International Journal of Research in Agriculture and Forestry
Volume 7, Issue 5, 2020, PP 10-22
ISSN 2394-5907 (Print) & ISSN 2394-5915 (Online)

SRYAHWA PUBLICATIONS

Genomic Mapping in Polyploid Plants

Werkissa Yali*

Department of Horticulture and Plant Science, Ambo University, Ambo, P.O., Box 19, Ethiopia

Corresponding Author: *Werkissa Yali, Department of Horticulture and Plant Science, Ambo University, Ethiopia*

ABSTRACT

The fusion of two or more genomes within one nucleus results in polyploidy, each cell contains more than two pairs of homologous chromosomes. Polyploidy is an important evolutionary force. Recent estimates suggest that 70% of all angiosperms have experienced one or more episodes of polyploidization. Polyploid organisms carry more than two copies of each chromosome, a condition rarely tolerated in animals but which occurs relatively frequently in the plant kingdom. One of the principal challenges faced by polyploid organisms is to evolve stable meiotic mechanisms to faithfully transmit genetic information to the next generation upon which the study of inheritance is based. Surprising given that a large proportion of domesticated plant species are polyploid. The current polyploid analytic toolbox includes software for assigning marker genotypes (and in particular, estimating the dosage of marker alleles in the heterozygous condition), establishing chromosome-scale linkage phase among marker alleles, constructing (short-range) haplotypes, generating linkage maps, performing genome-wide association studies (GWAS) and quantitative trait locus (QTL) analyses, and simulating polyploid populations. These tools can also help elucidate the mode of inheritance (disomic, polysomic or a mixture of both as in segmental allopolyploids) or reveal whether double reduction and multivalent chromosomal pairing occur. An increasing number of polyploids (or associated diploids) are being sequenced, leading to publicly-available reference genome assemblies. Much work remains in order to keep pace with developments in genomic technologies. However, such technologies also offer the promise of understanding polyploid genomes at a level which yet has remained elusive. Therefore, this paper was prepared with the objectives of reviewing the polyploidy mapping in genomics.

Keywords: *Polyploidy genetics, polyploid software tools, autopolyploid, allopolyploid, polyploidy*

INTRODUCTION

Polyploids are organisms with more than two sets of chromosomes. They are very important in agriculture and play a fundamental role in evolutionary processes, such as differentiation of species (Soltis*et al.*, 2016). One of the most fundamental descriptions of any organism is its ploidy level and chromosome number. Plant scientists in particular will be familiar with this representation of the chromosomal constitution of the sporophyte generation (i.e. the adult plant). The second term in this seemingly simple equation describes the normal complement of chromosomal copies possessed by a member of that species, which is generally 2x ("two times") for diploids. Species where this number exceeds two are collectively referred to as polyploids. Not unexpectedly, each polyploid individual is the product of the fusion of gametes from two parents, just like their diploid counterparts. In other words, polyploids can also be defined as individuals derived from non-haploid gametes (in the case of triploids derived from diploid × tetraploid crosses, only one gamete satisfies this condition). The transmission of non-haploid gametes is one of the main "complexifying" features of polyploidy, leading to a whole range of implications for the genetic analysis of these "hopeful monsters" (Goldschmidt, 1933).

Most genetic advances are made in model organisms, among which self-fertilising diploid species predominate. The ongoing genomics revolution can be seen as a rising tide which has also lifted the polyploid genetics boat, although not quite to the same level as for diploids. It is therefore not surprising that most tools and techniques for molecular- genetic studies are specific to diploids. However, polyploid species are particularly important to mankind in the provision of food, fuel, feed and fibre (not to mention "flowers", if ornamental plant species are also included), making the genetic analysis of polyploid species an important avenue of research for crop improvement. Although a collective term such as "polyploidy" has its uses, it tends to obscure some fundamental

differences between its members. For example, polyploids are generally subdivided into autopolyploids and allopolyploids (Kihara and Ono, 1926).

Autopolyploids arise through genomic duplication within a single species, generally through the production of unreduced gametes (Harlan and De Wet, 1975) and exhibit polysomic inheritance, meaning pairing and recombination can occur between all homologous copies of each chromosome during meiosis. One of the well-studied examples is autotetraploid potato (Solanumtuberosum L.). Allopolyploids, Allopolyploids, on the other hand, are the product of genomic duplication between species (usually through hybridisation involving unreduced gametes (Harlan and De Wet, 1975)) and display disomic inheritance, where more-related chromosome copies ("homologues") may pair and recombine during meiosis, whilst less-related chromosome copies ("homoeologues", also spelled "homeologues" (Glover et al., 2016)) do not. Among allopolyploids, allohexaploid wheat (Triticumaestivum L.) is probably the well-studied. If pairing and recombination between homoeologues occurs to a limited extent, the species may be referred to as "segmental allopolyploid", traditionally deemed to have arisen from hybridisation between very closely-related species (Chester et al., 2012) but which may also be the result of partially- diploidisedauto-polyploidy (Soltis et al., 2016). In many cases, a species cannot be clearly designated as one type or another, leading to uncertainty or debate on the subject (Barker et al., 2016; Doyle and Sherman-Broyles, 2016).

From the perspective of genetics and inheritance, allopolyploids behave much like diploid species and therefore many of the tools developed for diploids can be directly applied. The main challenge that faces allopolyploid geneticists is in distinguishing between homoeologous gene copies carried by sub-genomes within an individual (Kaur et al., 2012;van Dijk et al.,2012). Autopolyploids (and segmental allopolyploids) do not behave like diploids, and are therefore in most need of specialised methods and tools for subsequent genetic studies. In this review we focus primarily on the availability of tools and resources amenable to polysomic (and "mixosomic" (Soltis et al., 2016)) species, with less emphasis on allopolyploid-specific solutions. Experimental populations, in use since Mendel's ground-breaking work (Mendel, 1866), are traditionally derived from a controlled cross between two parental lines of interest (either directly studying the F1 or some later generation). Therefore, this paper was prepared with the objectives of reviewing the polyploidy mapping in genomics.

Literature Review

Polyploidy

Polyploidy is a significant evolutionary process in higher organisms. A polyploid is any organism that carries more than two copies of each chromosome, from the Greek "poly-" meaning much or many, and "–ploid" from "ploos", meaning fold, thus "many-fold". Polyploidy was discovered more than a century ago (Strasburger, 1910) and since then has been a topic of continued interest and debate. Diploidy, the state of possessing two copies of each chromosome, is considered the chromosomal "ground state" or norm for complex organisms. In diploids, parents transmit a single copy of each chromosome to each gamete, thereby re-establishing diploidy in the offspring. In polyploid organisms, more than one copy of each chromosome is transmitted, which is one of the main contributing factors to the complexity of polyploid genetics. There is a tendency for polyploid lineages to return to a diploid conformation over evolutionary timescales, a process termed "diploidisation" or "re-diploidisation" (Le Comber *et al.,* 2010).

In the time scales of interest to breeders and researchers however, polyploidy is effectively a permanent condition. Although the definition of polyploidy is quite unequivocal, there can be some confusion over the classification of a species as polyploid or not, particularly as many complex life-forms were polyploid at some point in the past (Van de Peer *et al.,* 2017). "Paleopolyploids" are species that were true polyploids millions of years ago but have since re-diploidised, and the term "neopolyploids" refers to newly- formed polyploids (Lloyd and Bomblies, 2016), possibly artificially generated for research purposes to understand how polyploids deal with theinitial "genomic shock" of having an extra genome. Neopolyploidy may also refer to recently-formed wild polyploid populations such as Spartinaanglica, an allopolyploid that arose when Spartinaalterniflora was introduced outside its native range and hybridised with local Spartina species (Soltis and Soltis, 2009). Some authors further distinguish "mesopolyploids" as re-diploidised species that underwent whole- genome duplication (WGD) at a less ancient timescale than paleopolyploids and which can be detected by genetic or cytogenetic analyses (Mandáková*et al.,*2010). In this thesis we are principally concerned with extant polyploid species that have not re-diploidised, yet have already passed the (presumed bumpy)

early generations i.e. they are no longer considered neopolyploid.

Among polyploids, two distinct types are generally recognized autopolyploids and allopolyploids. These terms can distinguish or emphasise two features, namely the origin of the polyploid (also termed the "taxonomic" definition), or how its chromosomes behave during meiosis (the "genetic" definition) (Ramsey and Schemske, 2002). Autopolyploids are generally-speaking derived from a single species and exhibit polysomic inheritance. Polysomic inheritance means that all possible combinations of alleles are equally likely to end up in a gamete – although we will return to this question in more detail in the next section as it is one of the fundamental points of interest in this thesis. Allopolyploids on the other hand are derived from at least two species and exhibit disomic inheritance (where disomic means diploid-like inheritance, the result of exclusive pairing and recombination between homologous chromosomes and an absence of pairing and recombination between homoeologous chromosomes). Although also important, they are not the focus of study here. It should be noted that classifying a species as either autopolyploid or allopolyploid is not always straightforward, as demonstrated for example in the debate about the correct classification of the polyploid ancestor ofsoybean (Glycine spp.) (Barker *et al.,* 2016;Doyle and Sherman-Broyles, 2016).

In other words, the taxonomic and genetic definitions do not always neatly overlap, particularly in species with a long history of inter-specific hybridisation among progenitor species of varying relatedness. A large body of polyploid research is aimed at understanding how different polyploid lineages arose, and how these newly-wed genomes adapted and evolved to accommodate each other and their changing environment. There is a third category of polyploid, namely the "segmental allopolyploid" as it was originally termed (Stebbins, 1947). Again this category can be defined from a taxonomic perspective or a genetic perspective – as a hybridisation between two very closely-related species or subspecies, or as a polyploid which demonstrates a meioti cpairing behaviour that cannot be classified as fully disomic or fully polysomic (recently termed "mixosomic" (Soltis*et al.,* 2016)). Throughout this thesis we rely on the genetic definition, as it is the pairing behaviour that influences how homologues recombine, upon which our methods to study inheritance are ultimately based. Although it is interesting to speculate upon how or why such differences arose, in the end we are primarily interested in understanding what happens, as this is the most solid ground upon which to build a model.

Polyploidy occurs in animals, plants and fungi, with the ancestors of all angiosperms and vertebrates thought to have experienced at least two whole-genome duplications (Putnam *et al.,* 2008; Jiao *et al.,* 2011). In the plant kingdom there are numerous examples of extant polyploids. There are fewer known examples of polyploid animals, which some suggest is due to difficulties in re-establishing a balance in chromosomal sex-determination systems following genome duplication (Muller, 1925). However, examples do exist, particularly in amphibians and fish (but much less so in other vertebrates) (Mable *et al.,* 2011). Among fish, it is now well-established that a whole genome duplication (WGD) occurred in the ancestor of all salmonids (e.g. salmon, trout etc.) between 50 and 100 million years ago (Allendorf*et al.,* 2015). Polyploidy is also sometimes artificially induced in animals, for example in Pacific oysters (Crassostreaegigas) (Benabdelmouna and Ledu, 2015) or the silkmoth (Bombyxmori L.). There continues to be debate about whether any polyploid mammals exist, with the initial claim that the Argentinian red vizcacha rat (Tympanoctomysbarrerae) is tetraploid being more recently challenged in light of new data (Evans *et al.,* 2017).

Polyploidy occurs widely among plant species, with recent advances in whole-genome sequencing allowing a detailed analysis of recent and ancient polyploidisation events in an increasingly large number of plant lineages (Van de Peer *et al.,* 2017). In natural populations of plants there is always a small possibility of a new polyploid species arising (usually although not exclusively through unreduced (2n) gametes (Harlan and De Wet, 1975)). In the case of autotetraploids, one possible path for their establishment is through the initial formation of a triploid bridge (from a fusion of n + 2n gametes) (Ramsey and Schemske, 1998; Schinkel*et al.,* 2017). These triploids, although generally infertile, may also produce 2n gametes and hence through selfing or pollination by 2n gametes from diploids, tetraploids may be formed (Ramsey and Schemske, 1998). Note that for a new polyploid lineage to establish and diversify, an even-numbered ploidy is required. An exception to this is when plants exclusively reproduce vegetatively or apomictically, thereby avoiding the disruptions that odd-numbered ploidies pose to balanced meiotic division.

However, such lineages would be expected to evolve more slowly than sexually-reproducing ones (McDonald *et al.,* 2016). The fusion of unreduced pollen with an unreduced egg cell, both from diploid parents, is also theoretically possible ("one-step" tetraploids (Ramsey and Schemske, 1998)).

Induced polyploidy (man-made) can also occur through somatic chromosome doubling, although its status as a means to polyploid formation in natural populations is uncertain (Harlan and De Wet, 1975). An interesting alternative pathway that has only recently been explored is the possibility of polyspermy, where more than one sperm cell fertilises an ovule (Dresselhaus and Johnson, 2018). Interestingly, stressed plants are found to produce more unreduced gametes (such stresses may relate to environmental variables like extreme temperatures, wounding, drought or nutrient deficiency (Ramsey and Schemske, 1998)). Unreduced gametes are thought to arise as a result of defective spindle fibres or cell-plate formation, both of which have been shown to regularly occur at extreme (particularly higher) temperatures (Bomblies*et al.,* 2015)). In most cases, neopolyploid plants are usually at an immediate disadvantage, being un-adapted and reproductively isolated (what is termed "minority cytotype disadvantage" (Husband, 2000)). The speed at which newly-established polyploid lineages prosper and diversify varies, with indications that there may be a significant time lag before this occurs (Schranz*et al.,* 2012).

Polyploids are particularly common among domesticated crops (Salman-Minkov*et al.,* 2016), a fact that has helped drive interest to better understand these species. In many cases polyploidy is deliberately induced – for example modern ornamental breeding often relies on inter-specific hybrids to create novel varieties, which are often "polyploidised" (through colchicine treatment for mitotic polyploidisation, or through selection of 2n gametes) to overcome sterility in the F1. Fruit breeders also generate seedless fruit by crossing parents of different ploidy levels, resulting in sterile fruit-bearing (usually triploid) offspring (Bradshaw, 2016). Many of the native attributes of polyploids may also have endeared them to the early agriculturalists, *e.g.* larger organs (tubers, fruits, flowers *etc.*), also known as the "gigas" effect (Sattler *et al.,* 2016), or their ability to be clonally propagated. We therefore find ourselves in the position of relying on some of the most genetically-complex species to provide us with the basic necessities for life. Examples of some globally-important polyploid crops include allopolyploids such as wheat (*Triticumaestivum*L.), cotton (*Gossypiumhirsutum*L.), coffee (*Coffeaarabica*L.), oilseed rape (*Brassica napus*L.), oats (*Avena sativa* L.), peanut (*Arachishypogaea*L.), strawberry (*Fragaria*× *ananassa*L.) and autopolyploids such as potato (*Solanumtuberosum*L.), alfalfa (*Medicago sativa* L.), sweetpotato (*Ipomoea batatas*L.), leek (*Allium ampeloprasum*L.), blueberry (*Vacciniu-mcorymbosum*L.), chrysanthemum (*Chrysan-themum* spp.) and rose (*Rosa* × *hybrid* L.).

Polyploidy Genotyping

One of the most crucial aspects in the study of polyploid genetics is the generation of accurate genotypic data. However, it is also fraught with difficulties, not least the detection of multiple loci when only a single locus is targeted (Mason, 2015; Limborg*et al.,* 2016). Various technologies exist, with almost all current applications aimed at identifying single nucleotide polymorphisms (SNPs). Although many genomic "service-providers" (*e.g.* companies or institutes that offer DNA sequencing) have their own tools to analyse and interpret raw data, these tools are not always suitable for use with polyploid datasets.

Genotyping Technologies

Although gel-based marker technologies continue to be used and have certain advantages (*e.g.* low costs associated with small marker numbers, requiring only basic laboratory facilities, multi-allelism*etc.*), most studies now rely on SNP markers for genotyping due to their great abundance over the genome, their high-throughput capacity and their low cost per data point. Targeted genotyping such as SNP arrays (a.k.a. "SNP chips") rely on previously-identified and selected polymorphisms, usually identified from a panel of individuals chosen to represent the gene pool under investigation. In contrast, untargeted genotyping generally uses direct sequencing of individuals, albeit after some procedure to reduce the amount of DNA to be sequenced (*e.g.* by exome sequencing (Ng *et al.,* 2009) or target enrichment (Mamanova*et al.,* 2010)).

The disadvantages of targeted approaches have been well explored (particularly regarding ascertainment bias, where the set of targeted SNPs on an array poorly represents the diversity in the samples under investigation due to biased methods of SNP discovery) (Moragues*et al.,* 2010; Lachance and Tishkoff, 2013), although there are advantages and disadvantages to both methods (Mason *et al.,* 2017). Apart from costs, differences exist in the ease of data analysis following genotyping, with sequencing data requiring greater

curation and bioinformatics skills as well as potentially containing more erroneous and missing data (Spindel*et al.,*2013; Jones*et al.,* 2017). In polyploids, SNP arrays have been developed in numerous species, which include both autopolyploid (or predominantly polysomic polyploids) and allopolyploid species.

Physical Maps

Physical mapping involves finding a contiguous series (or 'contig') of cloned DNA fragments which contain overlapping portions of the genome. Arguably, one of the most important "tools" in current genomics studies is access to a high-quality reference genome assembly. Species for which a reference genome assembly exists have even been classified as "model organisms" (Seeb*et al.,* 2011), such is the importance and impact a genome can bring to research on that species. Without a reference sequence available, the scope of genomic research remains limited. For example, genome-wide association studies rely on knowledge of the relative position of SNP markers (usually on a physical map), and many sequencing applications rely on a reference assembly on which to map reads. A reference genome also facilitates the development of molecular markers (*e.g.* primer development), the comparison of results between different genetic studies (by providing a single reference map), as well as allowing comparisons of specific sequences such as genes, enabling prediction of gene function across related species.

Polyploid genomes are by definition more complex than diploid genomes, having multiple copies of each homologous chromosome. Many polyploid species are also outbreeding, leading to increased heterozygosity which is problematic in *de novo* assemblies and necessitates specialized approaches (Kajitani *et al.,* 2014). The most common solution until now has been to sequence a representative diploid species. In the case of allopolyploids, multiple diploid progenitor species are often sequenced instead (*e.g.* peanut (Bertioli*et al.,* 2016)).

Linkage Maps

Although the first genetic linkage map was developed over a hundred years ago (Sturtevant, 1913), their use in genetic and genomic studies has persisted into the "next-generation" era. This can be attributed to a number of factors. A linkage map is a description of the recombination landscape within a species, usually from a single experimental cross of interest. For breeders, knowledge of genetic distance is arguably more important than physical distance, as it reflects the recombination frequencies in inheritance studies as well as describing the extent of linkage drag around loci of interest. Many software for performing quantitative trait locus (QTL) analysis require linkage maps of the markers, not physical maps. This is because co-inheritance of markers and phenotypes within a population are assumed to be coupled – a physical map gives less precise information about the co-inheritance of markers than a linkage map does since physical distances do not directly translate to recombination frequencies (particularly in the pericentromeric regions). Another reason why linkage maps continue to be developed is that they are often the first genomic representation of a species, upon which more advanced representations can be built. They provide useful long-range linkage information over the whole chromosome, which is often missing from assemblies of short sequence reads. This fact has been repeatedly exploited in efforts at connecting and correctly orientating scaffolds during genome assembly projects (Bartholomé*et al.,* 2015).

Many linkage maps in polyploids have been based exclusively on 1:1 segregating markers, also known as simplex markers (because the segregating allele is in simplex condition (one copy) in one of the parents only). These markers possess a number of advantages over other marker segregation types, but also some distinct disadvantages. In their favour, coupling-phase simplex markers in polyploid species behave just like they would in diploid species, regardless of the mode of inheritance involved (repulsion-phase recombine-tion frequency estimates are not invariant across ploidy levels or modes of inheritance, but exert less influence on map construction due to lower LOD scores). The advantage of this is clear: in unexplored polyploid species for which the mode of inheritance is uncertain, simplex markers allow an "assumption-free" linkage map to be created, following which the mode of inheritance can be further explored. The only exception to this is if double reduction occurs, *i.e.* when a segment of a single chromosome gets transmitted with its sister chromatid copy to an offspring, a consequence of multivalent pairing and a particular sequence of chromatid segregation and division during meiosis (Mather, 1935). Double reduction occurs randomly in polysomic species and only introduces a small bias into recombination frequency estimates (Bourke *et al.,* 2015). This means that, ignoring the possible influence of double reduction, diploid mapping software can generally be used for simplex marker sets at any ploidy level and for

any type of meiotic pairing behaviour opening up a very wide range of diploid-specific software options (Cheema and Dicks, 2009).

Genome-Wide Association Studies

Genome-wide association studies (GWAS) have emerged as a powerful tool for detecting causative loci underlying phenotypic traits. They have been particularly popular in species where the generation of experimental populations is problematic (such as humans). GWAS has been readily adopted across a broad spectrum of species since then, due to the promise of increased mapping resolution, a more diverse sampling of alleles and a simplicity in population creation (no crossing required) (Bernardo, 2016). There are certain disadvantages though, particularly in how rare (and potentially important) variants can be missed (Ott et al., 2015) and the confounding effect of population structure on results (Korte and Farlow, 2013). Nevertheless, GWAS continues to be an important analytical option to help shed greater light on genotype – phenotype associations.

QTL Analysis

The term "QTL analysis" usually refers to studies that aim to detect regions of the genome (so-called quantitative trait loci (Geldermann, 1975)) that have a significant statistical association with a trait in specifically-constructed experimental populat-ions. These populations are most often created by crossing two contrasting parental lines ("bi-parental" populations), although there is increasing interest in using more complex population designs in order to increase the range of alleles and genetic backgrounds being studied (Huang *et al.,* 2015).

As already discussed, there is great difficulty in developing inbred lines by repeatedly selfing polyploids due to the sampling of alleles during polyploid gamete formation (in a diploid this sampling generates (21)=2 combinations; for a tetraploid this rises to (42)=6 and in a hexaploid (63)=20 combinations, resulting in protracted heterozygosity not to mention the problem of inbreeding depression associated with many outcrossing polyploid species. Therefore, most QTL analyses in polyploid species have been performed using the directly-segregating F1 progeny of a cross between heterozygous parents (a "full sib" population). This leads to poor resolution of QTL positions when compared to the more popular diploid inbred populations like RILs *etc.*, as well as the fact that populations must be vegetatively propagated if replication over years or different growing environments is desired. For many polyploid species, vegetative propagation is indeed possible (Herben et al., 2017) and F1 populations have the added advantage of being relatively quick and simple to develop, while, because of a generally high level of hetero-zygosity, many loci will be segregating in the F1. Therefore, despite their drawbacks, F1 populations remain the bi-parental population of choice for mapping studies.

The methods for QTL analysis in diploid species have become increasingly convoluted (van Eeuwijk et al., 2010); in polyploid species such theoretical complexities have yet to be attempted, given the more immediate difficulties in accurately genotyping as well as modelling polyploid inheritance. Just like for linkage mapping and GWAS, the range of software tools available for QTL analysis in polyploids remains rather limited, although there are a number of recent developments that are helping transform the field. One of the only dedicated software for tetraploid QTL analysis is the already-mentioned Tetraploid Map software (Hackett et al., 2007). This software enables interval mapping to be performed in autotetraploid F1 populations (as well as a simple single-marker ANOVA test), using a restricted range of markers (1x0, 2x0 and 1x1 markers only, where 1x0 denotes a marker dosage of 1 in one parent and 0 in the other, *etc.*). Although still available, it has been superseded by the TetraploidSNPMap software (Hackett et al., 2017).

TetraploidSNPMap (TSNPM) uses SNP dosage data to either construct a linkage map (as already described) or perform QTL interval mapping. In contrast to its predecessor, TSNPM can analyse all marker segregation types, and allows the user to explore different QTL models at detected peaks. At its core is an algorithm to determine identity-by-descent (IBD) probabilities for the offspring of the population, which are then used in a weighted regression performed across the genome. An independent software tool that has been developed to determine IBD probabilities in tetraploids is TetraOrigin (Zheng et al., 2016), implemented in the Mathematica programming language. Tetra Origin relaxes the assumption of random bivalent pairing during meiosis (which TSNPM employs) to allow for both preferential chromosomal pairing as well as multivalent formation and the possibility of double reduction. Although not programmed in a user-friendly format like TSNPM, it is relatively straightforward to use, taking an integrated linkage map and marker dosage matrix as input. It does not perform QTL analysis directly, but the resulting IBD probabilities can then be used to model genotype effects in a QTL scan either using a weighted regression approach like TSNPM, or in

a linear mixed model setting. IBD probabilities allow interval mapping since they can be interpolated at any desired intervals on the linkage map.

For ploidy levels other than tetraploid, there are currently no dedicated software tools available for QTL analysis or IBD probability estimation. Single-marker approaches such as ANOVA on the marker dosages (assuming additivity – various dominant models could also be explored; see *e.g.* (Rosyara et al., 2016)) are of course possible and require access to basic statistical software packages such as R (or even Excel). However, such approaches are not ideal – they are only effective if marker alleles are closely linked in coupling with QTL alleles, and offer no ability to predict the QTL segregation type or mode of gene action as is done for example in TSNPM (Hackett *et al.,* 2017). Approaches such as pedigree-informed analyses, implemented for diploids in the FlexQTL software (Bink*et al.,* 2008), could overcome some of the limitations imposed by the restrictions on population types in software for polyploids. However, it may take some time before such tools become translated to the polyploid level.

Evolutionary Implications

The genetic and evolutionary implications of recurrent polyploidization and genome reshuffling are obvious in that both processes represent important sources of genetic variation. Population-level genetic studies of polyploidy plants and animals indicate that polyploidization should no longer be viewed as a rare event producing a polyploid species of unique origin and uniform genotype. Instead, polyploid species can maintain high levels of segregating genetic variation through the incorporation of genetic diversity from multiple populations of their diploid progenitors. Polyploid genotypes ultimately come into contact via migration and hybridize – with subsequent segregation and recombination generating even more genetic complexity. Concomitantly, the genome reshuffling of polyploid genomes is an important additional source of genetic diversity in polyploid species (Soltis and Soltis, 1999).

Polyploidy as Transilience

Templeton (1980) suggested that diploid speciation involves transilience, a period during which the genome is more amenable to or tolerant of change, such as recombination. Growing evidence suggests that polyploidy should also be viewed as transilience. The extensive genomic change detected in only five generations in synthetic allopolyploid Brassica, as well as the chromosomal rearrangements detected in hybrid Nicotiana (Leitch and Bennett, 1997), support this view. The presence of the same chromosomal changes throughout the populations of a polyploid species suggests that genome reorganization accompanied speciation, or occurred shortly thereafter. Species-wide intergenomic translocat- ions have been detected in several angiosperms, including tetraploid wheat and tobacco. Other chromosomal changes detected in tobacco and wheat are not present in all populations analysed and apparently represent subsequent chromoso- mal divergence that occurred in different populations after speciation (Soltis and Soltis, 1999).

The chromsomal and gene-level changes reviewed above are made possible by polyploidization; that is, polyploidy could represent a source of novel evolutionary processes. Rather than being stable, non-interacting enti- ties, two or more divergent genomes in a common ‘polyploid’ nucleus could facilitate intergenomic interactions, ultimately resulting in new chromosomal and gene arrange- ments. Polyploidization might be a source of genomic stress that facilitates rapid evolution.

Polyploidy and Transposable Elements

Transposable elements (TEs) might facilitate rapid genome restructuring after polyploidization. (Matzke and Matzke, 1998) argue that polyploidy permits extensive gene modification by TEs because, by nature, polyploid genomes contain duplicate copies of all genes; hence, they are well buffered from the deleterious consequences of transposi- tion. Transposable elements will tend to multiply and be maintained in polyploids because the additional copies of genes they maintain will compensate for the loss of altered expression of genes that might result from TE insertion. The end result could be higher genomic restructuring in polyploids compared with their diploid progenitors. Recent studies suggest the spread of DNA repeat fam- ilies from one parental diploid genome to the other in allopolyploid cotton, Gossypium. In cotton, most dispersed repeat families are restricted to A-genome diploids and are absent from D-genome diploids. However, in the allo- tetraploids (which combine the A and D genomes), the A-genome repeats have spread to the D genome, perhaps by replicative transposition (Spring, 1997).

Transposable elements might also have been the driv- ing force in the evolution of gene silencing mechanisms, such as methylation and heterochromatin formation, throughout eukaryotes in general. These ‘global repres- sion’ mechanisms

might have evolved as adaptive responses to the selfish drive of TEs to expand in number in a host genome (McDonald, 1998). (Matzke and Matzke, 1998) argue that if TEs are indeed the primary targets of methylation and other global repression mechanisms and that polyploids tolerate transposition because of their duplicate genes, then it follows that polyploid genomes will not only contain moreTEs than diploid genomes, but will also be more highly methylated. (Matzke and Matzke, 1998) suggest that a rough cor- relation exists. Widespread ('global') methylation is found in vertebrate genomes, which represent several rounds of polyploidization, as well as in polyploid plant genomes, which contain a high number of TEs. In contrast, 'fractional' (partial), rather than global, methylation occurs in inverte- brates and true diploid plants. For example, Arabidopsis, which is diploid and has a small genome, also has a small number of TEs, whereas ¬50% of the maize genome (an ancient polyploid) is composed of interspersed repetitive DNAs, primarily nested retrotransposons that insert between genes.

Genomic Prediction and Genomic Selection

There has been much attention given to the advantages of using *all* marker data to help predict phenotypic performance, rather than focusing on single markers (or haplotypes) that are linked to QTL as was previously advocated. The motivation behind this is clear – many of the most important traits in domesticated animal and plant species are highly quantitative, with far too many small-effect loci present to be able to tag them all with single markers (Bernardo, 2008). One of the most important traits in any breeding program is also a famously quantitative trait: yield. It has been suggested that despite many years of phenotypic selection, crop yield in tetraploid potato has essentially remained unchanged (Slater et al., 2016). This is a remarkable indictment of traditional selection methods, yet offers much-needed impetus for the development and deployment of new paradigms in breeding for quantitative traits.

Genomic prediction first arose in animal breeding circles (Meuwissen et al., 2001), where the concept of estimating breeding values from known pedigrees was already well-established. However, the estimation of breeding values in polyploid species requires special consideration due to the complexity of polysomic inheritance and the possibility of double reduction. In practice, breeding values are usually estimated using restricted maximum likelihood (REML) to solve mixed model equations, requiring the generation of an inverse additive relationship matrix ***A-1***, also called the numerator relationship matrix. The form of ***A-1*** depends on, among other things, whether the inheritance is polysomic or disomic, and whether double reduction occurs (Kerr et al., 2012; Hamilton and Kerr, 2017). Recently, the R package polyAinv was released which computes the appropriate ***A-1*** as well as the kinship matrix ***K*** and the inbreeding coefficients *F* (Hamilton and Kerr, 2017). However, in one study of nine common traits in autotetraploid potato, the inclusion of double reduction, or even the adoption of an autotetraploid-appropriate relationship matrix was found to have a minimal impact on the results (Slater et al., 2014). Studies which ignore the specific complexities of autopolyploids may still benefit from genomic prediction and selection, as for example was demonstrated in tetraploid potato (Sverrisdóttir et al., 2017). Commonly-used software tools for estimating breeding values at the diploid level include ProGeno (Maenhout, 2018) and ASreml (V.S.N. International, 2018) which could be suitable for polyploid breeding programs, although this has yet to be conclusively demonstrated.

Mode of Inheritance

The term "mode of inheritance" refers to the randomness of meiotic pairing processes that give rise to gametes, and is often used to distinguish between disomic (diploid-like) inheritance, and polysomic (all allele combinations equally possible) inheritance. As alluded to already, intermediate modes of inheritance are theoretically possible if partially-preferential pairing occurs between homologues, resulting in on average more recombinations between certain homologues, and less between others (putative homoeologues). This intermediate inheritance pattern, originally termed segmental allopolyploidy (Stebbins, 1947) and more recently termed mixosomy (Soltis et al., 2016), poses additional challenges over those of purely polysomic or disomic behaviour. One of the main complications is the lack of fixed segregation ratios to test markers against (Allendorf and Danzmann, 1997), which is often used as a measure of marker quality (Pompanon et al., 2005).

Currently there are no dedicated tools available to ascertain the most likely mode of inheritance in polyploids. Some "traditional" approaches to predict the mode of inheritance are summarised in (Bourke et al., 2017), many of which are relatively straightforward to implement using a statistical programming environment like R (R Core Team, 2016). In that study, TetraOrigin (Zheng et al., 2016) was used to estimate the most likely pairing

configuration that gave rise to each offspring in an F1 tetraploid population. This enabled the authors to test whether there were deviations from the expected patterns of homologue pairing under a tetrasomic model (Bourke et al., 2017). A simple alternative using closely-linked repulsion-phase simplex marker pairs was also proposed and has been implemented in the polymapRpackage. Apart from preferential pairing, TetraOrigin can also predict whether marker data arose from bivalent or multivalent pairing during meiosis, facilitating an analysis of the distribution of double reduction products. However, apart from its restriction to tetraploid data, an integrated linkage map is required before TetraOrigin can be employed. In severe cases of mixosomy, it is not obvious how a reliable linkage map should be generated. Corrections for mixosomy in a tetraploid linkage analysis are possible in polymapR, but in extreme cases marker clustering will also be affected, making map construction quite challenging. A confounding complication is the possibility of variable chromosome counts (aneuploidy), as for example encountered in sugarcane (Grivet et al., 1996; Grivet and Arruda, 2002) or in ornamentals such as *Alstroemeria* (Buitendijk et al., 1997), which makes the diagnosis of the mode of inheritance even more difficult.

Summary and Recommendation

Polyploidy has played a major role in the evolution of many eukaryotes. Studies have dramatically reshaped views of polyploid evolution, demonstrating that most polyploid species examined, both plant and animal, have formed recurrently from different populations of their progenitors. Populations of independent origin can subsequently come into contact and hybridize, generating new genotypes. Because of the frequency of polyploidy in plants, many recognized species are probably polyphyletic. Extensive and rapid genome restructuring can occur after polyploidization. Such changes can be mediated by transposons. Polyploidization could represent a period of transilience, during which genomic changes occur, potentially producing new gene complexes and facilitating rapid evolution.

The discoveries of the past few years set the stage for a new series of questions surrounding the genetic and genomic aspects of polyploid evolution. It involves the evolutionary and ecological consequences of multiple origins of polyploid species. Progress in the area of comparative genome organization will be facilitated greatly by large-scale genomic projects already under way for model organisms such as maize, wheat, Arabidopsis and members of Solanaceae. These studies will provide fine-scale genomic maps for polyploid plants including maize, the polyploid brassicas, potato and tobacco. Such data will provide additional in- sights into genome evolution in polyploids. It is important, however, that comparative genetics be applied not only to crops and close relatives, but also to diploids and their polyploid derivatives in natural populations.

A move from diploid-based reference genomes to fully polyploid (and haplotype-resolved) reference genomes would also help broaden the boundaries of polyploid genetics away from the diplo-centric view of genomics which currently dominates. Although there have been many exciting discoveries and developments in polyploid genetics in the past decade or more, we feel its golden age has yet to arrive, an age which will be heralded all the sooner by the provision of robust and user-friendly tools for the genetic dissection of these fascinating groups of organisms.

References

[1] Allendorf, F.W., Bassham, S., Cresko, W.A., Limborg, M.T., Seeb, L.W., and Seeb, J.E. (2015). Effects of crossovers between homeologs on inheritance and population genomics in polyploid-derived salmonid fishes. Journal of Heredity 106, 217-227.

[2] Allendorf F. W., Danzmann R. G. (1997). Secondary tetrasomic segregation of MDH-B and preferential pairing of homeologues in rainbow trout. *Genetics* 145 1083–1092.

[3] Barker, M.S., Arrigo, N., Baniaga, A.E., Li, Z., and Levin, D.A. (2016). On the relative abundance of autopolyploids and allopolyploids. New Phytologist 210, 391-398.

[4] Bartholomé, J., Mandrou, E., Mabiala, A., Jenkins, J., Nabihoudine, I., Klopp, C., et al. (2015). High-resolution genetic maps of Eucalyptus improve Eucalyptus grandis genome assembly. New Phytologist 206, 1283-1296

[5] Behrouzi, P., and Wit, E.C. (2017a). De novo construction of q-ploid linkage maps using discrete graphical models. arXiv preprint arXiv:1710. 01063.

[6] Behrouzi P., Wit E. C. (2017b). netgwas: an R package for network-based genome-wide association studies. arXiv preprint arXiv:1710 01236.

[7] Benabdelmouna, A., and Ledu, C. (2015). Autotetraploid Pacific oysters (Crassostrea gigas) obtained using normal diploid eggs: induction and impact on cytogenetic stability. Genome 58, 333-348.

[8] Bennett, M.D. Perspectives on polyploidy in plants Ancient and neo. Biol. J. Linn. Soc.2004, 82, 411-423.

[9] Bernardo R. (2008). Molecular markers and selection for complex traits in plants: learning from the last 20 years. *Crop Sci.* 48 1649–1664. 10.2135 /cropsci2008.03.0131

[10] Bernardo, R. (2016). Bandwagons I, too, have known. Theoretical and Applied Genetics 129, 2323-2332.

[11] Bertioli, D.J., Cannon, S.B., Froenicke, L., Huang, G., Farmer, A.D., Cannon, E.K., et al. (2016). The genome sequences of Arachis duranensis and Arachis ipaensis, the diploid ancestors of cultivated peanut. Nature Genetics 48, 438-446.

[12] Bomblies, K., Higgins, J.D., and Yant, L. (2015). Meiosis evolves: adaptation to external and internal environments. New Phytologist 208, 306-323.

[13] Bourke, P.M. (2014). QTL analysis in polyploids. MSc thesis, Wageningen University, Wageningen.

[14] Bourke, P.M., Arens, P., Voorrips, R.E., Esselink, G.D., Koning-Boucoiran, C.F.S., Van 'T Westende, W.P.C., et al. (2017). Partial preferential chromosome pairing is genotype dependent in tetraploid rose. The Plant Journal 90, 330-343.

[15] Bourke, P.M., Voorrips, R.E., Kranenburg, T., Jansen, J., Visser, R.G., and Maliepaard, C. (2016). Integrating haplotype-specific linkage maps in tetraploid species using SNP markers. Theoretical and Applied Genetics 129, 2211-2226.

[16] Bourke, P.M., Voorrips, R.E., Visser, and Maliepaard, C. (2015). The Double Reduction Landscape in Tetraploid Potato as Revealed by a High-Density Linkage Map. Genetics 201, 853-863.

[17] Bradshaw, J.E. (2016). Plant Breeding: Past, Present and Future. New York: Springer.

[18] Buitendijk J. H., Boon E. J., Ramanna M. S. (1997). Nuclear DNA content in twelve species of *Alstroemeria* L. and some of their hybrids. *Ann. Bot.* 79 343–353. 10.1006/anbo.1996.0345

[19] Cheema, J., and Dicks, J. (2009). Computational approaches and software tools for genetic linkage map estimation in plants. Briefings in bioinformatics 10, 595-608.

[20] Chen, Z.J. Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. Annu. Rev. Plant. Biol. 2007, 58, 377–406.

[21] Chester, M., Gallagher, J.P., Symonds, V.V., Da Silva, A.V.C., Mavrodiev, E.V., Leitch, A.R., et al. (2012). Extensive chromosomal variation in a recently formed natural allopolyploid species, Tragopogon miscellus (Asteraceae). Proceedings of the National Academy of Sciences 1176-1181.

[22] Comai, L. The advantages and disadvantages of being polyploid. Nat. Rev. Genet. 2005, 6, 836–846.

[23] Doyle, J.J., and Sherman-Broyles, S. (2016). Double trouble: taxonomy and definitions of polyploidy. New Phytologist, 213, 487-493.

[24] Dresselhaus, T., and Johnson, M.A. (2018). Reproduction: Plant Parentage à Trois. Current Biology 28, R28- R30.

[25] Evans, B.J., Upham, N.S., Golding, G.B., Ojeda, R.A., and Ojeda, A.A. (2017). Evolution of the Largest Mammalian Genome. Genome Biology and Evolution 9, 1711-1724.

[26] Galloway, L.F., and Etterson, J.R. (2007). Inbreeding depression in an autotetraploid herb: a three cohort field study. New Phytologist 173, 383-392.

[27] Galloway, L.F., Etterson, J.R., and Hamrick, J.L. (2003). Outcrossing rate and inbreeding depression in the herbaceous autotetraploid, Campanula americana. Heredity, 90, 308-315.

[28] Geldermann H. (1975). Investigations on inheritance of quantitative characters in animals by gene markers I. Methods. *TAG Theor. Appl. Genet.* 46 319–330. 10.1007/BF00281673

[29] Glover, N.M., Redestig, H., and Dessimoz, C. (2016). Homoeologs: what are they and how do we infer them? Trends in plant science 21, 609-621.

[30] Goldschmidt, R. (1933). Some aspects of evolution. Science 78, 539-547.

[31] Grandke, F., Ranganathan, S., Van Bers, N., De Haan, J.R., and Metzler, D. (2017). PERGOLA: fast and deterministic linkage mapping of polyploids. BMC Bioinformatics 18:12.

[32] Grandke, F., Singh, P., Heuven, H.C.M., De Haan, J.R., and Metzler, D. (2016). Advantages of continuous genotype values over genotype classes for GWAS in higher polyploids: a comparative study in hexaploid chrysanthemum. BMC Genomics 17:672.

[33] Grivet L., Arruda P. (2002). Sugarcane genomics: depicting the complex genome of an important tropical crop. *Curr. Opin. Plant Biol.* 5 122–127. 10.1016/S1369-5266(02)00234-0

[34] Grivet L., D'hont A., Roques D., Feldmann P., Lanaud C., Glaszmann J. C. (1996). RFLP mapping in cultivated sugarcane (*Saccharum* spp.): genome organization in a highly polyploid and aneuploid interspecific hybrid. *Genetics* 142 987–1000

[35] Hackett, C.A., and Broadfoot, L.B. (2003). Effects of genotyping errors, missing values and segregation distortion in molecular marker data on the construction of linkage maps. Heredity 90, 33-38

[36] Hackett, C.A., and Luo, Z. (2003). TetraploidMap: construction of a linkage map in autotetraploid species. Journal of Heredity 94, 358-359.

[37] Hackett, C.A., Boskamp, B., Vogogias, A., Preedy, K.F., and Milne, I. (2017). TetraploidSNPMap: Software for Linkage Analysis and QTL Mapping in Autotetraploid Populations Using SNP Dosage Data. Journal of Heredity 108, 438-442.

[38] Hackett, C.A., Bradshaw, J., and McNicol, J. (2001). Interval mapping of quantitative trait loci in autotetraploid species. Genetics 159, 1819-1832.

[39] Hackett, C.A., Bradshaw, J., Meyer, R., McNicol, J., Milbourne, D., and Waugh, R. (1998). Linkage analysis in tetraploid species: a simulation study. Genetics Research 71, 143-153.

[40] Hackett, C.A., Bradshaw, J.E., and Bryan, G.J. (2014). QTL mapping in autotetraploids using SNP dosage information. Theoretical and Applied Genetics 127, 1885-1904.

[41] Hackett, C.A., Mclean, K., and Bryan, G.J. (2013). Linkage Analysis and QTL Mapping Using SNP Dosage Data in a Tetraploid Potato Mapping Population. PLoS One 8, e63939.

[42] Hackett, C.A., Milne, I., Bradshaw, J.E., and Luo, Z. (2007). TetraploidMap for Windows: linkage map construction and QTL mapping in autotetraploid species. Journal of Heredity 98, 727-729.

[43] Hackett, C.A., Pande, B., and Bryan, G. (2003). Constructing linkage maps in autotetraploid species using simulated annealing. Theoretical and Applied Genetics 106, 1107-1115.

[44] Haldane, J.B. (1930). Theoretical genetics of autopolyploids. Journal of Genetics 22, 359-372.

[45] Hamilton M. G., Kerr R. J. (2017). Computation of the inverse additive relationship matrix for autopolyploid and multiple-ploidy populations. *Theor. Appl. Genet.* 131 851–860. 10.1007/s00122-017-3041-y

[46] Harlan, J.R., and De Wet, J.M.J. (1975). On Ö. Winge and a prayer: the origins of polyploidy. The botanical review 41, 361-390.

[47] He Y., Xu X., Tobutt K. R., Ridout M. S. (2001). Polylink: to support two-point linkage analysis in autotetraploids. *Bioinformatics* 17 740–741. 10.1093/bioinformatics/17.8.740

[48] Hendry, A.P.; Bolnick, D.I.; Berner, D.; Peichel, C.L. Along the speciation continuum in sticklebacks. J. Fish. Biol. 2009, 75, 2000–2036.

[49] Herben T., Suda J., Klimešová J. (2017). Polyploid species rely on vegetative reproduction more than diploids: a re-examination of the old hypothesis. *Ann. Bot.* 120 341–349. 10.1093/aob/mcx009

[50] Huang B. E., Verbyla K. L., Verbyla A. P., Raghavan C., Singh V. K., Gaur P., et al. (2015). MAGIC populations in crops: current status and future prospects. *Theor. Appl. Genet.* 128 999–1017. 10.1007/s00122-015-2506-0

[51] Husband, B.C. (2000). Constraints on polyploid evolution: a test of the minority cytotype exclusion principle. Proceedings of the Royal Society of London. Series B: Biological Sciences 267, 217-223.

[52] Jannoo N., Grivet L., Dookun A., D'hont A., Glaszmann J. C. (1999). Linkage disequilibrium among modern sugarcane cultivars. *Theor. Appl. Genet.* 99 1053–1060.

[53] Jiao, Y., Wickett, N.J., Ayyampalayam, S., Chanderbali, A.S., Landherr, L., Ralph, P.E., et al. (2011). Ancestral polyploidy in seed plants and angiosperms. Nature 473, 97-100.

[54] Jones, D.B., Jerry, D.R., Khatkar, M.S., Raadsma, H.W., Van Der Steen, H., Prochaska, J., et al. (2017). A comparative integrated gene-based linkage and locus ordering by linkage disequilibrium map for the Pacific white shrimp, Litopenaeus vannamei. Scientific Reports 7.

[55] Kajitani, R., Toshimoto, K., Noguchi, H., Toyoda, A., Ogura, Y., Okuno, M., et al. (2014). Efficient de novo assembly of highly heterozygous genomes from whole-genome shotgun short reads. Genome Research 24, 1384-1395.

[56] Kaur, S., Francki, M.G., and Forster, J.W. (2012). Identification, characterization and interpretation of single- nucleotide sequence variation in allopolyploid crop species. Plant biotechnology journal 10, 125- 138.

[57] Kerr R. J., Li L., Tier B., Dutkowski G. W., Mcrae T. A. (2012). Use of the numerator relationship matrix in genetic analysis of autopolyploid species.124 1271–1282. 10.1007/s00122-012-1785-y

[58] Kihara, H., and Ono, T. (1926). Chromosomenzahlen und systematische Gruppierung der Rumex-Arten. Zeitschrift für Zellforschung und mikroskopische Anatomie 4, 475-481.

[59] Kim S., Plagnol V., Hu T. T., Toomajian C., Clark R. M., Ossowski S., et al. (2007). Recombination and linkage disequilibrium in *Arabidopsis thaliana*. *Nat. Genet.* 39 1151–1155. 10.1038/ng2115

[60] Korte, A., and Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: a review. Plant methods 9:29.

[61] Lachance, J., and Tishkoff, S.A. (2013). SNP ascertainment bias in population genetic analyses: why it is important, and how to correct it. Bioessays 35, 780-786.

[62] Le Comber, S., Ainouche, M., Kovarik, A., and Leitch, A. (2010). Making a functional diploid: from polysomic to disomic inheritance. New Phytologist 186, 113-122.

[63] Leitch, I.J. and Bennett, M.D. (1997) Polyploidy in angiosperms, Trends Plant Sci. 2, 470–476

[64] Limborg, M.T., Seeb, L.W., and Seeb, J.E. (2016). Sorting duplicated loci disentangles complexities of polyploid genomes masked by genotyping by sequencing. Molecular ecology 25, 2117-2129.

[65] Lloyd, A., and Bomblies, K. (2016). Meiosis in autopolyploid and allopolyploid Arabidopsis. Current opinion in plant biology 30, 116-122.

[66] Mable, B., Alexandrou, M., and Taylor, M. (2011). Genome duplication in amphibians and fish: an extended synthesis. Journal of Zoology 284, 151-182.

[67] Mamanova, L., Coffey, A.J., Scott, C.E., Kozarewa, I., Turner, E.H., Kumar, A., et al. (2010). Target- enrichment strategies for next-generation sequencing. Nature Methods 7, 111-118

[68] Mandáková, T., Joly, S., Krzywinski, M.,

Mummenhoff, K., and Lysak, M.A. (2010). Fast diploidization in close mesopolyploid relatives of Arabidopsis. The Plant Cell 22, 2277-2290.

[69] Mason, A.S. (2015). "Challenges of Genotyping Polyploid Species," in Plant Genotyping: Methods and Protocols, ed. J. Batley. (New York, NY: Springer New York), 161-168.

[70] Mason, A.S., Higgins, E.E., Snowdon, R.J., Batley, J., Stein, A., Werner, C., et al. (2017). A user guide to the Brassica 60K Illumina Infinium™ SNP genotyping array. Theoretical and Applied Genetics 130, 621-633.

[71] Mather, K. (1935). Reductional and equational separation of the chromosomes in bivalents and multivalents. Journal of Genetics 30, 53-78.

[72] Matzke, M.A. and Matzke, A.J.M. (1998) Polyploidy and transposons, Trends Ecol. Evol. 13, 241

[73] Mayrose, I.; Zhan, S.H.; Rothfels, C.J.; Magnuson-Ford, K.; Barker, M.S.; Rieseberg, L.H.; Otto, S.P. Recently formed polyploid plants diversify at lower rates. Science 2011, 333, doi:10.1126/science.1207205.

[74] McDonald, J.F. (1998) Transposable elements, gene silencing, and macroevolution, Trends Ecol. Evol. 13, 94–95

[75] McDonald, M.J., Rice, D.P., and Desai, M.M. (2016). Sex speeds adaptation by altering the dynamics of molecular evolution. Nature 531, 233-236.

[76] Mendel, J.G. (1866). Versuche über Pflanzen-hybriden. Verhandlungen des naturforschenden Vereines in Brünn Bd. IV Abhandlungen, 3–47.

[77] Meuwissen T. H. E., Hayes B. J., Goddard M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157 1819–1829.

[78] Moragues, M., Comadran, J., Waugh, R., Milne, I., Flavell, A., and Russell, J.R. (2010). Effects of ascertainment bias and marker number on estimations of barley diversity from high-throughput SNP genotype data. Theoretical and applied genetics 120, 1525-1534.

[79] Muller, H. (1925). Why polyploidy is rarer in animals than in plants.The American Naturalist 59,346-353.

[80] Ng, S.B., Turner, E.H., Robertson, P.D., Flygare, S.D., Bigham, A.W., Lee, C., et al. (2009). Targeted capture and massively parallel sequencing of 12 human exomes. Nature 461, 272-276.

[81] Osborn, T.C.; Pires, J.C.; Birchler, J.A.; Auger, D.L.; Chen, Z.J.; Lee, H.-S.; Comai, L.; Madlung, A.; Doerge, R.W.; Colot, V.; et al. Understanding mechanisms of novel gene expression in polyploids. Trends Genet. 2003, 19, 141–147.

[82] Ott, J., Wang, J., and Leal, S.M. (2015). Genetic linkage analysis in the age of whole-genome sequencing. Nature Reviews Genetics 16, 275-284.

[83] Pompanon F., Bonin A., Bellemain E., Taberlet P. (2005). Genotyping errors: causes, consequences and solutions. *Nat. Rev. Genet.* 6 847–859. 10.1038/nrg1707

[84] Preedy, K.F., and Hackett, C.A. (2016). A rapid marker ordering approach for high-density genetic linkage maps in experimental autotetraploid populations using multidimensional scaling. Theoretical and Applied Genetics 129, 2117-2132.

[85] Putnam, N.H., Butts, T., Ferrier, D.E., Furlong, R.F., Hellsten, U., Kawashima, T., et al. (2008). The amphioxus genome and the evolution of the chordate karyotype. Nature 453, 1064-1071.

[86] Ramsey, J., and Schemske, D.W. (2002). Neopolyploidy in flowering plants. Annual review of ecology and systematics, 589-639.

[87] Ramsey, J Schemske, D.W. Neopolyploidy in flowering plants. Annu. Rev. Ecol. Syst. 2002,33, 589–639.

[88] Ramsey, J.; Schemske, D.W. Pathways, mechanisms, and rates of polyploid formation in flowering plants. Annu. Rev. Ecol. Syst. 1998, 29, 467–501.

[89] R Core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

[90] Rosyara U. R., De Jong W. S., Douches D. S., Endelman J. B. (2016). Software for genome-wide association studies in autopolyploids and its application to potato. *Plant Genome* 9 1–10. 10.3835/plantgenome2015.08.0073

[91] Salman-Minkov, A., Sabath, N., and Mayrose, I. (2016). Whole-genome duplication as a key factor in crop domestication. Nature plants 2:16115.

[92] Sattler, M.C., Carvalho, C.R., and Clarindo, W.R. (2016). The polyploidy and its key role in plant breeding. Planta 243, 281-296.

[93] Schinkel, C.C.F., Kirchheimer, B., Dullinger, S., Geelen, D., De Storme, N., and Hörandl, E. (2017). Pathways to polyploidy: indications of a female triploid bridge in the alpine species Ranunculus kuepferi (Ranunculaceae). Plant Systematics and Evolution 303, 1093-1108.

[94] Schranz, M.E., Mohammadin, S., and Edger, P.P. (2012). Ancient whole genome duplications, novelty and diversification: Current opinion in plant biology 15, 147- 153

[95] Seeb, J., Carvalho, G., Hauser, L., Naish, K., Roberts, S., and Seeb, L. (2011). Single- nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. Molecular Ecology Resources 11, 1-8.

[96] Shen J., Li Z., Chen J., Song Z., Zhou Z., Shi Y. (2016). SHEsisPlus, a toolset for genetic studies on polyploid species. *Sci. Rep.* 6:24095. 10.1038/srep24095

[97] Slater A. T., Cogan N. O., Forster J. W., Hayes B. J., Daetwyler H. D. (2016). Improving genetic gain with genomic selection in autotetraploid potato.

Plant Genome 9 1–15. 10.3835/plantgenome 2016.02.0021

[98] Slater A. T., Wilson G. M., Cogan N. O., Forster J. W., Hayes B. J. (2014). Improving the analysis of low heritability complex traits for enhanced genetic gain in potato. *Theor. Appl. Genet.* 127 809–820. 10.1007/s00122-013-2258-7

[99] Soltis, D.E., Buggs, R.J.A.; Doyle, J.J.; Soltis, P.S. What we still don't know about polyploidy. Taxon 2010, 59, 1387–1403.

[100] Soltis, D.E., Visger, C.J., Marchant, D.B., and Soltis, P.S. (2016). Polyploidy: Pitfalls and paths to a paradigm. American Journal of Botany 103, 1146-1166.

[101] Soltis, P.S., and Soltis, D.E. (2000). The role of genetic and genomic attributes in the success of polyploids. Proceedings of the National Academy of Sciences 97, 7051-7057.

[102] Soltis, D. E. and Soltis, P. S. (1999) 'Polyploidy : recurrent formation and genome evolution', 14(9), pp. 348–352.

[103] Soltis, P.S., and Soltis, D.E. (2009). The role of hybridization in plant speciation. Annual review of plant biology 60, 561-588.

[104] Spindel, J., Wright, M., Chen, C., Cobb, J., Gage, J., Harrington, S., et al. (2013). Bridging the genotyping gap: using genotyping by sequencing (GBS) to add high-density SNP markers and new value to traditional bi-parental mapping and breeding. Theoretical and Applied Genetics 126, 2699-2716.

[105] Spring, J. (1997) Vertebrate evolution by interspecific hybridizationare polyploid FEBS Lett. 400, 2–8

[106] Stebbins, G.L. (1947). Types of polyploids: their classification and significance. Advances in Genetics 1, 403- 429.

[107] Strasburger, E. (1910). Sexuelle und apogame Fortpflanzung bei Urticaceen. Jahrb Wiss Bot 47, 491-502.

[108] Sturtevant, A.H. (1913). The linear arrangement of six sex-linked factors in Drosophila, as shown by their mode of association. Journal of Experimental Zoology Part A: Ecological Genetics and Physiology 14, 43-59.

[109] Sverrisdóttir E., Byrne S., Sundmark E. H. R., Johnsen H. Ø., Kirk H. G., Asp T. et al. (2017). Genomic prediction of starch content and chipping quality in tetraploid potato using genotyping-by-sequencing. *Theor. Appl. Genet.* 130 2091–2108. 10.1007/s00122-017-2944-y

[110] Templeton, S.R. (1980) The theory of speciation via the founder principle, Genetics 94, 1011–1038

[111] Van De Peer, Y., Mizrachi, E., and Marchal, K. (2017). The evolutionary significance of polyploidy. Nature Reviews Genetics 18, 411-424.

[112] Van Dijk, T., Noordijk, Y., Dubos, T., Bink, M.C., Meulenbroek, B.J., Visser, R.G., et al. (2012). Microsatellite allele dose and configuration establishment (MADCE): an integrated approach for genetic studies in allopolyploids. BMC Plant Biology 12:25.

[113] Van Eeuwijk F. A., Bink M. C., Chenu K., Chapman S. C. (2010). Detection and use of QTL for complex traits in multiple environments. *Curr. Opin. Plant Biol.* 13 193–205. 10.1016/j.pbi.2010. 01.001

[114] Van Geest, G., Bourke, P.M., Voorrips, R.E., Marasek-Ciolakowska, A., Liao, Y., Post, A., et al. (2017a). An ultra-dense integrated linkage map for hexaploid chrysanthemum enables multi-allelic QTL analysis. Theoretical and Applied Genetics 130, 2527-2541.

[115] Van Ooijen, J.W. (2006). JoinMap® 4, Software for the calculation of genetic linkage maps in experimental populations. Kyazma B.V., Wageningen, The Netherlands.

[116] Van Ooijen, J.W., and Jansen, J. (2013). Genetic mapping in experimental populations. Cambridge University Press.

[117] VSN International (2018). *ASreml.* Hempstead: VSN International Ltd.

[118] Yu, J., Pressoir, G., Briggs, W.H., Vroh Bi, I., Yamasaki, M., Doebley, J.F., et al. (2006). A unified mixed- model method for association mapping that accounts for multiple levels. 38, 203-208.

[119] Zhang, Z., Ersoz, E., Lai, C.-Q., Todhunter, R.J., Tiwari, H.K., Gore, M.A., et al. (2010b). Mixed linear model approach adapted for genome-wide association studies. Nature Genetics 42, 355-360.

[120] Zheng, C., Voorrips, R.E., Jansen, J., Hackett, C.A., Ho, J., and Bink, M.C. (2016). Probabilistic Multilocus Haplotype Reconstruction in Outcrossing Tetraploids. Genetics 203, 119-131

[121] Zhou, C., Olukolu, B., Gemenet, D., Wu, S., Gruneberg, W., Cao, M.D., et al. (2017). Assembly of Whole- Chromosome Pseudomolecules For Polyploid Plant Genomes Using Outcrossed Mapping Populations. bioRxiv,

Citation: *Werkissa Yali, "Genomic Mapping in Polyploid Plants", International Journal of Research in Agriculture and Forestry, 7(5), 2020, pp. 10-22.*

Copyright: *© 2020 Werkissa Yali. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.*