Patterns of linkage disequilibrium in multiple populations

Dreisigacker S 1, Arief V 2, DeLacy I 2,3, Davenport G 1, Manes Y 1, Reynolds M 1, Ravi S 1, Dieters M 2, Crossa J 1
1International Maize and Wheat Improvement Center (CIMMYT), 2School of Land Crop and Food Sciences, The University of Queensland, 3Australian Centre for Plant Functional Genomics (ACPFG)

ABSTRACT

Application of association analyses in plant breeding populations has the potential to revolutionize crop genetics. To determine the optimal strategies for implementing association analysis in wheat (Triticum aestivum L. subsp. aestivum) we analysed the structure of linkage disequilibrium (LD) across the genome in two distinct populations of CIMMYT wheat breeding germplasm. LD within a total of 645 advanced lines bred during the last 2 decades and distributed via CIMMYT Elite Spring Wheat Yield Trials and 160 synthetic hexaploid wheat (SHW) have been examined. Plant materials were genotyped with DArT markers plus additional microsatellite markers on a smaller set of lines to compare LD measures among marker technologies. Variable pattern of LD among chromosomes and the two marker technologies was scored. Different patterns of LD decay were observed for advanced wheat lines and SHW. LD dropped faster in SHW and a lower percentage of loci pairs in LD was detected demonstrating the usefulness of SHW for high-resolution association analyses.

INTRODUCTION

Association analyses based on linkage disequilibrium (LD) is an emerging field of genetic mapping that has the potential for population-genetic model discovery of alleles underlying quantitative traits. Linkage disequilibrium refers to the non-random association of alleles in different loci which typically decays with the physical distance across chromosomes. The decay of LD in a genome determines the resolution of quantitative trait loci detection in association analyses studies and indicates the required marker density. Genetic and demographic factors such as mating system, population admixture or subdivision, drift, directional selection, and population history affect LD. This has lead to LD at variable levels depending upon the species, populations, or genomic region under consideration A single study determining LD for a species can therefore not be projected to all populations of the species. Inference of LD levels across the genome of a population can also be misleading because LD patterns are variable among chromosomes and over distance. This is due to effects of selection, heterozygosity, and effective recombination rates varying between and within chromosomes. E.g directional selection for a trait controlled by multiple loci on a single chromosome will fix these loci in the population, resulting in little variation.

Effective detectable recombination rate and LD decay decreases as homozygosity increases. Because higher levels of homozygosity occur in autogamous species than in allogamous species it is expected to see increased LD in self-pollinated plants. Recent studies have confirmed this hypothesis. For barley, soybean, and rice significant genome-wide LD degrade at 90-212 kb, 90kb to >500kb2 and 75kb to > 500kb3, respectively. In wheat, LD measured as a physical distance using microsatellite marker polymorphisms have shown to decay across the genome within 10-50 cM4,5,6.

Although the levels and patterns of LD in wheat have begun to be characterized, there is still little information on LD in this crop species across diverse populations. In addition, a direct comparison of the extent of LD between different types of genome-wide distributed molecular markers in wheat is lacking. In this study we describe the extent and genomic distribution of LD in two populations of hexaploid wheat undergoing diverse recombination history and compare the decay of LD measured using DArT and microsatellites in a subset of population.

MATERIALS AND METHODS

Plant materials included 645 advanced wheat lines and 160 synthetic hexaploid wheat (SHW). The wheat lines were bred and distributed by CIMMYT during the last 2 decades within the Elite Spring Wheat Yield Trials (ESWYT). The synthetic hexaploid wheats were developed by interspecific hybridization of diploid Ae. tauschii and tetraploid durum wheat accessions by CIMMYT during the last decade.

Plant materials were genotyped with DArT markers generated by Triticarte Pty. Ltd. (Canberra, Australia; http://www.triticarte.com.au). Four hundred thirty four and 293 pairwise DArT estimates were scored for the advanced wheat lines and SHW, respectively with rare alleles (<5%) omitted in the LD analyses. Linkage disequilibrium was calculated by means of the pairwise estimates r2 and D′ using the software PowerMarker (www.powermarker.net). The LD was examined for all adjacent locus pairs (linked) as well as for all pairwise comparisons of loci across the genome (unlinked). To assess the influence of interspecific hybridization and polyploidization events on pairwise estimates, LD was additionally determined for the AB and D genome separately. Linkage disequilibrium was calculated for the entire population of plant materials and the largest subgroup of each population determined via pattern analyses (Arief et al., this proceedings). Map distances were determined using a DArT consensus map. Eighty nine lines within the ESWYT were additionally genotyped with 49 microsatellites for LD analyses.
RESULTS AND DISCUSSION

Different patterns of LD decay were observed for advanced lines and SHW. The distance over which LD decayed below a population-specific threshold as the 95th percentile of the distribution of \( r^2 \) was within approx. 40 cM for the advanced wheat lines and within approx. 30 cM for the SHW indicating that LD degrade faster in SHW (data not shown). The percentage of loci pairs found genome-wide and among chromosomes in significant LD (\( r^2 \) and \( D' \), \( p<0.001 \)) is reported in Table 1. Genome-wide LD analysis revealed that overall, equivalent or fewer than 5% of DArT loci pairs showed significant LD across all populations. Among chromosomes, advanced lines showed a higher percentage of loci pairs in significant LD for both methods of pairwise comparisons than the SHW. Evaluating LD statistics within the largest subgroup of each population, the percentage of locus pairs in significant LD within each subpopulation was additionally reduced and was superior in advanced lines than in SHW confirming a faster degrade of LD in SHW. Figure 1 provides a comparison of LD across the length of the chromosome of homologous group 1. Variable pattern of LD among the three chromosomes was observed. The limited LD on SHW across the chromosome regions is also apparent when \( r^2 \) is plotted against distance.

Wheat is a model species of allopolyploids. A number of reports have shown that interspecific hybridizations lead to substantial genomic changes, which most likely facilitated the successful establishment of allopolyploids in nature\(^5\). Loss and gain of many DNA fragments in the generations immediately following polyploidy formation have been found and several potential mechanisms that could generate these changes have been listed including homologous recombination, point mutations, transposon activation and gene conversion-like events. These genomic and epigenomic changes are rapid events, which conceivably affect LD and might be responsible for the faster decay in SHW.

Reduced LD in subpopulations could be accounted for by the effect of population structure which has been reported in previous studies in wheat\(^4,10\).

To access the influence of interspecific hybridization and polyploidy in the different genomes, LD statistics were calculated for the AB and D genome separately.

| Table 1: Percentage, ratio and mean of DArT loci pairs in significant (\( P<0.001 \)) linkage disequilibrium within a diverse population of advanced CIMMYT wheat lines grown in Elite Spring Wheat Yield Trials (ESWYT) and a synthetic hexaploid wheat (SHW). |
|---|---|---|---|---|---|---|---|
| Group | No. of entries | Unlinked | Linked | Ratio\(^1\) | Unlinked | Linked | Unlinked | Linked | Ratio\(^1\) | Unlinked | Linked |
| ESWYT | 645 | 5 | 28 | 0.41 | 0.224 | 0.448 | 4 | 27 | 0.48 | 0.731 | 0.855 |
| ESWYT subgroup | 253 | 5 | 22 | 0.33 | 0.108 | 0.388 | 1 | 11 | 1.42 | 0.950 | 0.958 |
| ESWYT (AB genome) | 645 | 5 | 28 | 0.48 | 0.210 | 0.438 | 4 | 27 | 0.54 | 0.739 | 0.856 |
| ESWYT (D genome) | 645 | 5 | 33 | 1.73 | 0.398 | 0.802 | 5 | 31 | 1.72 | 0.678 | 0.938 |
| SHW | 162 | 5 | 22 | 0.34 | 0.193 | 0.489 | 3 | 19 | 0.60 | 0.795 | 0.906 |
| SHW subgroup | 97 | 4 | 19 | 0.40 | 0.127 | 0.523 | 1 | 11 | 1.25 | 0.415 | 0.528 |
| SHW (AB genome) | 162 | 5 | 22 | 0.42 | 0.201 | 0.221 | 3 | 18 | 0.68 | 0.795 | 0.914 |
| SHW (D genome) | 162 | 5 | 13 | 0.67 | 0.344 | 0.735 | 3 | 22 | 1.70 | 0.797 | 0.882 |

\(^1\)Ratio linked vs. unlinked loci pairs in LD
REFERENCES


