DR CHRISTIAN PARISOD (Orcid ID : 0000-0001-8798-0897)

Article type : - Research Review

Research review

Chromosome restructuring among hybridizing wild wheats

Christian Parisod (orcid.org/0000-0001-8798-0897)[1*] and Ekaterina D. Badaeva (orcid.org/0000-0001-7101-9639)[2]

[1]Institute of Plant Sciences, University of Bern, Altenbergrain 21, 3013 Bern, Switzerland

[2]Vavilov Institute of General Genetics, Russian Academy of Sciences, Gubkin St. 3, Moscow, Russia, 119991.

[*]Corresponding author: Institute of Plant Sciences, Altenbergrain 21, University of Bern, 3013 Bern, Switzerland, Phone : +41 31 631 4949, e-mail: christian.parisod@ips.unibe.ch

Received: *25 October 2019*

Accepted: *29 December 2019*

Keywords: *Aegilops-Triticum* (wild wheats), chromosome evolution, hybrid speciation, genome dynamics, repeated sequences, transposable elements

Summary

The wheat group offers an outstanding system to address the interplay between hybridization, chromosomal evolution and biological diversification. Most diploid wild wheats originated following hybridization between the A-genome lineage and the B-genome lineage some 4MY ago, resulting in an admixed D-genome lineage that presented dramatic radiation accompanied by considerable changes in genome size and chromosomal rearrangements. Comparative profiling of low-copy genes, repeated sequences and transposable elements among those divergent species characterized by different karyotypes highlights high genome dynamics and shed new light on processes underlying chromosomal evolution in wild wheats. One of the hybrid clades presents upsizing of metacentric chromosomes going along with the proliferation of specific repeats (i.e. “genomic obesity”), whereas other species show stable genome size associated with increasing chromosomal asymmetry. Genetic and ecological variation in those specialized species suggest that genome restructuring was coupled with adaptive processes to support the evolution of a majority of acrocentric chromosomes. This synthesis of current knowledge on genome restructuring across the diversity of wild wheats paves the way towards surveys based on latest sequencing technologies to characterize valuable resources and address the significance of chromosomal evolution in species with complex genomes.

Introduction

The wheat group (*Aegilops* L. and *Triticum* L.) includes a dozen of annual diploid species that naturally occur across the Mediterranean region and Central Asia (van Slageren, 1994; Kilian *et al.*, 2011). Besides genetic resources of great interest for food safety, wild wheats represent an outstanding model system to address the interplay between hybridization, chromosomal evolution and biological diversification. As comprehensively summarized by Feldman and Levy (2015), most domesticated wheats such as *Triticum turgidum* (genome BBAA) and *T. aestivum* (BBAADD) are indeed of hybrid origin, having evolved after the merging and duplication of chromosomes (i.e. allopolyploidy) from genetically divergent diploid species: *Ae. speltoides* (genome SS=BB), *T. urartu* (AA), and *Ae. tauschii* (DD). Besides allopolyploid wheats and their reticulate evolution, recent phylogenetic works demonstrated that homoploid hybridization preceded the radiation of most diploid wild wheat species (discussed in Huynh *et al.*, 2019). As will be addressed below, although all wild wheats present x = 7, comparative cytogenetics highlights considerable chromosomal rearrangements within and among diploid species (Badaeva *et al.*, 2007) that also show so-called "complex genomes" riffled with more than 80% of transposable elements (Tenaillon *et al.*, 2010; Senerchia *et al.*, 2013). After refined relationships among wild wheats are outlined to provide a firm comparative framework, we integrate available knowledge on their chromosomal evolution and address how genome restructuring supported the hybrid origin and the dramatic radiation of most diploid wild wheats. Finally, we discuss possibly underlying molecular and evolutionary processes, and revisit classical hypotheses on how genome dynamics may promote trajectories towards either "genomic obesity" (Bennetzen and Kellogg 1997) or adaptive "karyotype asymmetry" (Stebbins 1971).

Reticulate evolution of diploid wild wheats

The diversity of diploid wild wheat species has long been studied (Feldman & Levy, 2015). However, species phylogenetic relationships have remained difficult to disentangle until increasingly representative sets of species and plastid as well as nuclear loci were coupled with sophisticated inferential tools (Marcussen, 2014; Bernhardt *et al.*, 2017; Bernhardt *et al.*, 2019; Glémin *et al.*, 2019;

Huynh *et al.*, 2019). Besides delimitation of a clearly monophyletic *Aegilops-Triticum* clade, sister to the genus *Taeniatherum*, these studies convincingly demonstrated that diploid wild wheats have evolved through homoploid hybridization (Fig. 1).

Multispecies coalescent analyses that accounted for stochastic processes among multiple low-copy genes have highlighted a 4 to 5 MY-old hybridization event between diverging taxa of the A-genome lineage (i.e. ancestral *Triticum* diploids) and of the B-genome lineage (*Ae. speltoides* and *Ae. mutica*) at the origin of an admixed D-genome lineage that then diversified into the majority of current wild wheat species (Huynh *et al.*, 2019). As discussed in this study, other phylogenies have similarly highlighted reticulate evolution, but emphasized on "successively nested hybridization events" and suggested that *Ae. mutica* may have been the B-progenitor of the D-genome lineage (Bernhardt *et al*., 2019; Glémin *et al.*, 2019). Despite morphological distinctiveness used to justify the classification of *Ae. mutica* in different genus (van Slageren, 1994; Kilian *et al.*, 2011), this species has been demonstrated as closely related to *Ae. speltoides* and to have recently captured the plastid genome of *Ae. umbellulata* (ca. 0.4 MY ago; Huynh et al. 2019). The sister *Ae. speltoides* and *Ae mutica* indeed are the only outcrossing wild wheats that also distinctively possess B-chromosomes (Ohta, 1991) and present very similar karyotype structure showing similar major 45S rDNA loci and large clusters of GTT_n microsatellite sequences in pericentromeric regions (Badaeva et al. 1996b; Ruban & Badaeva 2018). Accordingly, phylogenies based on either full-parameterized multispecies networks *vs* topology or SNP approaches present rather shallow discrepancies due to ancestral population structure (see discussion in Huynh *et al*. 2019). They support an ancestral B-genome progenitor of the D-genome lineage related to both *Ae. mutica* and *Ae. speltoides* or sister to them, that likely shared their main chromosomal characteristics. Comparative analysis of karyotypes from species of the A-, B- and D-genome lineages is used below to shed further light on the events underlying the origin and evolutionary radiation of hybrid wild wheats.

Following Huynh *et al.* (2019), the hybrid D-genome ancestor underwent substantial radiation from ~3 MY onwards and yielded the majority of current *Aegilops* species that were previously assigned with different genomic formulas and classified in distinct sections of the genus (van Slageren, 1994; Kilian *et al.*, 2011). Despite uncertainties regarding the intensity and tempo of gene flow among wild

wheats, diploid species are reproductively well isolated and present considerable morphological, chromosomal and ecological differences (Kilian *et al.*, 2011). As shown in Fig. 1, phylogenetic evidence supports the divergence of three main clades early after hybridization at the origin of the D-genome lineage. Basal splits of hybrid ancestors indeed yielded (i) a monospecific *tauschii* clade consisting of *Ae. tauschii*, (ii) a homogeneous S* clade comprising species of the section *Sitopsis* (i.e. *Ae. bicornis*, *Ae. longissima*, *Ae. searsii*, *Ae. sharonensis*), and (iii) a diverse CUMN clade formed of remaining diploid species *Ae. caudata* (C genome), *Ae. umbellulata* (U), *Ae. comosa* (M) and *Ae. uniaristata* (N). As will be discussed below, such refined comparative framework highlights contrasted evolutionary trajectories towards either the expansion of metacentric chromosomes in S* species (i.e. "genomic obesity") or a majority of acrocentric chromosomes in CUMN species (i.e. "karyotype asymmetry"), offering a unique opportunity to address processes underlying genome restructuring and evolutionary radiation.

Hybrid speciation and chromosome restructuring in wild wheats of the D-genome lineage

Hybrid wild wheats of the D-genome lineage radiated with a constant chromosome number, although considerable karyotype reorganization is apparent through significant variation in chromosome sizes as well as positions of centromeres, heterochromatic and satellite regions (Fig. 1). The early evolution of hybrids between A- and B-genome species is nowadays largely confounded with the later divergence of all species. However, the basal *Ae. tauschii* presents metacentric chromosomes similar to those of all progenitor species from the A- and B-genome lineages, and it thus offers informative comparisons to identify events having early shaped D-genome species following hybridization. As summarized in Fig. 2, the sequencing of chromosome-scale subgenomes B, A and D of the cultivated bread wheats (International Wheat Genome Sequencing Consortium, 2018) and of *T. urartu* (Ling *et al*., 2018) and *Ae. tauschii* (Luo *et al*., 2017) greatly advanced comparative genomics in *Aegilops-Triticum* and confirmed a chromosomal structure typical of Triticeae (Dvořák, 2009).

Despite its hybrid origin, the karyotype of *Ae. tauschii* significantly differs from a strict combination (i.e. expected additivity) of A and B progenitors (Table 1). This is strikingly illustrated by the absence

of a major nucleolus organizer region (NOR) on chromosome 1D of *Ae. tauschii*, while it was expected based on its location in both progenitor species. Such evolution of non-additive features following hybridization contrasts with other major NORs of A- and B- genome species that were retained on chromosomes 5 in *Ae. tauschii* or chromosomes 1, 5 and 6 in other species of the D-genome lineage. Consistent with a largely stochastic process, species of D-genome lineage show differential retention and loss of specific NOR loci from progenitors. Reorganization of NOR loci has been commonly reported in interspecific hybrids (Pikaard, 2000) and recent work highlighted hierarchical nucleolar dominance among current wild wheats (Mirzaghaderi *et al.*, 2017). Much remains to be understood regarding molecular underpinnings as well as evolutionary consequences of such chromosome-scale reorganization.

Comparative genome organization based on tandem repeats provides reliable insights on chromosome restructuring in species such as wild wheats (Heslop-Harrison, 2000). As shown in Fig. 3, species of the hybrid D-genome lineage present karyotypes combining specificities of progenitors such as repeats from both A- (e.g. pTa-535) and B- (e.g. pSc.119.2) genome species. Chromosomes of *Ae. tauschii* however appear more similar to the A- than to the B-genome progenitor species. Overall dominance of the A-genome lineage along its chromosomes is particularly apparent through enrichment of specific repeats such as pAs1 (present in A-genome but rare among B-genome species) that are particularly abundant in the hybrid D-genome of *Ae. tauschii* (Rayburn & Gill, 1986; Badaeva *et al.*, 2015). Such a pattern is strikingly matching biased retention of maternal A-genome nuclear genes involved in cytonuclear enzyme complexes in this species (Li *et al.*, 2019). In contrast to co-adaptation of cytoplasmic and nuclear loci affecting specific genes, comparisons of karyotypes support largely conserved A-chromosomes that appear rather consistent with an elusive genome-wide process. Other species of the D-genome lineage also show differential retention of progenitor sequences, with *Ae. comosa* and *Ae. uniaristata* showing similar bias, whereas chromosomes of *Ae. umbellulata*, *Ae. comosa* and species of the S* clade are closer to the B-genome progenitor. Given that genome-wide retrotransposons showed differential methylation depending on their genome of origin in wild wheats (Senerchia *et al*., 2016), other drivers than cyto-nuclear interactions may have participated in shaping hybrid genomes.

The contribution of the B-genome lineage to the reticulate origin of hybrid wild wheats is evident from pSc119.2 repeats that are abundant in both *Ae. speltoides* and *Ae. mutica*, and all derived species of the D-genome lineage, except *Ae. tauschii* (Badaeva et al. 1996a). However, pSc119.2 loci are scattered along chromosomes of *Ae. speltoides* and appear exclusively terminal in species of the D-genome lineage and related grasses from Triticeae (Taketa *et al*., 2000), suggesting that ancestrally terminal pSc119.2 repeats amplified in the B-genome lineage after the hybridization event at the origin of the D-genome lineage. Similarly, satellite sequences that are scarce (Spelt-52) or even absent (Spelt-1) in derived species of the D-genome lineage are particularly abundant in B-genome species (Pestsova *et al*., 1998; Ruban & Badaeva 2018). Such particularly high abundance of specific sequences in B-genome species suggests high levels of repeat amplification and chromosomal restructuring in this clade. Accordingly, the exceptional abundance of Spelt-52 repeats in sister S* species (i.e. *Ae. longissima* and *Ae. sharonensis*) supports late hybridization with *Ae. speltoides*, as also suggested by Bernhardt *et al.* (2019), and highlights events that may have contributed to pervasive signals of gene flow at the origin of the D-genome lineage (Huynh *et al.*, 2019). Consistent with the hypothesis of extensive sequence turnover in the B-genome lineage, Raskina *et al.* (2004) identified active En/Spm transposons that, alone or in interaction with the relocation and amplification of 45S and 5S rDNA sites, form hot spots for chromosomal rearrangements in *Ae. speltoides*. Various families of transposable elements were later shown to vary tremendously in copy number among individuals and selfed progenies of that species (Belyayev *et al.*, 2010), as expected through the segregation of large-scale structural variants. Cytogenetic evidence accordingly supports tremendous variation in C-banding and GAA-patterns as well as heteromorphic homologous chromosomes likely caused by chromosomal rearrangements within that species (Fig. 3). Unfortunately, the now-recognized sister species *Ae. mutica* has not yet been thoroughly investigated and to what extent it shares such high genome dynamics with *Ae. speltoides* is largely unknown.

Despite a phylogenetic framework generally reconciling patterns observed at the gene and the chromosome scale (Dvořák, 2009), the high sequence turnover in wild wheat genomes (particularly of the B-genome lineage) makes the early evolution of hybrid D-genome species still challenging to accurately infer with available phylogenomic approaches (see Huynh *et al.*, 2019). On top of further characterization of the so far neglected *Ae. mutica*, ancestral genome reconstruction based on high-

quality chromosome-scale assemblies among wild wheats would add to our understanding of genome evolution towards translational studies for cultivated wheats and Triticeae (Pont *et al*., 2017). Drivers of the biased retention of progenitor genomes in hybrid species deserve further attention, as both molecular and evolutionary underpinnings remain elusive (Hu & Wendel, 2019). Provided that highly-expressed genes under strong stabilizing selection may be disproportionally retained following hybridization, unbalanced expression of genes from divergent progenitors has been postulated to drive differential evolution of hybrid genomes, and diploid wild wheats thus appear as a promising model system to further disentangle the relative impact of gene network regulation *vs* genome-wide silencing of repeats (Woodhouse *et al.*, 2014).

Obesity of metacentric chromosomes in S wild wheats*

Following hybrid speciation at the origin of the D-genome lineage, derived species diversified in clades that present strikingly different trajectories of chromosomal evolution. Unlike other hybrids, species of the S* clade retained metacentric chromosomes while specifically undergoing substantial genome upsizing (Table 1). Based on estimates of Eilam *et al.* (2007), these wild wheats all show an increase of ca. 20% in genome size (i.e. more than 1 Gb) as compared to progenitors and other species of the D-genome lineage. Such evolution towards "genomic obesity" within the last 3 MY is likely driven by the amplification of repeated sequences at a higher rate than can be removed by short deletions due to illegitimate recombination (Ma & Bennetzen, 2004; Chantret *et al.*, 2005; Schubert & Vu, 2016). Such a process dramatically increased the genome size of the wild rice *Oryza australiensis* following the amplification of specific retrotransposons (Piegu *et al.*, 2006) and likely predominates among S* wild wheats. In a study of four families of retrotransposons, (Hosid *et al.*, 2012) considered a few samples from the S*-lineage (i.e. *Ae. bicornis*, *Ae. longissima* and *Ae. sharonensis*) and reported considerable variation within and among species. As expected under genetic drift as the main driver of genetic variation, widely divergent arrangements of retrotransposons among those species was mostly reflecting their phylogenetic relationships.

Different genomic repeats present clear phylogenetic signals and, together with retrotransposons, have contributed to karyotype evolution among S* species (Fig. 4). Early diverging species such as *Ae. searsii* and *Ae. bicornis* presented repeats (pAs1 and pTa-535) that nearly disappeared in derived species *Ae. sharonensis* and *Ae. longissima*, whereas those two sister species further hosted the amplification of GAA_n microsatellites and of Spelt52 satellite repeats (Ruban & Badaeva, 2018). Large variation in heterochromatic C-bands was early reported near and away from centromeres (Friebe & Gill 1996) and those species are indeed characterized by high genome dynamics. Underlying processes however appear to have evenly affected chromosome arms of S* wild wheats. Such sequence turnover with a conserved metacentric structure is thus particularly coherent with a stochastic accumulation of neutral interspersed repeats and small-scale chromosomal rearrangements across the whole genome of those species. Comparative genomics among S* wild wheats would provide more quantitative insights to address how a dynamic balance between insertion and deletion could favor the amplification of specific repeated sequences and support increased genome size as well as retention of metacentric chromosomes.

Increased chromosome asymmetry in CUMN wild wheats

In contrast to S* species and their conserved metacentric chromosomes, species of the CUMN clade present evidence of large-scale genome restructuring and acrocentric organization of specific chromosomes, while keeping a constant number of centromeres (Fig. 5). This contrast with Robertsonian fusions and fissions known to involve acrocentric chromosomes but yielding derived species with different chromosome numbers (Jones, 1998). The pattern shown by CUMN wild wheats indicates an evolution from ancestral metacentric towards acrocentric chromosomes, allowing neutral vs adaptive processes underlying such increasingly asymmetrical karyotypes to be addressed.

Karyotype structures in species of the CUMN clade are highly diverse and even sister species may differ substantially in chromosome organization and morphologies. Consistent with phylogenetic relationships, *Ae. comosa* (M) and *Ae. uniaristata* (N) are both characterized by abundant pAs1 repeats and present submetacentric M chromosomes, whereas N chromosomes are predominantly

acrocentric (Fig. 5). Chromosome 2 of *Ae. comosa* shows considerable asymmetry following pericentric inversions or terminal intrachromosomal translocation (Nasuda *et al.*, 1998; Iqbal *et al.*, 2000). This is also the case in *Ae. uniaristata* (N), *Ae. umbellulata* (U) and *Ae. caudata* (C) that are characterized by further translocations and pericentric inversions (Danilova *et al.*, 2017) having supported the evolution of two or three additional acrocentric chromosomes. Accordingly, chromosomes of *Ae. caudata* are all nearly-acrocentric and represent the current extreme of this trend towards increasing karyotype asymmetry in CUMN wild wheats.

Asymmetrical chromosomes becoming the majority in CUMN wild wheats is in striking contrast to conserved metacentric S* chromosomes, offering a unique opportunity to address the molecular and evolutionary drivers of genome restructuring. Large-scale deletions being deleterious in diploids, either uneven amplification/deletion of repeated sequences or chromosomal rearrangements may be underlying the evolution of asymmetry between chromosome arms. Profiling of transposable elements highlighted several recently active retrotransposon families among diploid wild wheats (Senerchia *et al.*, 2014). Unlike *BARE1* that presented higher intraspecific diversity than genome-wide random loci and was thus identified as transpositionally active in *Ae. tauschii*, *Ae. caudata*, *Ae. comosa* and *Ae. umbellulata*, several other retrotransposons families presented species-specific evolutionary trajectories. Although such dynamics of retrotransposons may have participated to genome restructuring, biased insertion of transposable elements driving localized expansion of chromosome segments has not been documented in the wheat group and repeated sequences appear to rather accumulate widely across the genome (e.g. (Wicker *et al.*, 2018). Similarly, genome-wide deletions driven by illegitimate recombination appear to rule small-scale DNA loss and genome streamlining (Levin, 2002; Chantret *et al.*, 2005; Schubert & Vu, 2016). Accordingly, the limited variation in genome size reported among most CUMN species supports a balance between sequence amplification and deletion, indicating that the evolution of acrocentric chromosomes was chiefly uncoupled from processes of genome expansion-contraction.

Centromere shifts apparent among CUMN wild wheats most likely resulted from chromosomal rearrangements such as pericentric inversions and unequal translocations among non-homologous chromosome arms that have been documented in those species (Fig. 5). In particular, different *Ae.*

umbellulata accessions frequently present reciprocal translocations that often involve satellite chromosomes at interstitial breakpoints and suggest that considerable structural variation is segregating within that species (Badaeva *et al*. 2004). Although close association to centromeres causing meiotic drive to preferentially fix acrocentric-like arrangements was suggested in some cases (e.g. Zanders *et al.*, 2014), it remains unknown to what extent such a putative transmission advantage may yield a karyotype with a majority of acrocentric chromosomes in plants. In contrast to the long-term conservation of metacentric chromosomes reported in S* wild wheats, neutral genome dynamics may look unlikely to have promoted increased karyotype asymmetry such as exhibited by CUMN wild wheats.

Stebbins (1971) postulated that non-neutral processes were at work to shape asymmetrical karyotypes and that acrocentric chromosomes of CUMN wild wheats have been progressively build-up by adaptive processes. Provided that several recent studies have reported chromosomal rearrangements promoting selection on linked loci (i.e. supergenes underlying complex phenotypes; Thompson & Jiggins, 2014; Charlesworth, 2016; Coughlan & Willis, 2019), the accumulation of long-lasting clusters of genes on specific chromosome arms may be consistent with Stebbins' hypothesis. As diploid wild wheats were shown to strive in significantly different ecological niches (Huynh *et al.,* 2020), an adaptive scenario linking chromosomal restructuring and ecological specialization appears plausible. The evolution of acrocentric chromosomes in CUMN wild wheats would have accordingly been chiefly driven by large-scale restructuring events that promoted the accumulation of non-recombining clusters of adaptive traits supporting the filling of a specific ecological niche. To what extent inversions or translocations may support not only the clustering of distant loci but also the necessary long-range reduction in recombination that is necessary for such a process to be effective remains largely unknown in wild wheats and more generally in plants (Hoffmann & Rieseberg, 2008; Fishman *et al.*, 2013). Similarly, the impact of microchromosomal rearrangements on recombination, such as promoted by interspersed retrotransposons in other species (Choudhury *et al.*, 2019), remains elusive in wild wheats. Noticeably, ruderal, selfing species such as wild wheats, whose populations are preadapted to quickly grow in a specific ecological niche and show reduced effective recombination, may particularly benefit from such gene clusters (Grant & Flake, 1974).

Integration of phylogenomics and molecular cytogenetics in wild wheats reveals remarkably contrasted patterns of chromosomal restructuring among recently diverged S* species (i.e. genome expansion within a conserved metacentric karyotype) vs CUMN species (i.e. increasingly asymmetrical chromosomes). Molecular and evolutionary processes underlying such evolutionary trajectories however remain to be characterized and fully integrated. Future studies should address the drivers of linkage disequilibrium among co-adapted sets of alleles and the possible adaptive spread of asymmetrical chromosomes in support of ecological radiation in wild wheats. In particular, the possible impact of linked selection in so-called complex genomes of CUMN species should be characterized to yield a deeper understanding of the interplay between structural variation and adaptation within and among species.

Genome restructuring: driver or spandrel of diversification in hybridizing wild wheats

Comparative genomics among multiple, recently diverged species is still in its infancy, but available surveys in plants point to a fast sequence turnover and a rather conservative evolution of low-copy genes as compared to repeated sequences and large gene families. In wheat relatives, pericentromeric and proximal regions comprising mostly Gypsy retrotransposons and low-copy genes contrast with distal ends of chromosomes that harbor fast-evolving genes riffled with active Copia retrotransposons and CACTA transposons (Luo *et al*., 2017; Fig. 2). Although such chromosomal organization is consistent with different evolutionary strata supporting conserved proximal vs dynamic distal regions, not much is known beyond *Ae. tauschii* and genome-wide surveys are necessary to shed light on underling drivers. Although the causes and consequences of genome restructuring remain elusive, large-scale chromosomal rearrangements were detected to well mirror the evolution of rice species (Stein *et al.*, 2018) as well as species with complex genomes such as in wild wheats or sunflowers (Ostevik *et al.*, 2019). It remains unclear to what extent genome restructuring is a driver or a by-product (i.e. spandrel) of species diversification.

The high genome dynamics of S* wild wheats going along with their recent diversification across the Middle East is consistent with neutral expansion of chromosomes having promoted the origin of new

species (Raskina *et al.*, 2008). Raskina *et al.* (2004) reported limited impact of chromosomal rearrangements on reproductive isolation in S* species and further studies may thus take advantage of this clade to address the interplay between intrinsic genome dynamics and stochastic factors in shaping species diversification. On the other hand, coherent with Stebbins' hypothesis, large-scale chromosomal rearrangements that accumulated among CUMN wild wheats could chiefly be by-products of selection initiated by environmental triggers to yield ecologically specialized species. Wild wheats thus offer outstanding opportunities to shed further light on the adaptive role of genome restructuring. Related clades with such divergent evolutionary trajectories is indeed promising to address structural and functional consequences of genome restructuring. Integrative work is now necessary to understand the genomic substrate of evolutionary radiation in wild wheats and address the impact of stochastic vs adaptive processes in shaping chromosomes.

More generally, it remains unclear to what extent hybridization is a necessary stimulus for such species diversification. Species such as sunflowers or wild wheats that have been thoroughly investigated regarding genome restructuring and ecological specialization indeed share a history of reticulate evolution. Cycles of genome divergence – merging are postulated to combine and reshuffle large-effect haplotypes that may promote genomic conflicts (Abbott *et al.*, 2013) and/or facilitate adaptive radiation (Marques *et al.*, 2019), and thus support dramatic species diversification. In that context, it may be noticed that CUMN species have been particularly prone to further hybridization, having generated the majority of wild wheats through allopolyploidy (Kilian *et al.*, 2011; Senerchia *et al.*, 2014; Feldman & Levy, 2015). Although beyond the scope here, comparative analysis of allopolyploid wild wheats would shed new light on chromosomal evolution through time (Badaeva *et al.*, 2002; Badaeva *et al.*, 2004) and the impact of large-scale restructuring *vs* dynamics of repeated sequences on evolutionary radiation. Underpinnings of biased fractionation vs gene-flow in shaping subgenomes of allopolyploid wild wheats also remains largely to be explored (Zohary & Feldman, 1962; Senerchia *et al.*, 2014; Mirzaghaderi & Mason, 2017). Now that technological advances enable variation from nucleotide to large-scale rearrangements to be compared among complex genomes (e.g. Borrill *et al.*, 2019), additional surveys may soon shed light on the causes and consequences of chromosomal restructuring following the combination of divergent species.

Acknowledgments

This work was supported by the Swiss National Science Foundation (project 31003A-153388) and the Russian State Foundation for Basic Research (project 17-04-00087). We thank S. Huynh, T. Marcussen and anonymous reviewers for valuable comments while preparing this manuscript.

References

Abbott, R, Albach, D, Ansell, S, Arntzen, JW, Baird, SJE, Bierne, N, Boughman, J, Brelsford, A, Buerkle, CA, Buggs, R, Butlin, RK, Dieckmann, U, Eroukhmanoff, F, Grill, A, Cahan, SH, Hermansen, JS, Hewitt, G, Hudson, AG, Jiggins, C, Jones, J, Keller, B, Marczewski, T, Mallet, J, Martinez-Rodriguez, P, Möst, M, Mullen, S, Nichols, R, Nolte, AW, Parisod, C, Pfennig, K, Rice, AM, Ritchie, MG, Seifert, B, Smadja, CM, Stelkens, R, Szymura, JM, Väinölä, R, Wolf, JBW, Zinner, D. 2013. Hybridization and speciation. *Journal of Evolutionary Biology* **26**: 229–246.

Badaeva, ED, Amosova, AV, Goncharov, NP, Macas, J, Ruban, AS, Grechishnikova, IV, Zoshchuk, SA, Houben, A. 2015. A set of cytogenetic markers allows the precise identification of all A-genome chromosomes in diploid and polyploid wheat. *Cytogenetic and Genome Research* **146**: 71–79.

Badaeva, ED, Amosova, AV, Muravenko, OV, Samatadze, TE, Chikida, NN, Zelenin, AV, Friebe, B, Gill, BS. 2002. Genome differentiation in *Aegilops*. 3. Evolution of the D-genome cluster. *Plant Systematics and Evolution* **231**: 163–190.

Badaeva, ED, Amosova, AV, Samatadze, TE, Zoshchuk, SA, Shostak, NG, Chikida, NN, Zelenin, AV, Raupp, WJ, Friebe, B, Gill, BS. 2004. Genome differentiation in *Aegilops*. 4. Evolution of the U-genome cluster. *Plant Systematics and Evolution* **246**: 45–76.

Badaeva, ED, Dedkova, OS, Gay, G, Pukhalskyi, VA, Zelenin, AV, Bernard, S, Bernard, M. 2007. Chromosomal rearrangements in wheat. Their types and distribution. *Genome* **50**: 907–926.

Badaeva, ED., Friebe, B, Gill, BS. 1996a. Genome differentiation in *Aegilops*. 1. Distribution of highly repetitive DNA sequences on chromosomes of diploid species. *Genome* **39**: 293-306.

Badaeva, ED, Friebe, B, Gill, BS. 1996b. Genome differentiation in *Aegilops*. 2. Physical mapping of 5S and 18S-26S ribosomal RNA gene families in diploid species. *Genome* **39**: 1150-1158.

Badaeva, ED, Fisenko, AV, Surzhikov, SA, Yankovskaya, AA, Chikida, NN, Zoshchuk, SA, Belousova, MK, Dragovich, AY. 2019. Genetic heterogeneity of a diploid grass *Aegilops tauschii* revealed by chromosome banding methods and electrophoretic analysis of the seed storage proteins (gliadins). *Russian Journal of Genetics* **55**: 1315-1329.

Belyayev, A, Kalendar, R, Brodsky, L, Nevo, E, Schulman, AH, Raskina, O. 2010. Transposable elements in a marginal plant population. Temporal fluctuations provide new insights into genome evolution of wild diploid wheat. *Mobile DNA* **1**: 6.

Bennetzen, J. L., Kellogg, E. A. 1997. Do plants have a one-way ticket to genomic obesity? *The Plant Cell* **9**: 1509–1514.

Bernhardt, N, Brassac, J, Kilian, B, Blattner, FR. 2017. Dated tribe-wide whole chloroplast genome phylogeny indicates recurrent hybridizations within Triticeae. *BMC Evolutionary Biology* **17**: 141.

Bernhardt, N, Brassac, J, Dong, X, Willing, E-M, Poskara, CH, Kilian, B, Blattner, FR. 2019. Genome-wide sequence information reveals recurrent hybridization among diploid wheat wild relatives. *The Plant Journal* (online: 10 dec 2019) doi:**10.1111/tpj.14641**.

Borrill, P, Harrington, SA, Uauy, C. 2019. Applying the latest advances in genomics and phenomics for trait discovery in polyploid wheat. *The Plant Journal* **97**: 56–72.

Chantret, N, Salse, J, Sabot, F, Rahman, S, Bellec, A, Laubin, B, Dubois, I, Dossat, C, Sourdille, P, Joudrier, P, Gautier, M-F, Cattolico, L, Beckert, M, Aubourg, S, Weissenbach, J, Caboche, M, Bernard, M, Leroy, P, Chalhoub, B. 2005. Molecular basis of evolutionary events that shaped

the hardness locus in diploid and polyploid wheat species (*Triticum* and *Aegilops*). *The Plant Cell* **17**: 1033–1045.

Charlesworth, D. 2016. The status of supergenes in the 21st century: recombination suppression in Batesian mimicry and sex chromosomes and other complex adaptations. *Evolutionary Applications* **9**: 74–90.

Choudhury, RR, Rogivue, A, Gugerli, F, Parisod, C. 2019. Impact of polymorphic transposable elements on linkage disequilibrium along chromosomes. *Molecular Ecology* **28**: 1550–1562.

Contento, A, Heslop-Harrison, JS, Schwarzacher T. 2005. Diversity of a major repetitive DNA sequence in diploid and polyploid Triticeae. *Cytogenetic and Genome Research* **109**:34-42.

Coughlan, JM, Willis, JH. 2019. Dissecting the role of a large chromosomal inversion in life history divergence throughout the *Mimulus guttatus* species complex. *Molecular Ecology* **28**: 1343–1357.

Danilova, TV, Akhunova, AR, Akhunov, ED, Friebe, B, Gill, BS. 2017. Major structural genomic alterations can be associated with hybrid speciation in *Aegilops markgrafii* (Triticeae). *The Plant Journal* **92**: 317–330.

Dvořák, J. 2009. Triticeae genome structure and evolution. In Muehlbauer, GJ & Feuillet, C. (Eds.), Genetics and genomics of the Triticeae. New York: Springer US, 685–711.

Eilam, T, Anikster, Y, Millet, E, Manisterski, J, Sagi-Assif, O, Feldman, M. 2007. Genome size and genome evolution in diploid Triticeae species. *Genome* **50**: 1029–1037.

Feldman, M, Levy, AA. 2015. Origin and evolution of wheat and related Triticeae species. In: Molnár-Láng, M, Ceoloni, C & Doležel J. (edS.), Alien Introgression in Wheat. Cham: Springer International Publishing Switzerland, 21-76.

Fishman, L, Stathos, A, Beardsley, PM, Williams, CF, Hill, JP. 2013. Chromosomal rearrangements and the genetics of reproductive barriers in *Mimulus* (monkey flowers). *Evolution* **67**: 2547–2560.

Friebe, B, Gill, BS. 1996. Chromosome banding and genome analysis in diploid and cultivated polyploid wheats. In Jauhar PP. (ed.), Methods in genome analysis in plants: their merits and pitfalls. New York, London, Tokyo: Boca Ration, CRC Press, 39-60.

Glémin, S, Scornavacca, C, Dainat, J, Burgarella, C, Viader, V, Ardisson, M, Sarah, G, Santoni, S, David, J, Ranwez, V. 2019. Pervasive hybridizations in the history of wheat relatives. *Science Advances* **5**: eaav9188.

Grant, V, Flake, RH. 1974. Solutions to the cost-of-selection dilemma. *Proceedings of the National Academy of Sciences* **71**: 3863–3865.

Heslop-Harrison, JS. 2000. Comparative genome organization in plants: from sequence and markers to chromatin and chromosomes. *Plant Cell* **12**: 617–635.

Hoffmann, AA, Rieseberg, LH. 2008. Revisiting the impact of inversions in evolution. From population genetic markers to drivers of adaptive shifts and speciation? *Annual Review of Ecology, Evolution, and Systematics* **39**: 21–42.

Hosid, E, Brodsky, L, Kalendar, R, Raskina, O, Belyayev, A. 2012. Diversity of long terminal repeat retrotransposon genome distribution in natural populations of the wild diploid wheat *Aegilops speltoides*. *Genetics* **190**: 263–274.

Hu, G, Wendel, JF. 2019. Cis–trans controls and regulatory novelty accompanying allopolyploidization. *New Phytologist* **221**: 1691–1709.

Huynh, S, Broennimann, O, Guisan, A, Felber, F, Parisod, C. 2020. Eco-genetic additivity of diploids in allopolyploid wild wheats. *Ecology Letters* **in press**. https://doi.org/10.22541/au.157486393.30658173

Huynh, S, Marcussen, T, Felber, F, Parisod, C. 2019. Hybridization preceded radiation in diploid wheats. *Molecular Phylogenetics and Evolution* **139**: 106554.

International Wheat Genome Sequencing Consortium. 2018. Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* **361**: eaar7191.

Iqbal, N, Reader, SM, Caligari, PDS, Miller, TE. 2000. Characterization of *Aegilops uniaristata* chromosomes by comparative DNA marker analysis and repetitive DNA sequence in situ hybridization. *Theoretical and Applied Genetics* **101**: 1173–1179.

Jones, K. 1998. Robertsonian fusion and centric fission in karyotype evolution of higher plants. *Botanical Review* **64**: 273–289.

Kilian, B, Mammen, K, Millet, E, Sharma, R, Graner, A, Salamini, F, Hammer, K, Özkan, H. 2011. Aegilops. In *Wild Crop Relatives: Genomic and Breeding Resources*, edited by C. Kole: Springer-Verlag Berlin Heidelberg. p. 1–76.

Levin, DA. 2002. The role of chromosomal change in plant evolution. New York: Oxford University Press.

Li, C, Sun, X, Conover, JL, Zhang, Z, Wang, J, Wang, X, Deng, X, Wang, H, Liu, B, Wendel, JF, Gong, L. 2019. Cytonuclear coevolution following homoploid hybrid speciation in *Aegilops tauschii*. *Molecular Biology and Evolution* **36**: 341–349.

Ling, H-Q, Ma, B, Shi, X, Liu, H, Dong, L, Sun, H, Cao, Y, Gao, Q, Zheng, S, Li, Y, Yu, Y, Du, H, Qi, M, Li, Y, Lu, H, Yu, H, Cui, Y, Wang, N, Chen, C, Wu, H, Zhao, Y, Zhang, J, Li, Y, Zhou, W, Zhang, B, Hu, W, van Eijk, MJT, Tang, J, Witsenboer, HMA, Zhao, S, Li, Z, Zhang, A, Wang, D, Liang, C. 2018. Genome sequence of the progenitor of wheat A subgenome *Triticum urartu*. *Nature* **557**: 424–428.

Luo, M-C, Gu, YQ, Puiu, D, Wang, H, Twardziok, SO, Deal, KR, Huo, N, Zhu, T, Le Wang, Wang, Y, McGuire, PE, Liu, S, Long, H, Ramasamy, RK, Rodriguez, JC, Van, SL, Yuan, L, Wang, Z, Xia, Z, Xiao, L, Anderson, OD, Ouyang, S, Liang, Y, Zimin, AV, Pertea, G, Qi, P, Bennetzen, JL, Dai, X, Dawson, MW, Müller, H-G, Kugler, K, Rivarola-Duarte, L, Spannagl, M, Mayer, KFX, Lu, F-H, Bevan, MW, Leroy, P, Li, P, You, FM, Sun, Q, Liu, Z, Lyons, E, Wicker, T, Salzberg, SL, Devos, KM, Dvořák, J. 2017. Genome sequence of the progenitor of the wheat D genome *Aegilops tauschii*. *Nature* **551**: 498–502.

Ma, J, Bennetzen, JL. 2004. Rapid recent growth and divergence of rice nuclear genomes. *Proceedings of the National Academy of Sciences* **101**: 12404–12410.

Mascher, M, Gundlach, H, Himmelbach, A, Beier, S, Twardziok, SO, Wicker, T, Radchuk, V, Dockter, C, Hedley, PE, Russell, J, Bayer, M, Ramsay, L, Liu, H, Haberer, G, Zhang, X-Q, Zhang, Q, Barrero, RA, Li, L, Taudien, S, Groth, M, Felder, M, Hastie, A, Šimková, H, Staňková, H, Vrána, J, Chan, S, Muñoz-Amatriaín, M, Ounit, R, Wanamaker, S, Bolser, D, Colmsee, C, Schmutzer, T, Aliyeva-Schnorr, L, Grasso, S, Tanskanen, J, Chailyan, A, Sampath, D, Heavens, D, Clissold, L, Cao, S, Chapman, B, Dai, F, Han, Y, Li, H, Li, X, Lin, C, McCooke, JK, Tan, C, Wang, P, Wang, S, Yin, S, Zhou, G, Poland, JA, Bellgard, MI, Borisjuk, L, Houben, A, Doležel, J, Ayling, S, Lonardi, S, Kersey, P, Langridge, P, Muehlbauer, GJ, Clark, MD, Caccamo, M, Schulman, AH, Mayer, KFX, Platzer, M, Close, TJ, Scholz, U, Hansson, M, Zhang, G, Braumann, I, Spannagl, M, Li, C, Waugh, R, Stein, N. 2017. A chromosome conformation capture ordered sequence of the barley genome. *Nature* **544**: 427–433.

Marcussen, T, Sandve, S. R., Heier, L., Spannagl, M., Pfeifer, M.; International Wheat Genome Sequencing Consortium, Jakobsen, K.S., Wulff, B. B., Steuernagel, B., Mayer, K. F., Olsen, O. A. 2014. Ancient hybridizations among the ancestral genomes of bread wheat. *Science* **345**: 1250092.

Marques, DA, Meier, JI, Seehausen, O. 2019. A combinatorial view on speciation and adaptive radiation. *Trends in Ecology & Evolution* **34**: 531-544.

Mirzaghaderi, G, Abdolmalaki, Z, Zohouri, M, Moradi, Z, Mason, AS. 2017. Dynamic nucleolar activity in wheat × *Aegilops* hybrids: evidence of C-genome dominance. *Plant Cell Reports* **36**: 1277–1285.

Mirzaghaderi, G, Mason, AS. 2017. Revisiting pivotal-differential genome evolution in wheat. *Trends in Plant Science* **22**: 674–684.

Nasuda, S, Friebe, B, Bush, W, Kynast, RG, Gill, BS. 1998. Structural rearrangement in chromosome 2M of *Aegilops comosa* has prevented the utilization of the compair and related wheat-*Ae. comosa* translocations in wheat improvement. *Theoretical and Applied Genetics* **96**: 780–785.

Ohta, S. 1991. Phylogenetic relationship of *Aegilops mutica* Boiss with the diploid species of congeneric *Aegilops-Triticum* complex, based on the new method of genome analysis using its B-chromosomes. *Memoirs of the College of Agriculture Kyoto University* **137**: 1–116.

Ostevik, KL, Samuk, K, Rieseberg, LH. 2019. Ancestral reconstruction of sunflower karyotypes reveals dramatic chromosomal evolution. *bioRxiv*: http://dx.doi.org/10.1101/737155.

Pestsova, E. G., Goncharov, N. P., Salina, E. A. (1998) Elimination of a tandem repeat of telomeric heterochromatin during the evolution of wheat. *Theoretical and Applied Genetics* **97**:1380-1386.

Piegu, B, Guyot, R, Picault, N, Roulin, A, Sanyal, A, Kim, H, Collura, K, Brar, DS, Jackson, S, Wing, RA, Panaud, O. 2006. Doubling genome size without polyploidization. dynamics of retrotransposition-driven genomic expansions in *Oryza australiensis*, a wild relative of rice. *Genome Research* **16**: 1262–1269.

Pikaard, CS. 2000. Nucleolar dominance: uniparental gene silencing on a multi-megabase scale in genetic hybrids. *Plant Molecular Biology* **43**: 163–177.

Pont, C., Salse, J. 2017. Wheat paleohistory created asymmetrical genomic evolution. *Current Opinion in Plant Biology* **36**:29–37.

Raskina, O, Barber, JC, Nevo, E, Belyayev, A. 2008. Repetitive DNA and chromosomal rearrangements. speciation-related events in plant genomes. *Cytogenetics and Genome Research* **120**: 351–357.

Raskina, O, Belyayev, A, Nevo, E. 2004. Quantum speciation in *Aegilops*. Molecular cytogenetic evidence from rDNA cluster variability in natural populations. *Proceedings of the National Acadademy of Sciences* **101**: 14818–14823.

Ruban, AS, Badaeva, ED. 2018. Evolution of the S-genomes in *Triticum-Aegilops* alliance: Evidences from chromosome analysis. *Frontiers in Plant Science* **9**: 1756.

Salina EA, Adonina IG, Vatolina TY, Kurata N. 2004. A comparative analysis of the composition and organization of two subtelomeric repeat families in *Aegilops speltoides* Tausch. and related species. *Genetica* **122**: 227-237.

Schubert, I, Vu, GTH. 2016. Genome stability and evolution. Attempting a holistic view. *Trends in Plant Science* **21**: 749–757.

Senerchia, N, Felber, F, Parisod, C. 2014. Contrasting evolutionary trajectories of multiple retrotransposons following independent allopolyploidy in wild wheats. *New Phytologist* **202**: 975–985.

Senerchia, N, Felber, F, Parisod, C. 2015. Genome reorganization in F1 hybrids uncovers the role of retrotransposons in reproductive isolation. *Proceedings B* **282**: 20142874.

Senerchia, N, Wicker, T, Felber, F, Parisod, C. 2013. Evolutionary dynamics of retrotransposons assessed by high-throughput sequencing in wild relatives of wheat. *Genome Biology and Evolution* **5**: 1010–1020.

Stebbins, GL. 1971. Chromosomal evolution in higher plants. London: Adward Arnold.

Stein, JC, Yu, Y, Copetti, D, Zwickl, DJ, Zhang, L, Zhang, C, Chougule, K, Gao, D, Iwata, A, Goicoechea, JL, Wei, S, Wang, J, Liao, Y, Wang, M, Jacquemin, J, Becker, C, Kudrna, D, Zhang, J, Londono, CEM, Song, X, Lee, S, Sanchez, P, Zuccolo, A, Ammiraju, JSS, Talag, J, Danowitz, A, Rivera, LF, Gschwend, AR, Noutsos, C, Wu, C-C, Kao, S-M, Zeng, J-W, Wei, F-J, Zhao, Q, Feng, Q, El Baidouri, M, Carpentier, M-C, Lasserre, E, Cooke, R, Rosa Farias, Dd, da Maia, LC, Dos Santos, RS, Nyberg, KG, McNally, KL, Mauleon, R, Alexandrov, N, Schmutz, J, Flowers, D, Fan, C, Weigel, D, Jena, KK, Wicker, T, Chen, M, Han, B, Henry, R, Hsing, Y-IC, Kurata, N, Oliveira, AC de, **Panaud, O, Jackson, SA, Machado, CA, Sanderson, MJ, Long, M, Ware, D, Wing, RA. 2018.** Genomes of 13 domesticated and wild rice relatives highlight genetic conservation, turnover and innovation across the genus Oryza. *Nature Genetics* **50**: 285-296.

Taketa, S., Ando, H., Takeda, K., Harrison, G. E., Heslop-Harrison, J. S. (2000) The distribution, organization and evolution of two abundant and widespread repetitive DNA sequences in the genus *Hordeum*. *Theoretical and Applied Genetics* **100**: 169-176.

Tenaillon, MI, Hollister, JD, Gaut, BS. 2010. A triptych of the evolution of plant transposable elements. *Trends in Plant Science* **15**: 471–478.

Thompson, MJ, Jiggins, CD. 2014. Supergenes and their role in evolution. *Heredity* **113**: 1–8.

van Slageren, MW. 1994. Wild wheats: a monograph of *Aegilops* L. and *Ambylopyrum* (Jaub. & Spach) Eig (Poaceae). *Wageningen Agricultural University Papers* **94-7**: 1–513.

Wicker, T, Gundlach, H, Spannagl, M, Uauy, C, Borrill, P, Ramírez-González, RH, Oliveira, R de, **Mayer, KFX, Paux, E, Choulet, F. 2018.** Impact of transposable elements on genome structure and evolution in bread wheat. *Genome Biology* **19**: 103.

Woodhouse, MR, Cheng, F, Pires, JC, Lisch, D, Freeling, M, Wang, X. 2014. Origin, inheritance, and gene regulatory consequences of genome dominance in polyploids. *Proceedings of the National Academy of Sciences* **111**: 5283–5288.

Zanders, SE, Eickbush, MT, Yu, JS, Kang, J-W, Fowler, KR, Smith, GR, Malik, HS. 2014. Genome rearrangements and pervasive meiotic drive cause hybrid infertility in fission yeast. *Elife* **3**: e02630.

Zohary, D, Feldman, M. 1962. Hybridization between amphidiploids and the evolution of polyploids in the wheat (*Aegilops-Triticum*) group. *Evolution* **16**: 44–61.

Legends to Figures

Fig. 1. Phylogenetic relationships and chromosomal evolution among diploid wheats (*Aegilops-Triticum*). The genome composition of each species is indicated in between brackets and the reticulation event between species of the A- and the B-genome lineages at the origin of the D-genome lineage is shown as dotted lines following Huynh *et al*. (2019), with mean node ages and 95% credibility intervals in million years ago (MYA). Chromosomes were labeled with the A/D-genome-specific repeat pAs1 (green) and the B/S-genome-specific repeat pSc119.2 (red), and FISH karyotypes following Badaeva *et al*. (1996a, 2015, 2019) and Ruban & Badaeva (2018) show progenitor species from the A-genome (here, *T. monococcum*) and B-genome (here, *Ae. speltoides*) lineages together with selected karyotypes from the three clades nested within the hybrid D-genome lineage: (i) *Aegilops tauschii* shows the A-like karyotype of the *tauschii* clade (also see Fig. 3), (ii) *Ae. longissima* and its large metacentric chromosomes is representative of "genomic obesity" in the S* clade (also see Fig. 4) and (iii) *Ae. caudata* illustrates species of the variable CUMN clade and their acrocentric chromosomes (also see Fig. 5).

Fig. 2: Sequence data in wild wheats (*Aegilops-Triticum*) and genome organization in Triticeae. A: Comparative genomics based on chromosome-scale assemblies of subgenomes from cultivated wheats (i.e. BAD; International Wheat Genome Sequencing Consortium, 2018), of *T. urartu* (A-genome; Ling *et al*., 2018) and *Ae. tauschii* (D-genome; Luo *et al*., 2017) supports the typical a predominance of transposable elements in wild wheats (around 85%). Diploid species present noticeably different profiles of repeats that otherwise match abundances reported among subgenomes of bread wheat (Wicker *et al*., 2018). The gene space (i.e. gene + non-repeated sequences) is thus rather restricted with only 7.5% of the genome corresponding to high confidence genes in *Ae. tauschii*. **B:** The current assembly of *Ae. tauschii* comprises more than 95% of its genome (Luo *et al*., 2017) and matches the typical organization of chromosomes in Triticeae (Dvořák, 2009). Genes are rare across the 50 to 100 MB of pericentromeric regions, at low density in proximal regions of

chromosomes comprising mostly low-copy genes and at high density only towards subtelomeric regions. Matching recent insights from nuclear conformation of chromosomes in Barley (Mascher *et al*., 2017), mostly distal ends of chromosomes show frequent recombination and are enriched in gene families in *Ae. tauschii*. Such organization is generally coherent with the hypothesis of conservative vs dynamic evolutionary strata of Triticeae chromosomes (Dvořák 2009).

Tandem repeats appear abundant across specific regions of chromosomes such as pericentromeric microsatellites (GAA_n or GTT_n) or subtelomeric islands (Spelt-1, Spelt-52; Salina *et al*. 2004), although likely underestimated at less than 1% in genome assemblies (e.g. Liu *et al*. 2017). Despite conserved presence of such repeat islands in Triticeae, their organization and primary sequence varies even among species (e.g. pAs1, pSc.119.2, Spelt-1; Contento *et al*. 2005; Salina *et al*. 2004).

Transposable elements represent the vast majority (i.e. 84.4%) of the *Ae. tauschii* genome. On top of 16% of 'cut-paste' DNA transposon (mostly CACTA), more than 65% of its genome is composed of interspersed copies from some 550 divergent families of 'copy-paste' long-terminal repeat retrotransposons. As in other Triticeae, the Copia type (RLC) represents the minority, with relatively high copy densities across distal ends of chromosomes. Some abundant families such as the *BARE1* group (i.e. *Angela*, *BARE1* and *WIS*) present around 2000 intact copies suggestive of recent insertions across the gene space. The Gypsy type (RLG) accounts for more than half the 20'000 intact copies identified in *Ae. tauschii* and therefore comprises several families having recently proliferated. *Fatima* is an example of RLG having diversified with wild wheats to currently present divergent sequence pools among species and that is particularly abundant in *Ae. tauschii*. As in other Triticeae, the RLG *Cereba* interacts with CENH3 histones and is therefore strictly centromeric, despite around 200 recently inserted copies indicative of sequence turnover in *Ae. tauschii*. Although still poorly understood, the chromosomal organization of *Ae. tauschii* matches the typical Triticeae model with particularly abundant RLG copies in pericentromeric and proximal regions, whereas RLC show a more even distribution and are particularly abundant across distal ends of chromosomes.

Fig. 3. Contrasting karyotypes of putative progenitors A-genome species (*Triticum urartu* and *T. monococcum*) and B-genome species (*Aegilops speltoides* and *Ae. mutica*) compared with a representative species of the hybrid D-genome lineage, *Ae. tauschii*. As summarized in Table 1 and discussed in the text, the distribution of repeated sequences among samples of *Ae. tauschii* combines A-specific (e.g. pTA-535, pAs1) and B-specific (e.g. pSC.119.2) characteristics. Such repeats selected for their correlation between abundance and cytogenetic signals appear globally more similar to the karyotype of A-genome progenitors. In contrast, the organization of major nucleolus organizer regions (NOR, 5S) in *Ae. tauschii* differs from a combination of A- and B- genome progenitors. Probe combinations are shown on the top with colors corresponding to signals along homeologous chromosomes 1 to 7 (also see Table 1). Accession names are given on the bottom. FISH-karyotypes were constructed based on previous analyses as in Fig. 1.

Fig. 4. Karyotype evolution towards "genomic obesity" in the S* clade of the hybrid D-genome lineage of wild wheats. Enlarged chromosomes appears to match the increased genome size estimated in progenitor species (Table 1). Large variation in C-banding patterns and distribution of GAA_n microsatellites is visible within and among species. Subtelomeric Spelt52 repeats that are abundant in *Aegilops speltoides* (Fig. 3) are restricted to the sister species *Ae. sharonensis* and *Ae. longissima*. Probe combinations are shown on the top with colors corresponding to signals along homeologous chromosomes 1 to 7 and accession names are given on the bottom, based on previous analyses as in Fig. 1.

Fig. 5. Karyotype evolution towards a majority of asymmetrical chromosomes (i.e. "karyotype asymmetry") in the CUMN clade of the hybrid D-genome lineage of wild wheats. Submetacentric and predominantly acrocentric chromosomes of *Aegilops comosa* and *Ae. uniaristata* show abundant pAs1 repeats. On top of chromosome 2 asymmetry that is shared by all species, *Ae. umbellulata* and *Ae. caudata* present two and three additional acrocentric chromosomes, respectively. All chromosomes of *Ae. caudata* are therefore nearly-acrocentric, illustrating karyotype asymmetry. Probe combinations are shown on the top with colors corresponding to signals along homeologous chromosomes 1 to 7 (also see Table 1) and accession names are given on the bottom, based on previous analyses as in Fig. 1.

Table 1: Diploid wild wheat species and main characteristics of their karyotype as shown in Fig. 3, 4 and 5.

Lineage / Genomic composition	**Species** (synonyms)	**GS**[a] (pg)	**C-band**[b]	**GAA**[b]	**GTT**[b]	**NOR**[c] major /minor (chrom. pos)	**5S**[c] (chrom. pos)	**Sp52**[d]	**Sp1**[d]	**Ta**[d]	**As1**[d]	**Sc**[d]
A / A^m	*T. monococcum subsp. monococcum subsp. aegilopoides* (= *T. boeoticum*)	6.45	+	+	+	2 (1S+5S) / 3(5L+6S+7L)	2 (1S=5S)	+	0	+++	++	0
A / A^u	*T. urartu*	6.02	+	+	+	2 (1S+5S) / 3(5L+6S+7L)	2 (1S=5S)	+	0	+++	++	0
B / S	*Ae. speltoides*	5.81	+++	++	+++	2 (1S+6S) / 0	1 (5S)	+++	+++	0	0	+++
B / T	*Ae. mutica* (= *Amblyopyrum muticum*)	5.82	++	++	+++	2 (1S+6S) / 1 (7L)	2 (1S>5S)	na	0	0	+	+++
D / D	*Ae. tauschii*	5.17	+	+(0)	+	1 (5S) / 1 (7L)	2 (1S=5S)	0	0	+++	+++	+
D/ S*(S^b)	*Ae. bicornis*	6.84	++	++	+	2 (5S+6S) / 2 (1S+6L)	2 (1S>5S)	0	0	+	+	+++
D / S*(S^s)	*Ae. searsii*	6.65	++	++	+	2 (5S+6S) / 2 (1S+6L)	2 (1S=5S)	0	0	+	+	+++
D / S*(S^l)	*Ae. longissima*	7.48	+++	+++	+	2 (5S+6S) / 2 (1S+6L)	2 (1S=5S)	+++	0	+	+	+++
D / S*(S^l)	*Ae. sharonensis*	7.52	+++	+++	+	2 (5S+6S) / 2 (1S+6L)	2 (1S=5S)	+++	0	+	+	+++

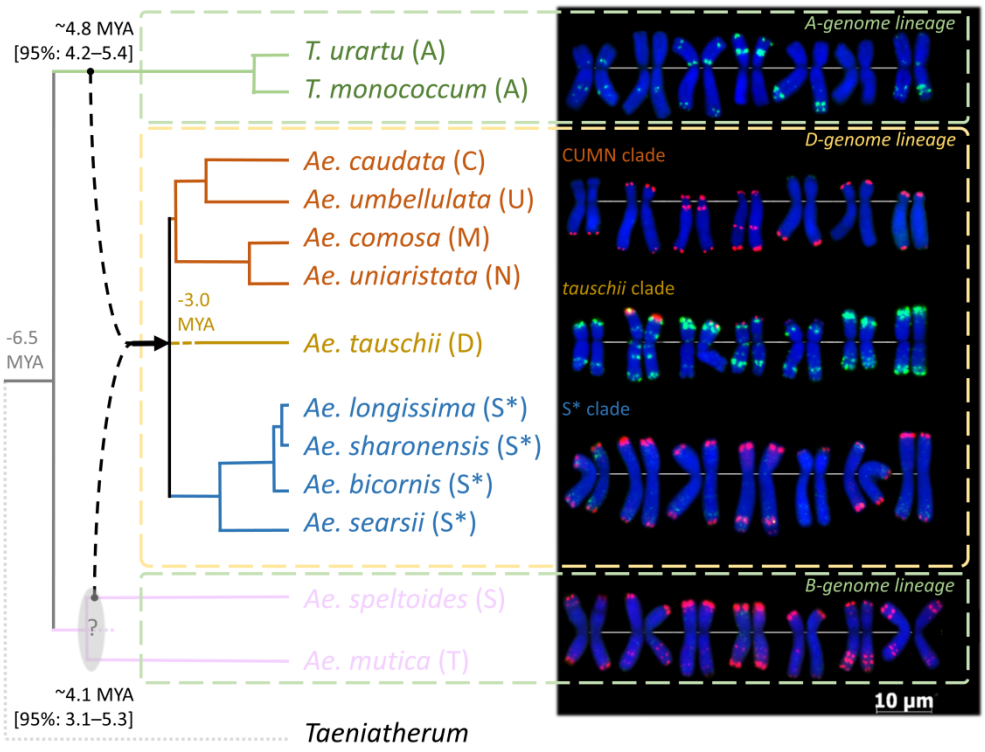
D / C	*Ae. caudata*	4.84	+++	+++	+++	2 (1S+5S) / 0	2 (1S=5S)	0	0	0	+	+++
D / U	*Ae. umbellulata*	5.38	+++	+++	+	2 (1S+5S) / 1 (6L)	2 (1S=5S)	0	0	0	+	+++
D / M	*Ae. comosa* (= *Ae. markgrafii*)	5.53	++	++	+	2 (1S+6S) / 5 (2S+3S+4L+7L)	2 (1S>5S)	0	0	++	++	++
D / N	*Ae. uniaristata*	8.82	+++	+++	+++	1 (5S) / 6 (1L+2L+3L+7L)	2 (1L>5S)	0	0	+++	+++	+++
Out	*Taeniatherum caput-medusae*	4.31	+	+	++	1 (1S)	2 (1L>5L)	0	0	+	+	++

[a]Genome Size (GS) in pg from Eilam et al. 2007. For *Taeniatherum caput-medusae* from https://cvalues.science.kew.org

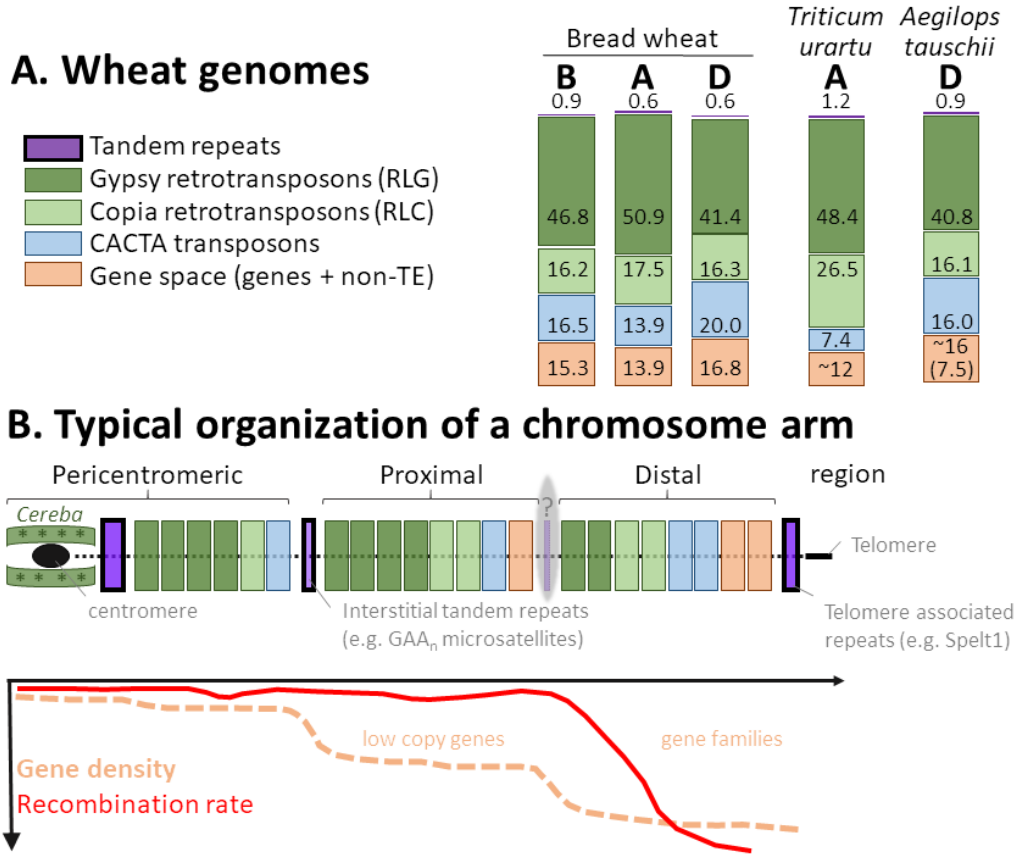
[b]Abundance of C-bands, largely co-locating with GAA_n microsatellites, presented as low (+), medium (++) or high (+++) following Friebe & Gill (1996) and Badaeva *et al*. (1996a).

[c]Positions of nucleolar organizing regions (NOR) and 5S ribosomal RNA gene (5S) on chromosome arms, following Badaeva *et al*. (1996b).

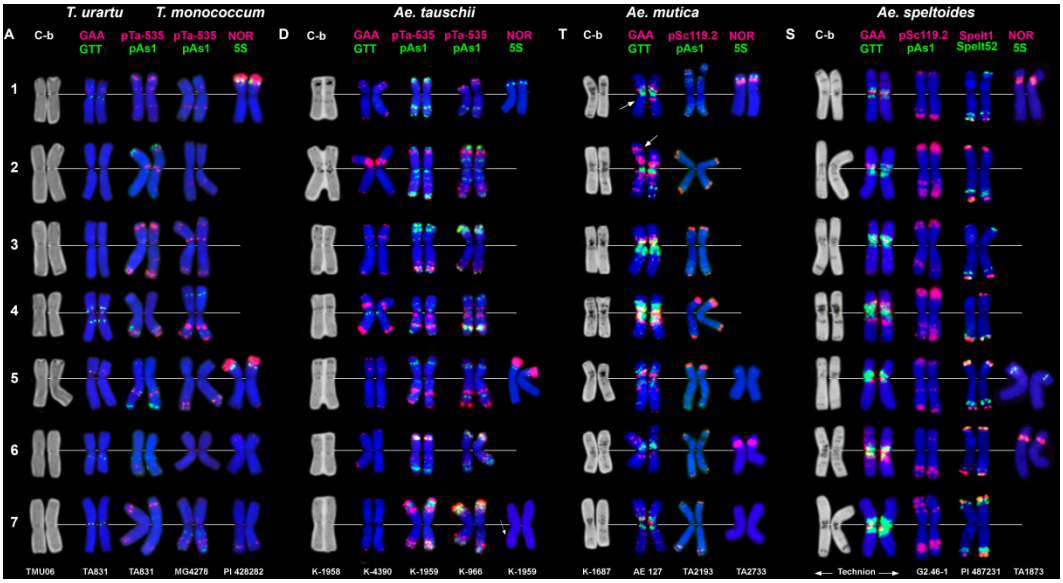
[d]Abundance of tandem repeats Spelt52 (Sp52), Spelt1 (Sp1), pTa-535 (Ta), pAs1 (As), pSc119.2 (Sc) characterized in Badaeva *et al*. (1996a) and Ruban & Badaeva (2018) and summarized here as absent (0), low (+), medium (++), high (+++) or not available (na) based on their correlation between abundance in genomes and intensity of FISH signals.



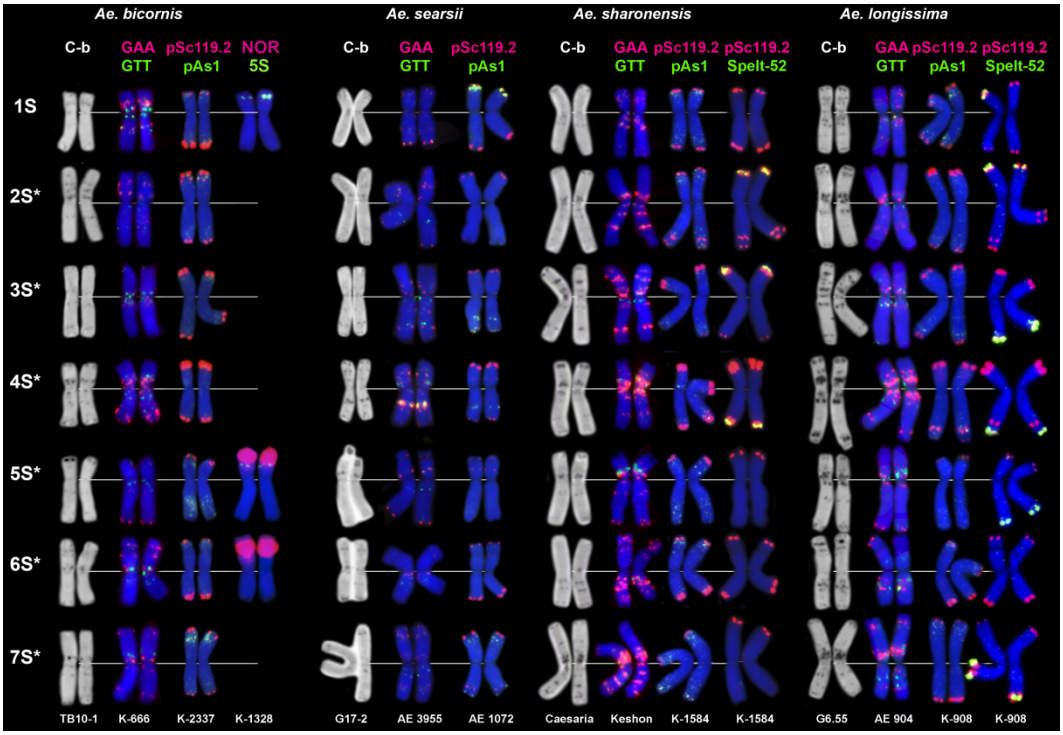
nph_16415_f1.tif



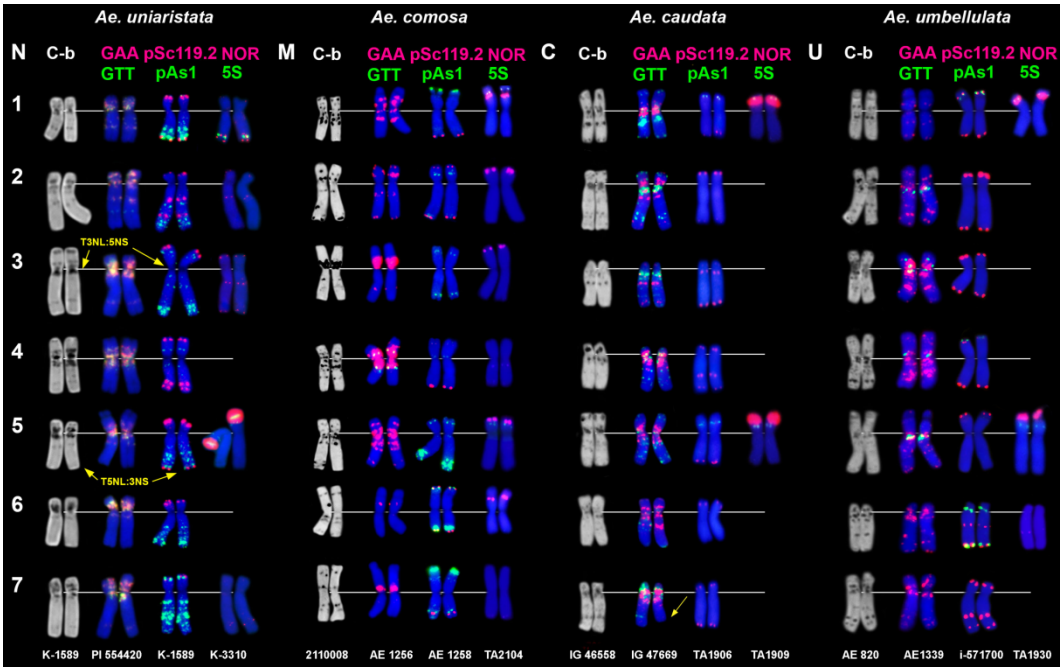
nph_16415_f2.tif



nph_16415_f3.tiff



nph_16415_f4.tiff



nph_16415_f5.tiff