

# Abstract

Synthetic neopolyploids offer an insight into the complex nature of polyploid formation and genetics during their early stages. The sudden introgression of new genetic material into the genome during allopolyploidization can lead to 'genomic shock' whereby mutations at various levels not typically seen in well-established polyploids arise. Here, I explored the effect of allopolyploidization on the genome in nascent allohexaploid mutation accumulation (MA) lines which had been formed from a cross between the natural allotetraploid *Arabidopsis suecica* (ÁÁBB) and diploid *Arabidopsis thaliana* Col-0 (AA). Crossing these two species had given rise to a triploid hybrid individual which underwent an unexpected spontaneous whole-genome duplication (WGD) (Matsushita et al. 2012). The progeny were used to initiate 37 MA lines which were propagated via single seed descent. My analysis focused on six of these lines where I investigated the genomic changes that occurred in the resulting MA lines up to 17 generations. The allohexaploid offspring genomes consist of two subgenomes from *A. suecica* and one from *A. thaliana* Col-0. and have a complex allopolyploid karyotype (AAÁÁBB;  $2n=6x=36$ ;  $4n=20$ ,  $2n=16$ ) with two copies of the *A. arenosa* genome (BB) and four copies of the *A. thaliana* genome (AAÁÁ) from two different accessions.

To facilitate the resequencing analysis of the allohexaploid offspring geno-

---

mes, I first generated a chromosome-level assembly of *A. suecica* Sue1 accession with an N50 of 19.06 Mb. By analyzing the resequencing data of the neoallohexaploid genomes in these six lines, I was able to gain insight into the evolutionary process of early polyploids. Although no clear patterns of rediploidization could be found in the allohexaploid data, a significant level of genetic material translocations between the homoeologous subgenomes of the allohexaploids could be identified, revealing instabilities not present or present to a much lower degree in well-established polyploids such as *A. suecica*. A coverage-based optimization method was used to predict whole chromosome gain or loss in the *A. thaliana* subgenome of the allohexaploid offspring. 22 out of the 42 individuals sequenced were found to have either a gain or loss of whole chromosome copies and these could be validated using the SNP allele frequencies between the parental subgenomes. This finding highlights the genomic instability of the neopolyploids as compared to well-established polyploids. Chromosome arm translocations between subgenomes were identified in 19 allohexaploid individuals, and the majority of them were found to be fixed or partially fixed within the lines. No evidence of a transposable element (TE) burst or high levels of TE proliferation resulting from 'genomic shock' from the allopolyploidization was observed.

This analysis shows that there is a high degree of genomic variability in the early hexaploids even between generations and sisters of the same line and

---

gives an insight into the early evolution of large-scale mutations in nascent polyploids towards rediploidization.

# Zusammenfassung

Synthetische Neopolyploide bieten einen Einblick in die komplexe Natur der polyploiden Bildung und Genetik in ihren frühen Stadien. Das plötzliche Eindringen von neuem genetischem Material in das Genom während der Allopolyploidisierung kann zu einem „genomischen Schock“ führen, wodurch Mutationen auf verschiedenen Ebenen entstehen, die bei etablierten Polyploiden normalerweise nicht zu sehen sind. Hier habe ich den Effekt der Allopolyploidisierung auf das Genom in naszierenden allohexaploiden Mutationsakkumulationslinien (MA) untersucht, die aus einer Kreuzung zwischen den natürlichen allotetraploiden *Arabidopsis suecica* (ÁÁBB) und diploiden *Arabidopsis thaliana* Col-0 (AA). Die Kreuzung dieser beiden Arten hatte zu einem triploiden Hybrid-Individuum geführt, das eine unerwartete spontane „Ganzgenom-Duplikation“ (Matsushita et al. 2012) durchlief. Die Nachkommen wurden verwendet, um 37 MA-Linien zu initiieren, die über Einzelsamenabstieg vermehrt wurden. Meine Analyse konzentrierte sich auf sechs dieser Linien, wobei ich die genomischen Veränderungen untersuchte, die in den resultierenden MA-Linien bis zu 17 Generationen auftraten. Die allohexaploiden Nachkommengenome bestehen aus zwei Subgenomen von *A. suecica* und eine aus *A. thaliana* Col-0. und haben einen komplexen allopolyploiden Karyotyp (AAÁÁBB;  $2n=6x=36$ ;  $4n=20$ ,  $2n=16$ ) mit zwei Kopien des *A. arenosa* Genom (BB) und

---

vier Kopien des *A. thaliana*-Genom (AAÁÁ) aus zwei verschiedenen Akzessionen.

Um die Resequenzierungsanalyse der allohexaploiden Nachkommengenome zu erleichtern, habe ich zuerst eine Genomassemblierung auf Chromosomenebene der *A. suecica* Sue1-Akzession mit einem N50 von 19,06 Mb generiert. Durch die Analyse der Resequenzierungsdaten der neoallohexaploiden Genome in diesen sechs Linien konnte ich einen Einblick in den Evolutionsprozess der frühen Polyploiden gewinnen. Obwohl in den allohexaploiden Daten keine klaren Muster der Rediploidisierung gefunden werden konnten, konnte ein signifikantes Maß an Translokationen von genetischem Material zwischen den homöologen Subgenomen der Allohexaploiden identifiziert werden, was Instabilitäten offenbart, die bei etablierten Polyploiden wie *A. suecica* fehlen oder in viel geringerem Maße vorhanden sind. Eine abdeckungsbasierte Optimierungsmethode wurde verwendet, um den Gewinn oder Verlust des gesamten Chromosoms im *A. thaliana*-Subgenom der allohexaploiden Nachkommen vorherzusagen. Bei 22 der 42 sequenzierten Individuen wurde entweder eine Zunahme oder ein Verlust ganzer Chromosomenkopien festgestellt, und diese konnten anhand der SNP-Allelfrequenzen zwischen den elterlichen Subgenomen validiert werden. Dieser Befund unterstreicht die genomische Instabilität der Neopolyploiden im Vergleich zu etablierten Polyploiden. Bei 19 allohexaploiden Individuen wurden Chromosomenarm-Translokationen zwi-

---

schen Subgenomen identifiziert, und die meisten von ihnen waren innerhalb der Linien fixiert oder teilweise fixiert. Es wurden keine Hinweise auf einen Ausbruch transponierbarer Elemente (TE) oder ein hohes Maß an Proliferation von TE infolge eines „genomischen Schocks“ durch die Allopolyploidisierung beobachtet.

Diese Analyse zeigt, dass bei den frühen Hexaploiden ein hohes Maß an genomischer Variabilität auch zwischen Generationen und Schwestern derselben Linie besteht und gibt einen Einblick in die frühe Evolution großräumiger Mutationen in naszierenden Polyploiden in Richtung Rediploidisierung.