Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Plant Biology

Comparative paleogenomics of crucifers: ancestral genomic blocks revisited

Martin A Lysak[1], Terezie Mandáková[1] and M Eric Schranz[2]

A decade ago the concept of the Ancestral Crucifer Karyotype (ACK) and the definition of 24 conserved genomic blocks was presented. Subsequently, 35 cytogenetic reconstructions and/or draft genome sequences of crucifer species (members of the Brassicaceae family) have been analyzed in the context of this system; placing crucifers at the forefront of plant phylogenomics. In this review, we highlight how the ACK and genomic blocks have facilitated and guided genomic analysis of crucifers in the last 10 years and provide an update of this robust model.

Addresses
[1] Plant Cytogenomics Group, CEITEC – Central European Institute of Technology, Masaryk University, Kamenice 5, Brno CZ-62500, Czech Republic
[2] Biosystematics Group, Wageningen University (WU), Droevendaalsesteeg 1, Wageningen 6708 PB, The Netherlands

Corresponding author: Schranz, M Eric (eric.schranz@wur.nl)

Current Opinion in Plant Biology 2016, **30**:108–115

This review comes from a themed issue on **Genome studies and molecular genetics**

Edited by **Yves Van de Peer** and **J Chris Pires**

For a complete overview see the Issue and the Editorial

Available online 4th March 2016

http://dx.doi.org/10.1016/j.pbi.2016.02.001

Introduction

Ten years ago, we proposed a model for comparative genomic and chromosomal analysis for the mustard family (Brassicaceae or crucifers) based on the concept of the Ancestral Crucifer Karyotype (ACK) comprising eight chromosomes and 24 genomic blocks [1]. The system was developed from a rather limited mixed-set of completed genomes, genetic linkage maps and/or cytogenetic reconstructions for *Arabidopsis thaliana*, *A. lyrata*, *Capsella rubella*, three close-relatives of *Arabidopsis* (*Hornungia*, *Neslia* and *Turritis*), and *Brassica* species [3,12,40].

The last decade has been an extremely fruitful period for plant genomics and for crucifers in particular. There have been quantum leaps in whole-genome sequencing using next-generation sequencing (NGS) platforms [4]. *Arabidopsis thaliana* was the first plant genome sequenced [40], making the close relatives of Arabidopsis and other crucifer species obvious targets for comparative genomic and cytogenetic studies (reviewed by [5]). Furthermore, the Arabidopsis-based genetic and genomic resources, small genome sizes [6], and low proportion of repetitive DNA elements organized in discrete genome regions make crucifer genomes and chromosomes relatively straightforward to analyze and interpret.

The phylogenetic patterns within the family have also become clearer in the last decade, including the establishment of *Aethionema* as sister to the rest of the family (also known as the crown-group), the identification of three (Lineages I, II and III; [7]) and later five (A to E; [8•]) major lineages within the crown-group, and the delimitation of 49 tribes [9]. Also, the presence and phylogenetic placement of ancient whole-genome duplications and triplications (WGDs and WGTs) have been elucidated. For example, it is now known that *Aethionema* shares the At-alpha duplication with the crown-group [10], the *Brassica* genome triplication is shared among genera of the tribe Brassicaceae [12], and that there are several post-alpha mesopolyploid WGD/T events across the crown-group [6,11]. Here we aim to explore how the newly accumulated comparative cytogenomic and phylogenomic data impacted the ACK genome concept and our view on the architecture of ancestral and modern crucifer genomes.

The GB framework then and now

The system of 24 genomic blocks (GBs) across the eight ancestral chromosomes was founded on four significant preceding reports: (1) the sequenced genome of Arabidopsis; (2) the concept of eight ancestral chromosomes resembling the genomes of *A. lyrata* and *C. rubella* of Lineage I; (3) 21 segments of shared synteny between Arabidopsis and the 19 chromosomes of *B. napus* of Lineage II [3]; and (4) comparative cytogenetic data for five close Arabidopsis relatives [2]. The initial set of 21 syntenic segments from the Arabidopsis/*Brassica* comparison was refined and extended to 24 ancestral GBs labeled by capital letter from A to X [1].

The framework of the ACK and 24 GBs have become a crucial reference point for new genomic sequencing projects and karyotype reconstructions by comparative chromosome painting across the family and to date has been utilized for the analysis of at least 35 crucifer species (Table 1). The unprecedented advance in our knowledge of genome evolution in the Brassicaceae allows us to now revisit the original concept of the Ancestral Crucifer Karyotype and 24 GBs. Data on the 35 species corroborate

Table 1

Presence/absence of ancestral genomic block associations in 35 crucifer species

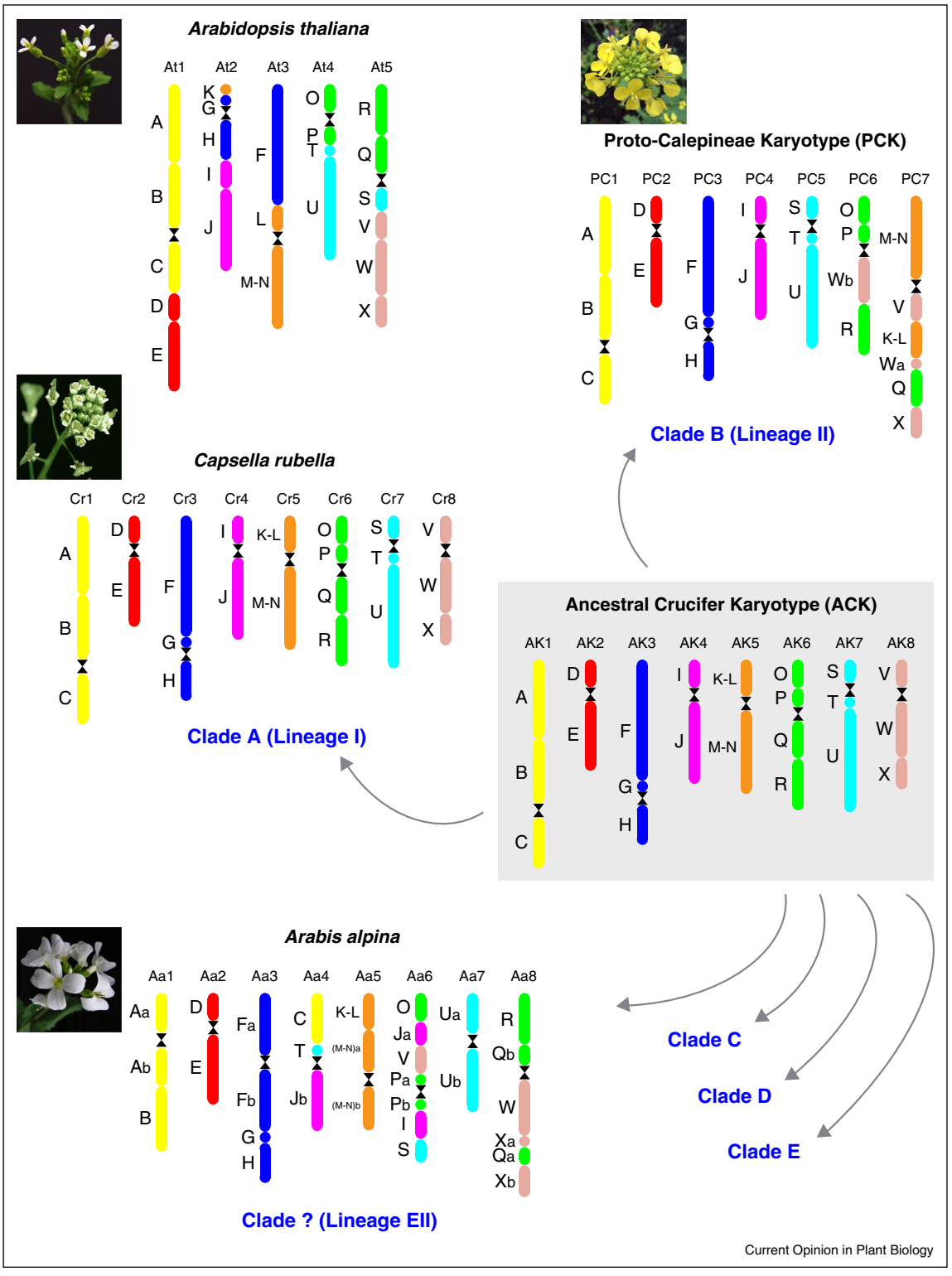
Species	Tribe clade[a] (lineage[b])	Ancestral genomic block association										Reference
		A–B	F–G	K–L	M–N	O–P	Q–R	T–U	W–X	Wb–R[c]	V–K–L–Wa–Q–X[c]	
ACK ancestor												
Arabidopsis lyrata	Camelineae A (I)	+	+	+	+	+	+	+	+	–	–	[2]
Arabidopsis thaliana	Camelineae A (I)	+	–	–	+	+	+	+	+	–	–	[1]
Ballantinia antipoda[d]	Microlepidieae A (I)	+	+	+	+	–	+	–	+	–	–	[23]
Boechera divaricarpa	Boechereae A (I)	+	+	+	+	+	+	+	+	–	–	[27,34]
Boechera stricta	Boechereae A (I)	+	+	+	+	+	+	+	+	–	–	[34]
Camelina sativa[e]	Camelineae A (I)	+	+	+	+	+	+	+	+	–	–	[25••]
Cardamine amara	Cardamineae A (I)	+	+	+	+	+	+	+	+	–	–	[26]
Cardamine flexuosa[d]	Cardamineae A (I)	+	+	+	+	+	+	+	+	–	–	[35]
Cardamine hirsuta	Cardamineae A (I)	+	+	+	+	+	+	+	+	–	–	[36]
Capsella rubella	Camelineae A (I)	+	+	+	+	+	+	+	+	–	–	[33]
Crucihimalaya wallichii	Crucihimalayeae A (I)	+	+	+	+	+	+	+	+	–	–	[23]
Hornungia alpina	Descurainieae A (I)	+	+	+	+	+	+	+	+	–	–	[2]
Neslia paniculata	Camelineae A (I)	+	+	+	+	+	+	+	+	–	–	[2]
Pachycladon exilis[d]	Microlepidieae A (I)	+	+	+	+	+	+	+	+	–	–	[37]
Stenopetalum lineare[d]	Microlepidieae A (I)	–	+	+	+	–	–	+	+	–	–	[23]
Stenopetalum nutans[d]	Microlepidieae A (I)	–	+	+	+	–	–	–	+	–	–	[23]
Transberingia bursifolia	Crucihimalayeae A (I)	+	+	+	+	+	+	+	+	–	–	[23]
Turritis glabra	Turritideae A (I)	+	+	+	+	+	+	+	+	–	–	[2]
PCK ancestor												
Brassica oleracea[e]	Brassiceae B (II)	–	–	+	–	+	–	–	–	+	+	[18]
Brassica rapa[e]	Brassiceae B (II)	–	–	+	–	+	–	–	–	+	+	[15•,18,22]
Brassica napus[d+e]	Brassiceae B (II)	–	–	+	–	+	–	–	–	+	+	[18]
Calepina irregularis	Calepineae B (EII)	+	+	+	+	+	–	+	–	+	+	[14]
Conringia orientalis	Conringieae B (EII)	+	+	+	+	+	–	+	–	+	+	[14]
Caulanthus amplexicaulis[d]	Thelypodieae? (EII)	+	–	?	+	–	–	+	–	+	+	[38]
Glastaria glastifolia	Isatideae B (II)	+	+	+	+	+	–	+	–	+	+	[14]
Goldbachia laevigata[d]	Calepineae B (EII)	+	+	+	+	–	–	+	–	+	+	[14]
Myagrum perfoliatum	Isatideae B (II)	+	+	+	+	+	–	+	–	+	+	[14]
Noccaea caerulescens	Coluteocarpeae B (EII)	+	+	+	+	–	–	–	–	+	+	[28]
Noccaea jankae[d]	Coluteocarpeae B (EII)	+	+	+	+	+	–	–	–	+	+	[28]
Ochthodium aegyptiacum	Sisymbrieae B (II)	+	+	+	+	+	–	+	–	+	+	[14]
Raparia bulbosa[d]	Coluteocarpeae B (EII)	+	+	+	+	+	–	–	–	+	+	[28]
Schrenkiella parvula	Unassigned B (EII)	+	+	+	+	+	–	+	–	+	+	[14]
Thellungiella salsuginea	Eutremeae B (EII)	+	+	+	+	+	–	+	–	+	+	[39]
Unresolved ancestral genome												
Arabis alpina	Arabideae ? (EII)	+	+	+	+	–	+	–	+	–	–	[31••]
Biscutella laevigata[d]	Biscutelleae C (?)	+	+	+	+	+	–	?	–	+	–	[24]

[a] According to [8•].
[b] According to [13].
[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 WGD.

the original idea of an ancestral genome with eight chromosomes — the ACK. An important new discovery was identification of a second most common ancestral genome, particularly in taxa belonging to expanded Lineage II [13] or Clade B [8•], with seven chromosomes and called the Proto-Calepineae Karyotype (PCK, [14]). It is more parsimonious to assume that the $n = 7$ PCK genome has descended from the older $n = 8$ (ACK) genome (Figure 1).

After the introduction of the PCK and its evolutionarily younger variant tPCK [14,15•], insights into *Brassica* evolution have been extended with genome publications for *B. rapa* [16], *B. oleracea* [17,18] and *B. napus* [19••]. These assemblies, and previously published data, reveal that all three sequenced *Brassica* genomes contain three or six (in *B. napus*) sub-genomes as the consequence of a shared or very similar whole-genome triplication event(s) (e.g., [3,12,20,21]). Newly available genomic data also allowed for a more accurate reconstruction of *Brassica* proto-genome structures before the WGT. Using the system of 24 genomic blocks, Cheng *et al.* [15•] reconstructed the evolution of the mesohexaploid genome of *B. rapa* as two hybridization events between three structurally identical or very similar tPCK genomes. Sequencing of triplicated and extensively shuffled genomic blocks

Figure 1



The revised Ancestral Crucifer Karyotype (ACK) comprising 22 genomic blocks and their position within the ancestral genome of PCK (Clade B/Lineage II) and genomes of *Arabidopsis thaliana*, *Capsella rubella* (both Clade A/Lineage I) and *Arabis alpina* (unresolved clade/expanded Lineage II). Note the shuffling of GBs associated with descending dysploidy events (PCK and *Arabidopsis*) and independent of dysploidy (*A. alpina*). Some rearrangements, including centromere repositioning (*A. alpina*), might result into partition of GBs to sub-blocks labeled by lower-case letters. Ancestral karyotypes for clades C, D, and E are yet unknown. Black sandglass-like symbols represent centromeres. Clades and lineages refer to the published family-wide phylogenies [8•,13]. The figure is based on genomic data adopted from [1], [14,15•,33] and [31••]; see Table 2 for the definition of the 22 GBs.

building up the *Brassica* genomes provided further support and refinement of the genomic blocks ([15•,22] and this paper).

The known genome structure of numerous crucifer species provides us with the opportunity to re-appraise whether all the 24 GBs have evolutionary significance as ancestral building blocks. There are eight paired-associations of GBs, each on a different chromosome arm within ACK that potentially could be treated as one unit rather than two adjacent blocks. To be functionally and evolutionary meaningful we considered only GBs forming a pair-wise ancestral association in ACK but which are separated from each other in at least two different genera. We thus evaluated the contiguity of paired associations using the genomic and cytogenetic results from the available 35 species (Table 1). From the 24 originally defined GBs [1], blocks K and L are separated only in Arabidopsis, whereas the two GBs are associated on a single chromosome arm in all other genomes. Similarly, blocks M and N are separated only in *Brassica* species, but positioned together on a single chromosome arm in all other species. We thus update block intervals based on recently published data and we merge the four GBs to form two new blocks (K–L and M–N) into a revised ACK genome containing 22 ancestral GBs (Table 2). In the revised system of the 22 GBs, 10 blocks span entire chromosome arms, whereas the remaining 12 GBs are associated-pairs on six different arms (A + B, F + G, O + P, Q + R, T + U, and W + X). Figure 1 shows the updated ancestral genomes (ACK and PCK) and two examples of modern crucifer genomes (Arabidopsis and *Arabis alpina*).

Frequent polyploidy, diploidization and block shuffling

In Brassicaceae, multiple independent WGD and WGT events have occurred during their roughly 40 million years of evolution (e.g., [6,10,11,23]). WGDs and WGTs give rise to genomic redundancy and an increased probability for recombination between homeologous chromosomes, presumably mainly between ubiquitous DNA repeats. Eventually this process of diploidization leads to pervasive reorganization of duplicated genomic blocks,

Table 2

Revisited 22 genomic blocks (GB) building up the Ancestral Crucifer Karyotype (ACK). The GB boundaries are defined by *A. thaliana* gene loci (AtXg…) and BAC clones retrieved from TAIR. Note that BAC intervals refer to painting probes used for comparative cytogenetics in crucifer taxa and can be shorter than the gene-based intervals. AK1–8: chromosomes of the ACK; At1–5: chromosomes of *A. thaliana*. The order, orientation, and color-coding of each block is based on its position in the ACK. In graphic representations (ideograms) of the ACK (Figure 1), each GB is considered to be in the upright orientation with start border on the top of a GB. This version of crucifer GB framework is based on data from Schranz *et al.* [1], Cheng *et al.* [15•] and Kim *et al.* [22]

AK	At	GB	Block borders			
			Start		End	
			Gene	BAC clone (GenBank Accession)	Gene	BAC clone (GenBank Accession)
AK1	At1	A	At1g01010	T25K16 (AC007323)	At1g19840	F14P1 (AC024609)
	At1	B	At1g19850	F6F9 (AC007797)	At1g37130	F12K21 (AC023279)
	At1	C	At1g43020	F2J6 (AC009526)	At1g56190	T6H22 (AC009894)
AK2	At1	D	At1g64670	F1N19 (AC009519)	At1g56210	F14G9 (AC069159)
	At1	E	At1g64960	F13O11 (AC006193)	At1g80950	F23A5 (AC011713)
AK3	At3	F	At3g01015	T4P13 (AC008261)	At3g25520	MWL2 (AB025639)
	At2	G	At2g05170	F16J10 (AC007289)	At2g07690	T25N22 (AC005693)
	At2	H	At2g10940	T10F5 (AC007063)	At2g20900	F5H14 (AC006234)
AK4	At2	I	At2g20920	F7O24 (AC007142)	At2g31035	T19L18 (AC004747)
	At2	J	At2g31040	T9J22 (AC002505)	At2g48150	T8I13 (AC002337)
AK5	At2	K-L	At2g01060	F2I9 (AC005560)	At2g05160	F3C11 (AC007167)
	At3	K-L	At3g25540	T5M7(AP001313)	At3g32960	T4A2 (AP002066)
	At3	M-N	At3g42180	T10D17 (AL353865)	At3g63530	F16M2 (AL138648)
AK6	At4	O	At4g00026	F6N15 (AF069299)	At4g05450	T1J1 (AF128393)
	At4	P	At4g12620	T1P17 (AL049730)	At4g07390	T3H13 (AF128396)
	At5	Q	At5g30510	T8M17 (AF296835)	At5g23010	T20O7 (AB026660)
	At5	R	At5g23000	MRN17 (AB005243)	At5g01010	F7J8 (AL137189)
AK7	At5	S	At5g42110	MPK23 (AB020748)	At5g32470	F5H8 (AB025605)
	At4	T	At4g12700	T20K18 (AL049640)	At4g16240	F18A5 (AL035528)
	At4	U	At4g16250	T6K21 (AL021889)	At4g40100	T5J17 (AL035708)
AK8	At5	V	At5g47810	MGC1 (AB028612)	At5g42130	MJC20 (AB017067)
	At5	W	At5g47820	K16F13 (AB024025)	At5g60800	MUP24 (AB005246)
	At5	X	At5g60805	MSL3 (AB008269)	At5g67640	K9I9 (AB013390)

including their potential loss, and to reduction of chromosome numbers to diploid-like levels. The best-characterized example of this process is the diploidized Arabidopsis genome, reduced to only five linkage groups after the At-alpha duplication [40]. However, the most extensive chromosomal diploidization so far revealed is for the Australian endemic species *Stenopetalum nutans* (from $n = 16$ to 4; [23]), followed by the 2.1-fold decrease in *Brassica* (from $n = 21$ to 10; [15•,16]) or 1.7-fold reduction of chromosome number in *Biscutella leavigata* (from $n = 16$ to 9; [24]). By contrast, three sub-genomes within the mesohexaploid genome of *Camelina sativa* ($n = 20$) remained structurally intact since the WGT event [25••]. Accordingly, polyploidy-driven multiplication of genomic blocks does not always result in an extensive and rapid loss of duplicated GBs. In the *B. rapa* genome, all but one block were detected as three, though sometimes fractionated, genomic copies and only block G as two copies [15•,16]. Among the mesotetraploid Australian crucifers, the most reduced genome of *S. nutans* ($n = 4$) retained all 48 GBs, whereas the other two, less diploidized, species with $n = 5$ and $n = 6$ chromosomes lost four and eight GBs, respectively [23]. Further research is needed to uncover key drivers underlaying the differential rate of diploidization, including the retention and loss of duplicated genomic blocks.

Sequence analysis of mesopolyploid genomes undergoing the post-polyploidy process of diploidization is providing us with invaluable information on the mobility of genomic blocks, their sequence characteristics as well as on mechanisms of genomic reshuffling. The detailed structures of mesopolyploid genomes also allows for continuous re-appraisal and refinement of genomic blocks and the ACK concept.

Breaks in within genomic blocks

Reconstructing the evolution of modern crucifer genomes, three basic trajectories can be observed (Figure 1). Ancestral genomes (i.e. ACK and PCK) remained stable throughout million of years (*A. lyrata*, *C. rubella*) or have undergone extensive shuffling of genomic blocks without changing the ancestral number of chromosomes (*A. alpina*), or have been modified through descending dysploidy (chromosome number reduction) from $n = 8$ to 7, 6 and 5 (*A. thaliana*). The two latter patterns of karyotype evolution were accompanied not only by the block-defining breakpoints, but also by rarer in-block breaks (see PCK and *A. arabis* genome structures in Figure 1). Factors triggering in-block breakages remain elusive, ranging in diploid genomes from one in-block break in *Cardamine amara* ($n = 8$; [26]) to at least eight breaks in *A. alpina* ($n = 8$; Figure 1). Due to the redundancy of duplicated or triplicated GBs followed by diploidization, in-block breaks should be more frequent in dynamic mesopolyploid genomes. Indeed, this is the case in the mesohexaploid genome of *B. rapa* ($n = 10$; [15•,16]) and *C. sativa* [25••] with approx. 26 and 21 in-block breakpoints, respectively. However, in *C. sativa* the 21 in-block breaks occurred presumably already during the evolution of the three diploid progenitor genomes and not as a result of diploidization [25••]. The 4-fold reduction of chromosome number in the mesotetraploid genome of *S. nutans* [23] was accompanied by only nine in-block breakpoints and only two such breaks were observed during diploidization of the mesotetraploid *B. laevigata* ($n = 9$; [24]). We conclude, that most chromosomal rearrangements (CRs) involve genomic block breakpoints and that the frequency of rarer in-block breakpoints varies both among diploids and mesopolyploids.

Evolutionary mobility of genomic blocks: mechanisms

Ectopic recombination between tandem repeats (such as of rDNA arrays of the Nucleolar Organizer Regions) and transposable elements frequently mediates CRs. Therefore the very existence of GBs is to a lesser or greater extent predetermined by the arrays of repetitive DNA. Obvious repeat-rich borders of GBs are pericentromeres and telomeres, less common are repeats positioned interstitially along euchromatic chromosome arms. Non-allelic recombination between these repeats is resulting in chromosomal reshuffling with or without associated descending dysploidy. Inversions and translocations are most frequent types of CRs in crucifers. Both peri- and paracentric inversions, changing mutual orientation of GBs, can occur without the reduction of chromosome number [27,28], or as a first step towards terminal chromosome translocations (TCTs) mediating descending dysploidy [1,2,29].

Two main types of TCTs were common during evolution of the modern crucifer genomes. A whole-arm (Robertsonian) translocation between two telo- or acrocentric chromosomes may result in two unequally sized products: a larger mono- or dicentric translocation chromosome and a mini-chromosome [1,2,30]. The mini-chromosome, without essential genes, fails to pair regularly at meiosis, and as such is eliminated. TCTs may also involve (sub)-telomeric regions of two different chromosomes. This event results in the origin of a translocation chromosome comprising two centromeres. To stabilize a dicentric chromosome, one of the centromeres is presumably epigenetically inactivated and/or deleted through intra- or inter-chromosomal recombination (e.g., [23,29]).

In a few instances, centromere repositioning is the most parsimonious explanation for centromere relocations with ancestral chromosome collinearity being conserved. An evolutionary new centromere may be moved to block boundary [26] or into a position within a block, as in *A. alpina* (Figure 1; [31••]).

Future prospects

While the last ten-years has seen major advances in the study of crucifer genomes, the future holds even more promise. Several ambitious genomic projects like the Brassicaceae Map Alignment Project (BMAP) will surely double, if not triple, the number of available sequenced genomes within the next few years. Technological innovations are bringing the divide between molecular biology, genome sequencing and cyogenetics closer together. The emergence of so called 'third generation' long-molecule sequencing such as PacBio and Oxford NanoPore technology are setting the stage for complete genome assemblies including repeat-rich regions such as pericentromeres where CRs often occur. Cytogenetic analysis is also now possible using short custom-synthesized oligonucleotides so that only genes are hybridized and thus avoiding difficult signal to noise ratios from BAC-based painting caused by repeats [32•]. Thus, it will be possible to perform comparative cytogenetic experiments using data from any or all crucifer genomes and no longer be limited to the use of Arabidopsis BAC contigs. Alternatively, lineage-specific expansion of repetitive elements can posses phylogenetically informative signals and can be correlated with CRs [41]. Other exciting developments include the use of genome mapping technologies such as the BioNano Genomics platform 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. It will be possible to resolve rearrangements down to the nucleotide level, even when changes have occurred in highly repetitive regions. For example, it should be feasible to address the exact nature of rearrangements in polyploid genomes.

An important unresolved question is the genomic structure of *Aethionema* species, the first-branching Brassicaceae lineage, and that of Lineage III (mainly Clade E species from [8•]). All data to date supports the concept of the $n = 8$ ACK genome for the crown-group Brassicaceae in Lineages I and II (and unplaced species such as *A. alpina*). *Aethionema* in particular is expected to have undergone a largely independent trajectory of genome evolution because it diverged from the crown-group very soon after the At-alpha paleopolyploidy event [10]. By extending even further and making connections to more distant species such as members of the sister-family Cleomaceae, it is also feasible to do an ancestral genome reconstruction of the pre-At-alpha non-polyploid genome of crucifers. Such a resource will be extremely valuable for translational research from the Arabidopsis model to other angiosperm families.

Acknowledgements

MAL and TM were supported by a research grant from the Czech Science Foundation (grant no. P501/12/G090). MES was supported by the Netherlands Organization for Scientific Research (NWO VIDI Grant 864.10.001).

References and recommended reading

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

- • of special interest
- •• of outstanding interest

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

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

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

4. Thudi M, Li Y, Jackson SA, May GD, Varshney RK: **Current state-of-art of sequencing technologies for plant genomics research**. *Brief Funct Genomics* 2012, **11**:3-11.

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

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

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

8. • 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: **Resolution of Brassicaceae phylogeny using nuclear genes uncovers nested radiations and supports convergent morphological evolution**. *Mol Biol Evol* 2016, **33**:394-412.
A family-wide phylogenetic study based on nuclear markers retrieved from newly sequenced transcriptomes of 55 Brassicaceae species. Six major clades were resolved.

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

10. Schranz ME, Mohammadin S, Edger PP: **Ancient whole genome duplications, novelty and diversification: the WGD Radiation Lag-Time Model**. *Curr Opin Plant Biol* 2012, **15**:147-153.

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

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

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

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

15. • Cheng F, Mandáková T, Wu J, Xie Q, Lysak MA, Wang X: **Deciphering the diploid ancestral genome of the mesohexaploid *Brassica rapa***. *Plant Cell* 2013, **25**:1541-1554.
Parsimonious reconstruction of three ancestral genomes, shuffling of triplicated genomic blocks and fate of lost centromeres in the mesohexaploid *Brassica rapa*.

16. 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: **The genome of the mesopolyploid crop species *Brassica rapa***. *Nat Genet* 2011, **43**:1035-1039.

17. 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: **The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes**. *Nat Commun* 2014, **23**:3930.

18. 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: **Transcriptome and methylome profiling reveals relics of genome dominance in the mesopolyploid *Brassica oleracea***. *Genome Biol* 2014, **15**:R77.

19. •• 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: **Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome**. *Science* 2014, **345**:950-953.

Sequence analysis of the recent and ancient polyploidization events in the *Brassica* lineage. The work demonstrates the role and extent of homeologous exchanges between the A and C sub-genomes.

20. Arias T, Beilstein MA, Tang M, McKain MR, Pires JC: **Diversification times among *Brassica* (Brassicaceae) crops suggest hybrid formation after 20 million years of divergence**. *Am J Bot* 2014, **101**:86-91.

21. Sharma S, Padmaja KL, Gupta V, Paritosh K, Pradhan AK *et al.*: **Two plastid DNA lineages — Rapa/Oleracea and Nigra — within the tribe Brassicaceae 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* 2014, **9**:e09326.

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

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

24. Geiser C, Mandáková T, Arrigo N, Lysak MA, Parisod C: **Repeated whole-genome duplication, karyotype reshuffling and biased retention of stress-responding genes in Buckler Mustard**. *Plant Cell* 2016, **28**:17-27.

25. •• 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: **The emerging biofuel crop *Camelina sativa* retains a highly undifferentiated hexaploid genome structure**. *Nat Commun* 2014, **23**:3706.

Sequence analysis of the hexaploid genome of *C. sativa* provides a valuable analogy to the mesohexaploid genomes of *Brassica* species.

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

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

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

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

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

31. •• 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: **Genome expansion of *Arabis alpina* linked with retrotransposition and reduced symmetric DNA methylation**. *Nat Plants* 2015, **1** http://dx.doi.org/10.1038/nplants.2014.23.

The sequenced genome of *Arabis alpina* represents as-yet unknown evolutionary trajectory of the ACK genome, characterized by centromere repositioning and neocentromere formation.

32. • Han Y, Zhang T, Thammapichai P, Weng Y, Jiang J: **Chromosome-specific painting in *Cucumis* species using bulked oligonucleotides**. *Genetics* 2015, **200**:771-779.

Study showing the use of bulked short custom-synthesized oligonucleotides for chromosome-specific painting in plant species with a sequenced genome.

33. 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: **The *Capsella rubella* genome and the genomic consequences of rapid mating system evolution**. *Nat Genet* 2013, **45**:831-835.

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

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

36. 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: ***Cardamine hirsuta*: a versatile genetic system for comparative studies**. *Plant J* 2014, **78**:1-15.

37. Mandáková T, Heenan PB, Lysak MA: **Island species radiation and karyotypic stasis in *Pachycladon* allopolyploids**. *BMC Evol Biol* 2010, **10**:367.

38. Burrell AM, Taylor KG, Williams RJ, Cantrell RT, Menz MA, Pepper AE: **A comparative genomic map for *Caulanthus***

***amplexicaulis* and related species (Brassicaceae)**. *Mol Ecol* 2011, **20**:784-798.

39. 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: **Insights into salt tolerance from the genome of *Thellungiella salsuginea***. *Proc Natl Acad Sci U S A* 2012, **109**:12219-12224.

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

41. Zhang Y, Cheng C, Li J, Yang S, Wang Y, Li Z, Chen J, Lou Q: **Chromosomal structures and repetitive sequences divergence in Cucumis species revealed by comparative cytogenetic mapping**. *BMC Genomics* 2015, **16**:730.