Research Statement

Martin Mascher

Research focus

My research group at IPK Gatersleben studies domestication and adaptation processes and their interaction with genetic diversity in crops and their wild relatives. Research is mainly focused on the temperate cereals barley, wheat and rye, for which a large number of wild and domesticated accessions are maintained in the IPK genebank. The main research goals are:

1. Elucidating the relationship between crops and their extant wild relatives and progenitors;
2. Tracing the demographic development and adaptation of cereal crops as they expanded their range from the initial site of domestication in the Fertile Crescent into Europe;
3. Understanding the molecular consequences of domestication on patterns of nucleotide diversity, gene expression and gene regulation and recovering lost variation for plant breeding.

Towards these aims, we apply methods of population genetics, genome informatics and gene expression analysis to genome-wide sequence, genetic marker and transcriptome datasets. We are also involved in the development of genomic sequence and map resources to enable the unbiased assessment of genetic variation in highly diverse germplasm collections and facilitate the genomic contextualization of sequence data.

Past achievements

1. **Reference sequence assembly for wheat and barley**

   The International Barley Genome Sequencing Consortium has recently reported a reference genome sequence assembly of the barley cv. Morex. We have established a general workflow for constructing chromosome-scale assemblies of Triticeae genomes using three-dimensional (3D) contact probabilities obtained by means of chromosome conformation capture sequencing (Hi-C). This allowed us to derive, for the first time, a linear order of the vast non-recombining regions of the barley genome. The barley genome sequence will arguably represent the most important resource for barley genetics and genomics in the coming years. We applied the computational methods developed for assembling the barley genome to wheat. Chromosome-scale assemblies were constructed for a wild emmer wheat accession, an elite durum variety, and the reference accession of bread wheat (Chinese Spring).

2. **Genomic analysis of 6000-year old barley**

   Together with Israeli archeologists and German archaeogeneticists, we analyzed DNA sequences retrieved from barley grains recently excavated in a Chalcolithic cave site in the Judean desert. The ancient samples were put into the context of extant barley diversity by joint analysis with exome sequences from 267 wild and domesticated barleys, mainly from IPK's *ex situ* collection. Our key finding was that the ancient barley samples are closely related to present-day landraces from the Southern Levant, presumably because key domestication and adaptation processes had been finished by 4000 BC (6000 years after the inception of barley domestication).

Ongoing and future work

1. **Wheat and barley population genomics**
   We are currently leading bioinformatics efforts in several highly synergistic projects to expand the understanding of genetic diversity in cereal crops at various levels.
   In collaboration with We have genotyped 20,000 wild and domesticated barley accessions. This dataset has (i) highlighted a history of recent intentional admixture between major germplasm groups due to breeders' efforts, (ii) enabled the mapping of highly morphological traits and virus resistance loci by means of GWAS, and (iii) led to the assembly of genetically defined core collections for future deep analysis.
   As a next step, we are currently (i) constructing chromosome-scale reference assemblies of three barley genotypes (one cultivar, one landrace, one wild barley); (ii) to collect low-pass whole-genome shotgun (WGS) data for >200 wild and domesticated accessions; and (iii) to analyze structural variations based on deep WGS data, Hi-C and complementary resources (genetic maps, optical maps). The high-resolution whole-genome SNP datasets for accessions from different stages of domestication and crop improvement will enable the utilization of advanced algorithms for selection scans, demographic inference and genome-wide association studies developed in human genetics.
2. **Rye and oat genomics**
   Rye and oats can be considered the 'orphans crops' of cereal genomics. Compared to their relatives wheat and barley, their genomic resources are underdeveloped and no representative high-resolution surveys of global genetic diversity in these crops have been assembled so far. I have recently joined an international consortium on oat genome sequencing where I contribute expertise in genome assembly and analysis of resequencing data of diversity collections. A PhD student in my group is currently studying the population structure of rye and its relationship to wild relatives based on GBS data from a core collection of IPK's collection. We also found evidence for crop-wild hybridization and incomplete lineage sorting across the genus *Secale*, either due to recent, and possibly incomplete speciation, or frequent inter-species geneflow.
3. **Regulatory architecture in wild and domesticated barley**
   We study the relative contributions of cis- and trans-effects on gene regulation in wild and domesticated barleys and their interaction with an environmental factor (low temperature). Eight wild barleys and landraces have been crossed to the common parent Morex. RNA-seq is then used to assay allele-specific gene expression in parents and the F_1 hybrids to dissect the relative contributions of cis- and transacting factors to gene regulation and their response to cold treatment. RNA-seq data and genomic exome capture sequences of the parents and hybrids have been collected and are currently being analyzed.
4. **Chromatin organization across the plant cell cycle**
   While we have established Hi-C mainly for chromosome-scale assembly of Triticeae genomes, this method was originally developed to interrogate the 3D organization of chromatin. In collaboration with Jaroslav Doležel (IEB Olomouc, Czech Republic), we have analyzed Hi-C data from flow-sorted mitotic chromosomes. We found the relationship between contact probability and linear distance between genomic loci in mitotic plant chromosomes to be fundamentally different from that of mammalian mitotic chromosomes. To expand on these preliminary results, we plan to assay chromatin organization with the complementary methods (Hi-C and fluorescence *in situ* hybridization) throughout the barley cell cycle.