# Whole-Genome Resequencing of a Worldwide Collection of Rapeseed Accessions Reveals the Genetic Basis of Ecotype Divergence

**Dezhi Wu[1,8], Zhe Liang[2,8], Tao Yan[1], Ying Xu[1], Lijie Xuan[1], Juan Tang[3], Gang Zhou[3], Ulrike Lohwasser[4], Shuijin Hua[5], Haoyi Wang[1], Xiaoyang Chen[7], Qian Wang[1], Le Zhu[1], Antony Maodzeka[1], Nazim Hussain[1], Zhilan Li[1], Xuming Li[3], Imran H Shamsi[1], Ghulam Jilani[6], Linde Wu[3], Hongkun Zheng[3], Guoping Zhang[1], Boulos Chalhoub[1], Lisha Shen[2,*], Hao Yu[2,*], and Lixi Jiang[1,*]**

[1]Institute of Crop Science, Zhejiang University, China. [2]Temasek Life Sciences Laboratory and Department of Biological Science, National University of Singapore, Singapore. [3]Biomarker Technologies Corporation, China. [4]Department of Genebank, Leibniz Institute of Plant Genetics and Crop Plant Research, Germany. [5]Institute of Crop and Nuclear Agricultural Sciences, Zhejiang Academy of Agricultural Sciences, China. [6]Office of Research, Innovation & Commercialization, PMAS-Arid Agricultural University Rawalpindi, Rawalpindi-Pakistan. [7]Institute of Crop Science, Jinhua Academy of Agricultural Sciences, China. [8]These authors contributed equally to this work.

The work was supported by the National Key Basic Research Project (No. 2015CB150205), Natural Science Foundation of China (No. 31671597, 31370313, 31670283), Sino-German Science Centre for Research Promotion (GZ 1099), Jiangsu Collaborative Innovation Center for Modern Crop Production, and the Singapore National Research Foundation Investigatorship Programme (NRF-NRFI2016-02).

## Introduction

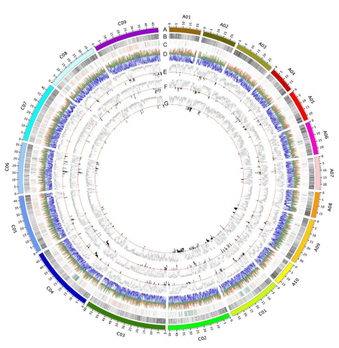


Fig 6. Circos Plot Showing Genetic Diversity and Selective Signals among Three Ecotype Groups.

**Background:**

Rapeseed (*Brassica napus* L.), an important oilseed crop, has adapted to diverse climate zones and latitudes by forming three main ecotype groups, namely winter, semi-winter, and spring types. Although genetic variations among winter, semi-winter, and spring ecotypes of rapeseed have been reported, the key allelic variations underlying the divergence of the different *B. napus* ecotypes are not fully understood.

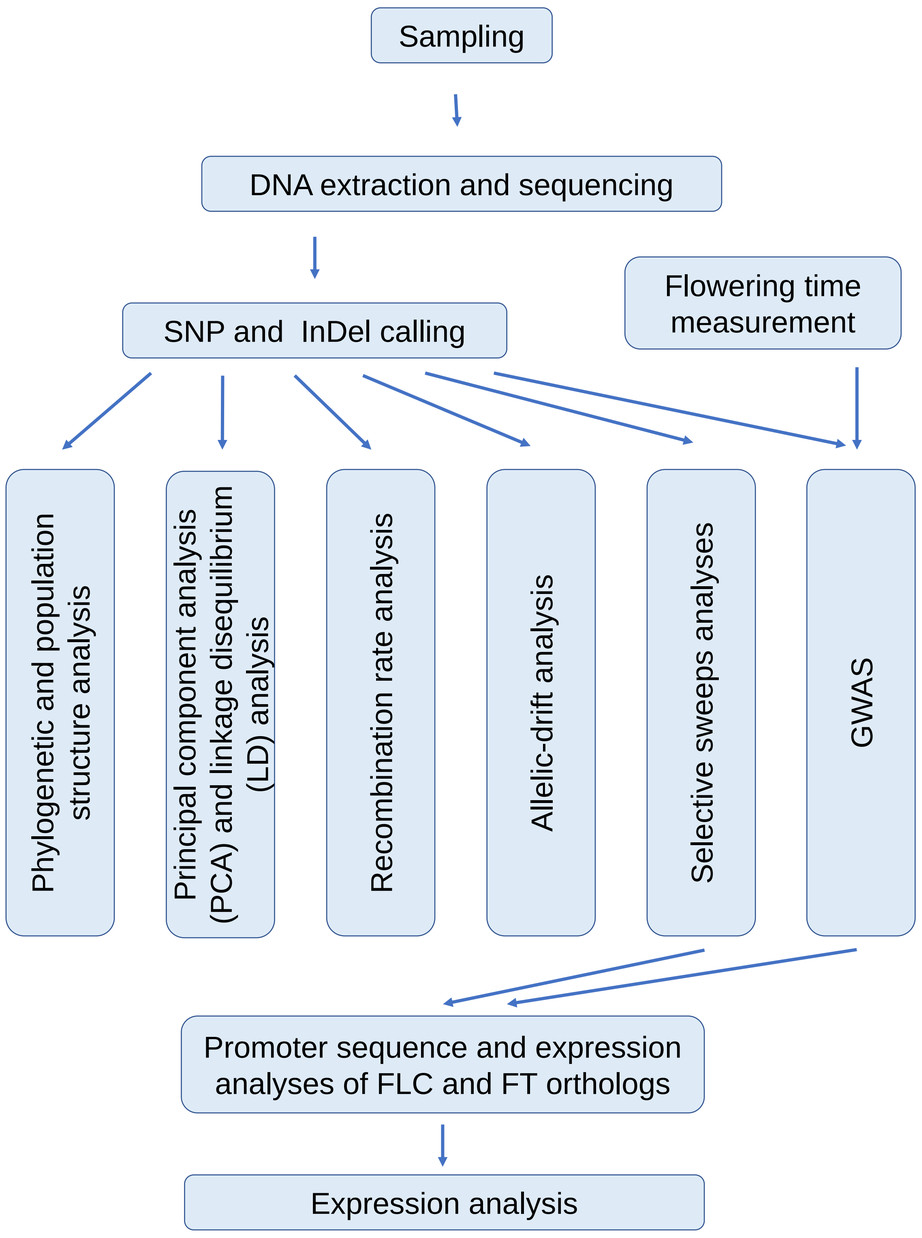
**Objectives:**

(1) To identify the global pattern of genetic polymorphism in rapeseed revealed by resequencing a worldwide collection of 991 germplasm accessions.

(2) To shed new light on the genomic footprints generated during the process of natural and artificial selection and flowering-time divergence among three major ecotype groups.

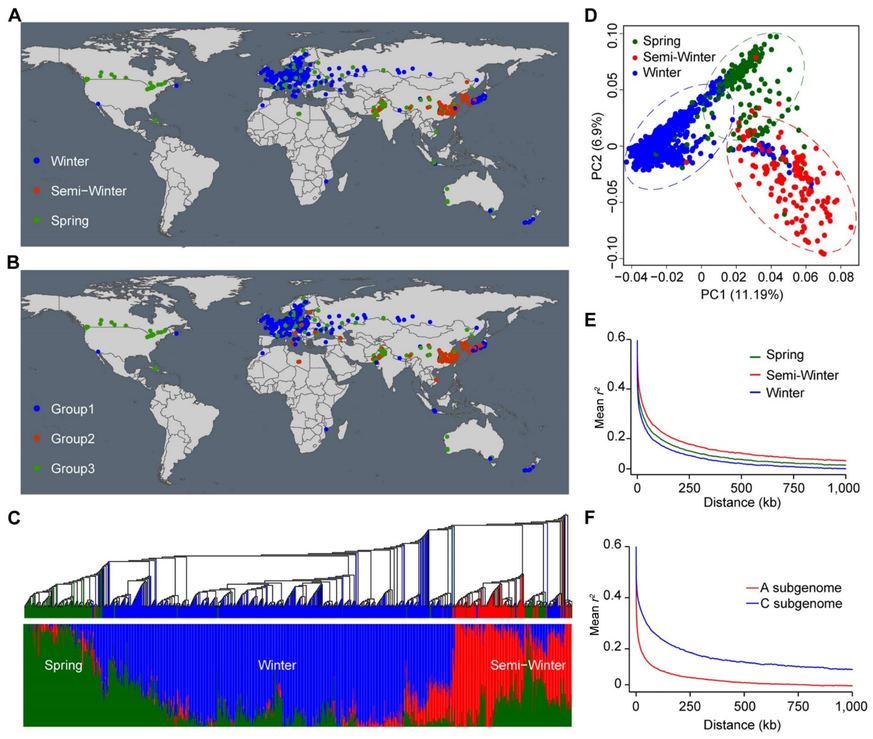
(3) To provide by far the largest genetic resource for screening molecular markers for the genetic improvement of rapeseed.

## Methodology




## Results & Discussion

### 1. Resequencing, Population Structure and Genomic Variation among the 991 Accessions




We generated a total of 7.82 Tb of clean reads from whole-genome resequencing of 991 *B. napus* accessions originating from 39 countries across the world with an average of 6.6-fold-coverage. After SNP and InDel calling, 5.56 million SNPs and 1.86 million InDels were detected.

Figure 1. Distribution, Population Structure, PCA, and LD Decay of the 991 Rapeseed Germplasm Accessions.
The geographic distribution of the 991 rapeseed accessions based on country of origin and ecotype information(A) and population structure and PCA (B). (C) Population structure analysis using 293 498 SNPs (missing data < 50%, MAF > 5%, r2 < 0.2). (D) PCA plot distinguishes three types. (E) Genome-wide average LD decay in the winter, semi-winter, and spring ecotypes. (F) Genome-wide average LD decay for the two subgenomes of *B. napus*.

### 2. Paths of Allelic Drift among the Major Sites of Origin

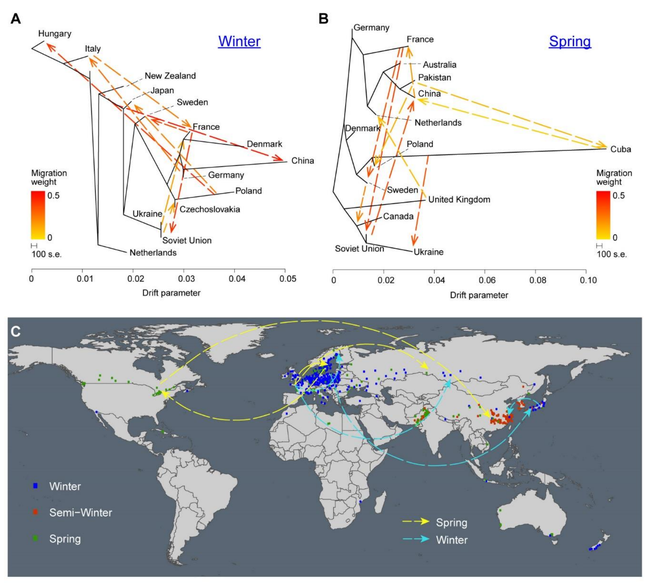



Figure 2. Allelic-Drift Paths Showing Population Splits and Mixtures among the Major Sites of Origin of the Accessions.
The direction of gene flow for the winter-type rapeseed accessions was from France to Russia, France to Sweden via Germany, Poland to Japan, and Japan to China, while spring-type rapeseed spread from France to Russia, France to Sweden via Germany, Sweden to Canada, and Canada to China.

### 3. Selective Signals and Genes Associated with Flowering Time.

The strongest signal of a selective sweep was found on Chr.A10 between the winter and spring types, including an FLC ortholog (BnaA10g22080D). Another strong selective signal was found in a large region on Chr.A02, including orthologs of FT (BnaA02g12130D).

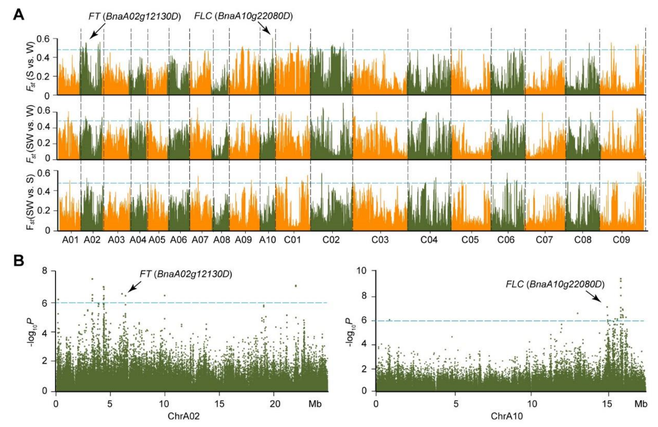



Figure 3. Selective-Sweep Signals between the Three *B. napus* Ecotype Groups.
(A) Genome-wide distribution of selective-sweep signals identified through comparisons between three ecotypes.
(B) Manhattan plots of GWAS results for the flowering-time trait on Chr.A02 and Chr.A10. The identified SNPs were in located the regions from 4 to 7 Mb on Chr.A02 (left) and the regions from 14 to 17 Mb on Chr.A10 (right).

### 4. Identification of SNPs Corresponding to Ecotype Divergence

The ecotype-specific SNPs are associated with the expression levels of FT and FLC orthologs, and could contribute to the differences in flowering time observed among the three ecotype groups.

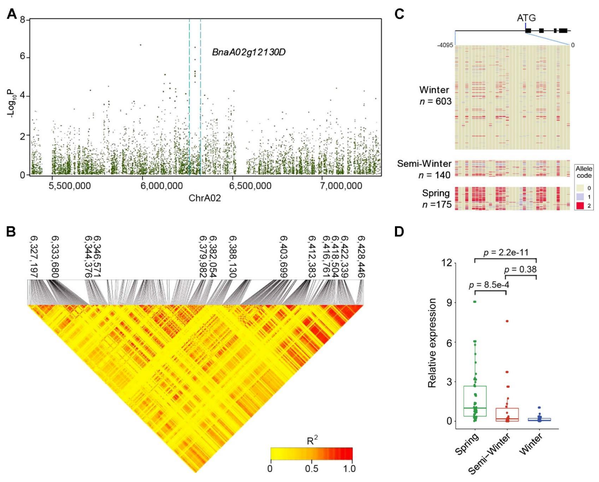



Figure 4. The SNPs Responsible for Flowering-Time Variation in the Region from 6.37 to 6.38 Mb on Chromosome A02.

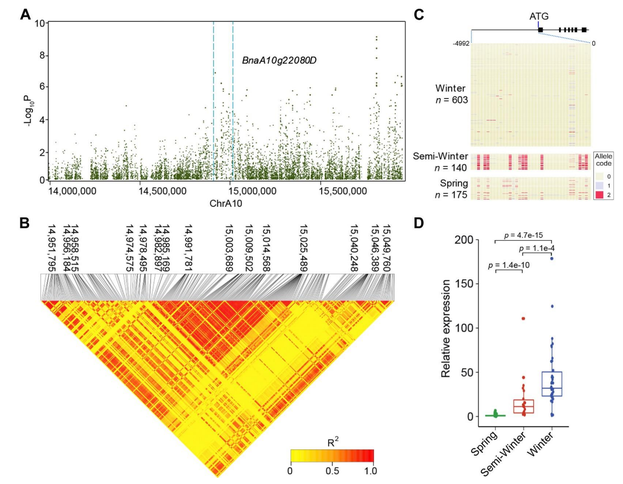



Figure 5. The SNPs Responsible for Flowering-Time Variation in the Region from 14.99 to 15.00 Mb on Chromosome A10.

## Conclusion

- Resequenced a worldwide collection of 991 *B. napus* germplasm accessions, including 658 winter types, 145 semi-winter types, and 188 spring types, from 39 countries.
- Identified the global pattern of genetic polymorphism in *B. napus* and revealed the paths of allelic drift showing the splits and mixtures of populations among the major origins.
- Studied the selective sweeps produced during natural and artificial selection, and uncovered the genetic basis underlying the divergence of the main ecotypes. GWAS of the flowering-time trait identified SNPs in the promoter regions of FT and FLC orthologs, which specifically correspond to the three rapeseed ecotype groups.
- Provided important insights into the genomic footprints of rapeseed evolution and flowering-time divergence among the three ecotype groups, and will facilitate screening of molecular markers for accelerating rapeseed breeding.