae by transcriptome sequencing. *PLoS One* 10: e0128181.

Tang HB, Bowers JE, Wang XY, Ming R, Alam M, Paterson AH (2008) Perspective - Synteny and collinearity in plant genomes. *Science* 320: 486-488.

Tobgy HA (1943) A cytological study of *Crepis fuliginosa*, *C. neglecta*, and their F1 hybrid, and its bearing on the mechanism of phylogenetic reduction in chromosome number. *J Genet* 45: 67-111.

Vanneste K, Baele G, Maere S, Van de Peer Y (2014) Analysis of 41 plant genomes supports a wave of successful genome duplications in association with the Cretaceous-Paleogene boundary. *Genome Res* 24: 1334-1347.

Wang W, Huang S, Liu Y, Fang Z, Yang L, Hua W, Yuan S, Liu S, Sun J, Zhuang M, Zhang Y, Zeng A (2012) Construction and analysis of a high-density genetic linkage map in cabbage (*Brassica oleracea* L. var. *capitata*). *BMC Genomics* 13: 523.

Wang X, Wu J, Liang J, Cheng F, Wang X (2015) Brassica database (BRAD) version 2.0: integrating and mining Brassicaceae species genomic resources. *Database*: pii: bav093.

Wang X, Wang H, Wang J, Sun R, Wu J, Liu S, Bai Y, Mun JH, Bancroft I, Cheng F, Huang S, Li X, Hua W, Wang J, Wang X, Freeling M, Pires JC, Paterson AH, Chalhoub B, Wang B, Hayward A, Sharpe AG, Park BS, Weisshaar B, Liu B, Li B, Liu B, Tong C, Song C, Duran C, Peng C, Geng C, Koh C, Lin C, Edwards D, Mu D, Shen D, Soumpourou E, Li F, Fraser F, Conant G, Lassalle G, King GJ, Bonnema G, Tang H, Wang H, Belcram H, Zhou H, Hirakawa H, Abe H, Guo H, Wang H, Jin H, Parkin IA, Batley J, Kim JS, Just J, Li J, Xu J, Deng J, Kim JA, Li J, Yu J, Meng J, Wang J, Min J, Poulain J, Wang J, Hatakeyama K, Wu K, Wang L, Fang L, Trick M, Links MG, Zhao M, Jin M, Ramchiary N, Drou N, Berkman PJ, Cai Q, Huang Q, Li R, Tabata S, Cheng S, Zhang S, Zhang S, Huang S, Sato S, Sun S, Kwon SJ, Choi SR, Lee TH, Fan W, Zhao X, Tan X, Xu X, Wang Y, Qiu Y, Yin Y, Li Y, Du Y, Liao Y, Lim Y, Narusaka Y, Wang Y, Wang Z, Li Z, Wang Z, Xiong Z, Zhang Z; *Brassica rapa* Genome Sequencing Project Consortium (2011) The genome of the mesopolyploid crop species *Brassica rapa*. *Nat Genet* 43: 1035-1039.

Warwick SI, Al-Shehbaz IA (2006) *Brassicaceae*: Chromosome number index and database on CD-Rom. *Pl Syst Evol* 259: 237-248.

Willing E-M, Rawat V, Mandáková T, Maumus F, Velikkakam James G, Nordström KJV, Becker C, Warthmann N, Chica C, Szarzynska B, Zytnicki M, Albani MC, Kiefer C, Bergonzi S, Castaings L, Mateos JL, Berns MC, Bujdoso N, Piofczyk T, de Lorenzo L, Barrero-Sicilia C, Mateos I, Piednoël M, Hagmann J, Chen-Min-Tao R, Iglesias-Fernández R, Schuster SC, Alonso-Blanco C, Roudier F, Carbonero P, Paz-Ares J, Davis SJ, Pecinka A, Quesneville H, Colot V, Lysak MA, Weigel D, Coupland G, Schneeberger K (2015) Genome expansion of *Arabis alpina* linked with retrotransposition and reduced symmetric DNA methylation. *Nat Plants* 1: doi:10.1038/nplants.2014.23.

Wu HJ, Zhang Z, Wang JY, Oh DH, Dassanayake M, Liu B, Huang Q, Sun HX, Xia R, Wu Y, Wang YN, Yang Z, Liu Y, Zhang W, Zhang H, Chu J, Yan C, Fang S, Zhang J, Wang Y, Zhang F, Wang G, Lee SY, Cheeseman JM, Yang B, Li B, Min J, Yang L, Wang J, Chu C, Chen SY, Bohnert HJ, Zhu JK, Wang XJ, Xie Q (2012) Insights into salt tolerance from the genome of *Thellungiella salsuginea*. *Proc Natl Acad Sci USA* 109: 12219-12224.

Yang R, Jarvis DE, Chen H, Beilstein MA, Grimwood J, Jenkins J, Shu S, Prochnik S, Xin M, Ma C, Schmutz J, Wing RA, Mitchell-Olds T, Schumaker KS, Wang X (2013) The reference genome of the halophytic plant *Eutrema salsugineum*. *Front Plant Sci* 4: 46.

Yogeeswaran K, Frary A, York TL, Amenta A, Lesser AH, Nasrallah JB, Tanksley SD, Nasrallah ME (2005) Comparative genome analyses of *Arabidopsis* spp.: inferring chromosomal rearrangement events in the evolutionary history of *A. thaliana*. *Genome Res* 15: 505-515.

Yu J, Zhao M, Wang X, Tong C, Huang S, Tehrim S, Liu Y, Hua W, Liu S (2013) Bolbase: a comprehensive genomics database for *Brassica oleracea*. *BMC Genomics* 14: 664.

Zhang B, Lv Z, Pang J, Liu Y, Guo X, Fu S, Li J, Dong Q, Wu HJ, Gao Z, Wang XJ, Han F (2013) Formation of a functional maize centromere after loss of centromeric sequences and gain of ectopic sequences. *Plant Cell* 25: 1979-1989.

Ziolkowski PA, Kaczmarek M, Babula D, Sadowski J (2006) Genome evolution in Arabidopsis/*Brassica*: conservation and divergence of ancient rearranged segments and their breakpoints. *Plant J* 47: 63-74.

8. Applicant's publications cited

Cheng F, Mandáková T, Wu J, Xie Q, **Lysak MA**, Wang X (2013) Deciphering the diploid ancestral genome of the mesohexaploid *Brassica rapa*. *Plant Cell* 25: 1541-1554.

Fransz PF, de Jong JH, **Lysak M**, Ruffini Castiglione M, Schubert I (2002) Interphase chromosomes in *Arabidopsis* are organised as well-defined chromocenters from which euchromatin loops emanate. *Proceedings of the National Academy of Sciences USA* 99: 14584-14589.

Franzke A, **Lysak MA**, Al-Shehbaz IA, Koch MA, Mummenhoff K (2011) Cabbage family affairs: the evolutionary history of Brassicaceae. *Trends in Plant Science* 16: 108-116.

Geiser C, Mandáková T, Arrigo N, **Lysak MA**, Parisod C (2016) Repeated whole-genome duplication, karyotype reshuffling and biased retention of stress-responding genes in Buckler Mustards. *Plant Cell*, doi: http://dx.doi.org/10.1105/tpc.15.00791 (in press).

Haudry A, Platts AE, Vello E, Hoen D, Leclercq M, Williamson R, Joly-Lopez Z, Forczek E, Steffen J, Hazzouri K, Dewar K, Stinchcombe JR, Schoen DJ, Wang X, Schmutz J, Town CD, Edger PP, Pires JC, Schumaker KS, Jarvis DE, Mandáková T, **Lysak MA**, Schranz ME, van den Bergh E, Harrison P, Moses A, Bureau TE, Wright SI, Blanchette M (2013) An atlas of over 90,000 conserved noncoding sequences provides insight into crucifer regulatory regions. *Nature Genetics* 45: 891-898.

Hay AS, Pieper B, Cooke E, Mandáková T, Cartolano M, Tattersall AD, Ioio RD, McGowan SJ, Barkoulas M, Galinha C, Rast MI, Hofhuis H, Then C, Plieske J, Ganal M, Mott R, Martinez-Garcia JF, Carine MA, Scotland RW, Gan X, Filatov DA, **Lysak MA**, Tsiantis M (2014) *Cardamine hirsuta*: a versatile genetic system for comparative studies. *Plant Journal* 78: 1-15.

Hohmann N, Wolf E, **Lysak MA**, Koch MA (2015) A time-calibrated road map of Brassicaceae species radiation and evolutionary history. *Plant Cell* 27: 2770-2784.

Long Q, Rabanal FA, Meng D, Huber CD, Farlow A, Platzer A, Zhang Q, Vilhjalmsson BJ, Mandáková T, **Lysak MA**, Korte A, Nizhynska V, Voronin V, Korte P, Sedman L, Seren U, Hellmann I, Nordborg M (2013) Massive genomic variation and strong selection in *Arabidopsis thaliana* lines from Sweden. *Nature Genetics* 45: 884-890.

Lysak MA (2014) Live and let die: centromere loss during evolution of plant chromosomes. *New Phytologist* 203: 1082-1089.

Lysak MA, Lexer C (2006) Towards the era of comparative evolutionary genomics in Brassicaceae. *Plant Systematics and Evolution* 259: 175-198.

Lysak MA, Koch MA (2011) Phylogeny, genome and karyotype evolution of crucifers (Brassicaceae). In: Bancroft I, Schmidt R (eds.): Genetics and Genomics of the Brassicaceae. Springer, New York, pp. 1-31.

Lysak MA, Fransz PF, Ali HBM, Schubert I (2001) Chromosome painting in *Arabidopsis thaliana*. *Plant Journal* 28: 689-697.

Lysak MA, Pecinka A, Schubert I (2003) Recent progress in chromosome painting of Arabidopsis and related species. *Chromosome Research* 11: 195-204.

Lysak MA, Koch M, Pecinka A, Schubert I (2005) Chromosome triplication found across the tribe Brassiceae. *Genome Research* 15: 516-525.

Lysak MA, Berr A, Pecinka A, Schmidt R, McBreen K, Schubert I (2006) Mechanisms of chromosome number reduction in *Arabidopsis thaliana* and related Brassicaceae species. *Proceedings of the National Academy of Sciences USA* 103: 5224-5229.

Lysak MA, Cheung K, Kitschke M, Bures P (2007) Ancestral chromosomal blocks are triplicated in Brassiceae species with varying chromosome number and genome size. *Plant Physiology* 145: 402-410.

Lysak MA, Koch MA, Beaulieu JM, Meister A, Leitch IJ (2009) The dynamic ups and downs of genome size evolution in Brassicaceae. *Molecular Biology and Evolution* 26: 85-98.

Lysak MA, Mandáková T, Lacombe E (2010) Reciprocal and multi-species chromosome BAC painting in crucifers (Brassicaceae). *Cytogenetic and Genome Research* 129: 184-189.

Lysak MA, Mandáková T, Schranz EM (2016) Comparative paleogenomics of crucifers: ancestral genomic blocks revisited. *Current Opinion in Plant Biology* (in press).

Mandáková T, **Lysak MA** (2008) Chromosomal phylogeny and karyotype evolution in x=7 crucifer species (Brassicaceae). *Plant Cell* 20: 2559-2570.

Mandáková T, Joly S, Krzywinski M, Mummenhoff K, **Lysak MA** (2010) Fast diploidization in close mesopolyploid relatives of *Arabidopsis*. *Plant Cell* 22: 2277-2290.

Mandáková T, Heenan PB, **Lysak MA** (2010) Island species radiation and karyotypic stasis in *Pachycladon*. *BMC Evolutionary Biology* 10: 367.

Mandakova T, Mummenhoff K, Al-Shehbaz IA, Mucina L, Muehlhausee A, **Lysak MA** (2012) Whole-genome triplication and species radiation in the southern African tribe Heliophileae (Brassicaceae). *Taxon* 61: 989-1000.

Mandáková T, Shimizu Inatsugi R, Zozomová-Lihová J, Shimizu K, Kovařík A, Marhold K, **Lysak MA** (2013) The more the merrier: recent hybridization and polyploidy in *Cardamine*. *Plant Cell* 25: 3280-3295.

Mandáková T, Marhold K, **Lysak MA** (2014) The widespread crucifer species *Cardamine flexuosa* is an allotetraploid with a conserved subgenomic structure. *New Phytologist* 201: 982-992.

Mandáková T, Singh V, Kraemer U, **Lysak MA** (2015) Genome structure of the heavy metal hyperaccumulator *Noccaea caerulescens* and its stability on metalliferous and non-metalliferous soils. *Plant Physiology* 169: 674-689.

Mandáková T, Schranz ME, Sharbel TF, de Jong H, **Lysak MA** (2015) Karyotype evolution in apomictic *Boechera* and the origin of the aberrant chromosomes. *Plant Journal* 82: 785-793.

Mandáková T, Gloss A, Whiteman NK, **Lysak MA** (2016) How diplodization turned a tetraploid into a diploid. *American Journal of Botany* (in press).

Pecinka A, Schubert V, Meister A, Kreth G, Klatte M, **Lysak MA**, Fuchs J, 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.

Schranz EM, **Lysak MA**, Mitchell-Olds T (2006) The ABC's of comparative genomics in the Brassicaceae: building blocks of crucifer genomes. *Trends in Plant Science* 11: 535-542.

Schubert I, **Lysak MA** (2011) Interpretation of karyotype evolution should consider chromosome structural constraints. *Trends in Genetics* 27: 207-216.

Slotte T, Hazzouri KM, Ågren JA, Koenig D, Maumus F, Guo Y, Steige K, Platts AE, Escobar JS, Wang W, Mandáková T, Vello E, Steffen J, Takuno S, Brandvain Y, Coop G, Andolfatto P, Hu TT, Blanchette M, Clark RM, Quesneville H, Nordborg M, Gaut BS, **Lysak MA**, Jenkins J, Grimwood J, Prochnick S, Shu S, Rokhsar D, Schmutz J, Weigel D, Wright SI (2013) The *Capsella rubella* genome and the genomic causes and the genomic consequences of rapid mating system evolution. *Nature Genetics* 45: 831-835.

Tan EH, Henry IM, Ravi M, Bradnam KR, Mandakova T, Marimuthu MP, Korf I, **Lysak MA**, Comai L, Chan SW (2014) Catastrophic chromosomal restructuring during genome elimination in plants. *eLife* 4: e06516.

Willing E-M, Rawat V, Mandáková T, Maumus F, Velikkakam James G, Nordström KJV, Becker C, Warthmann N, Chica C, Szarzynska B, Zytnicki M, Albani MC, Kiefer C, Bergonzi S, Castaings L, Mateos JL, Berns MC, Bujdoso N, Piofczyk T, de Lorenzo L, Barrero-Sicilia C, Mateos I, Piednoël M, Hagmann J, Chen-Min-Tao R, Iglesias-Fernández R, Schuster SC, Alonso-Blanco C, Roudier F, Carbonero P, Paz-Ares J, Davis SJ, Pecinka A, Quesneville H, Colot V, **Lysak MA**, Weigel D, Coupland G, Schneeberger K (2015) Genome expansion of *Arabis alpina* linked with retrotransposition and reduced symmetric DNA methylation. *Nature Plants* 1: doi:10.1038/nplants.2014.23.