

High throughput SNP-based analysis of population structure and linkage disequilibrium in landrace barley

Moragues, M.¹, Flavell, A. J.¹, Booth, A.², Ceccarelli, S.³, Grando, S.³, Russell, J.²

¹Division of Plant Sciences, University of Dundee, UK, ²Scottish Crop Research Institute, Invergowrie, Dundee, UK

³International Center for Agricultural Research in the Dry Areas, P.O. Box 5466, Aleppo, Syria

The domestication of barley took place about 10,000 years ago in the “Fertile Crescent”. Since then, the repeated human and natural selection led to locally adapted landraces. Therefore, landraces form an important genepool for breeding programs. Additionally, the genetic structure of landrace collection, when linked to geographical and environmental data, may reveal genomic signatures of selection.

These samples were subjected to high throughput genotyping using the Illumina GoldenGate BeadArray platform. From the 1536 SNPs on BOPA1, allele calls were not possible for 59 SNPs, 198 are monomorphic in the sample set and 608 show very low allele frequency (below 0.10). Thus, 670 SNPs (43.6 % of the full BOPA1 marker set) were useful for linkage disequilibrium (LD) and population structure analysis.

Fig. 2. SNP based Q plot from STRUCTURE (Falush et al., 2003) for K = 3

Numbers on x-axis are the collection sites on the map. Numbers on the y-axis are the P of each individual to belong to each of the 3 subpopulations.

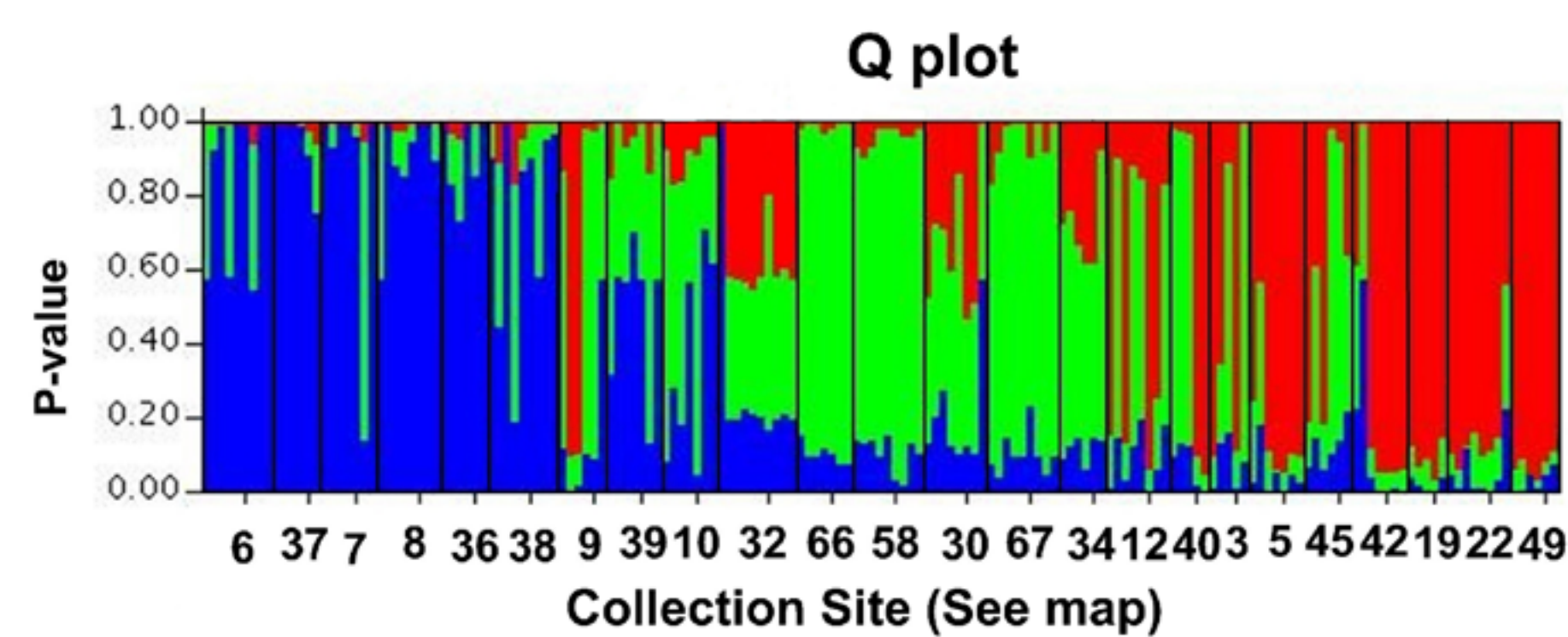


Fig. 1

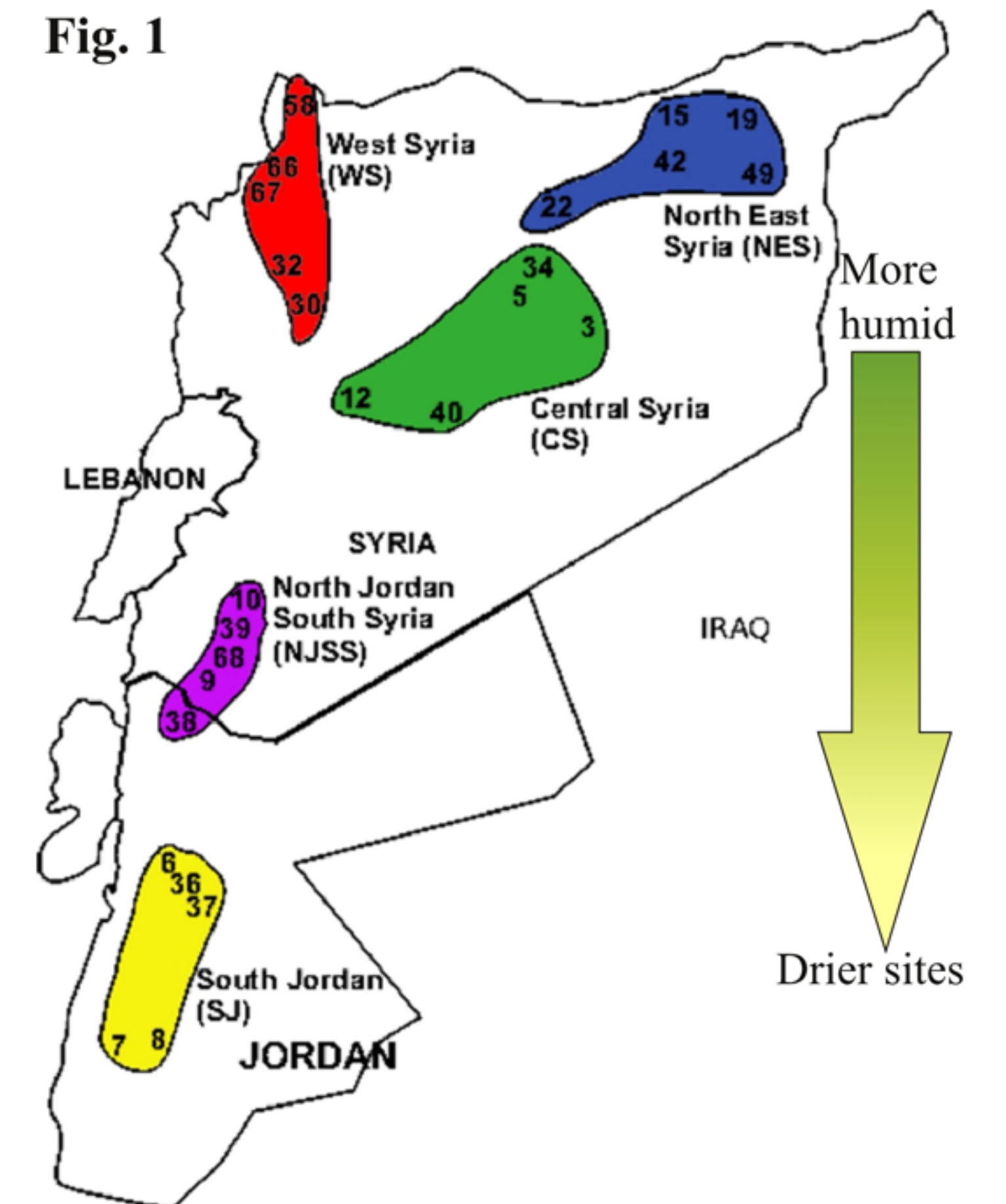
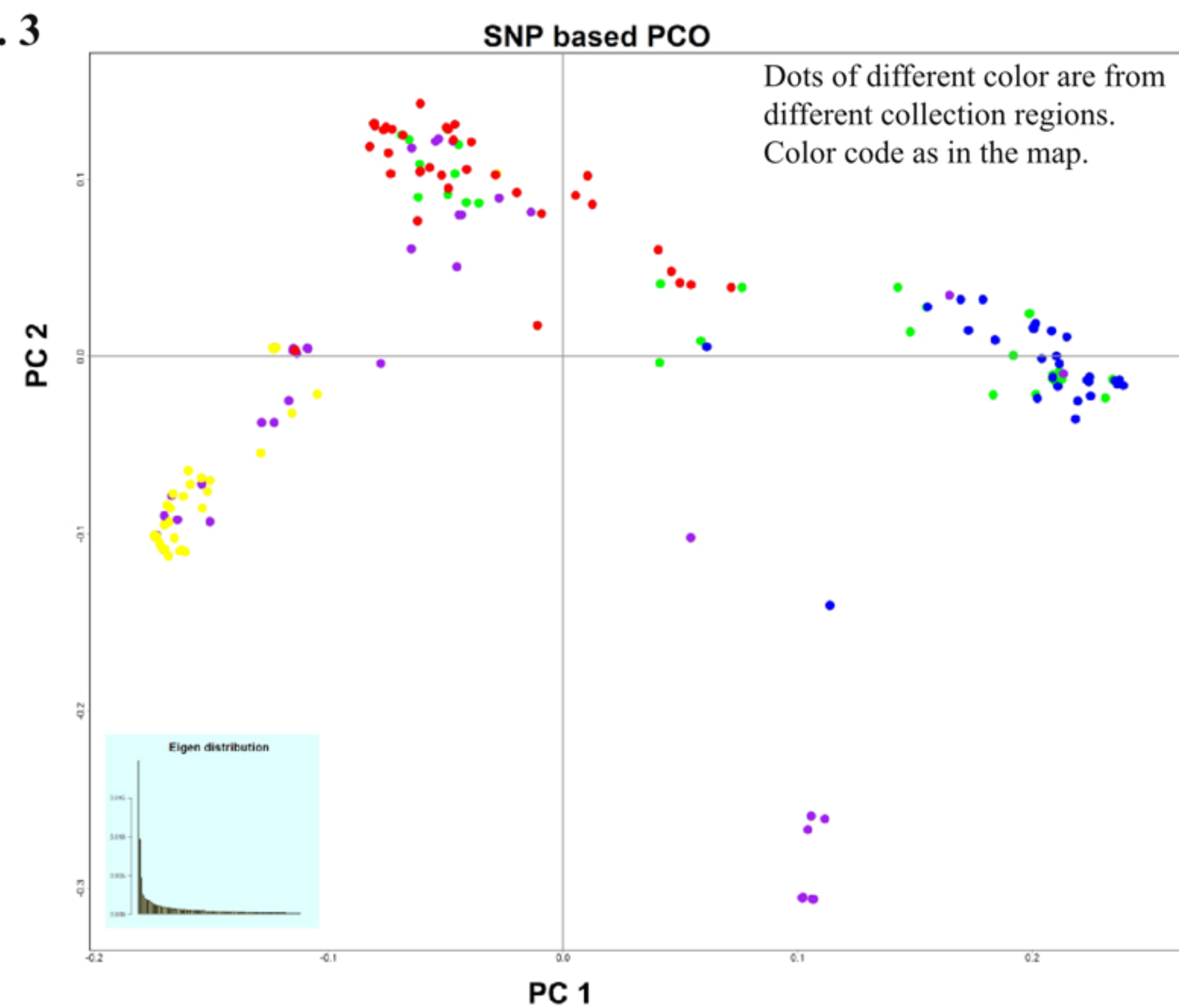


Fig. 3



The population structure was estimated by non-metric multidimensional scaling (Fig. 3) and STRUCTURE (Pritchard et al., 2001) (Fig. 2). The revealed population structure follows a geographical pattern. The lines from North East Syria and South Jordan differ markedly from each other but are very similar within collection regions. The landraces from the three other regions show diversity between the former two regions and display much more internal variability. Patterns of LD for these samples along the barley chromosomes show peaks of LD around the centromeres (Fig. 6), which corresponds with a decrease in diversity in those regions. In this property, this landrace collection performs in a similar way to modern cultivars and confirms the lack of recombination present in centromeric regions in barley. Additionally, the LD decays very rapidly as the genetic distance between markers increases (Fig. 5), which makes it suitable for fine mapping of QTLs.

Differences in the patterns of diversity between regions of origin were observed along the chromosome (Fig. 4). This may highlight selection signatures of adaptation to the environment and/or agronomic practices prevalent in those regions.

Fig. 4

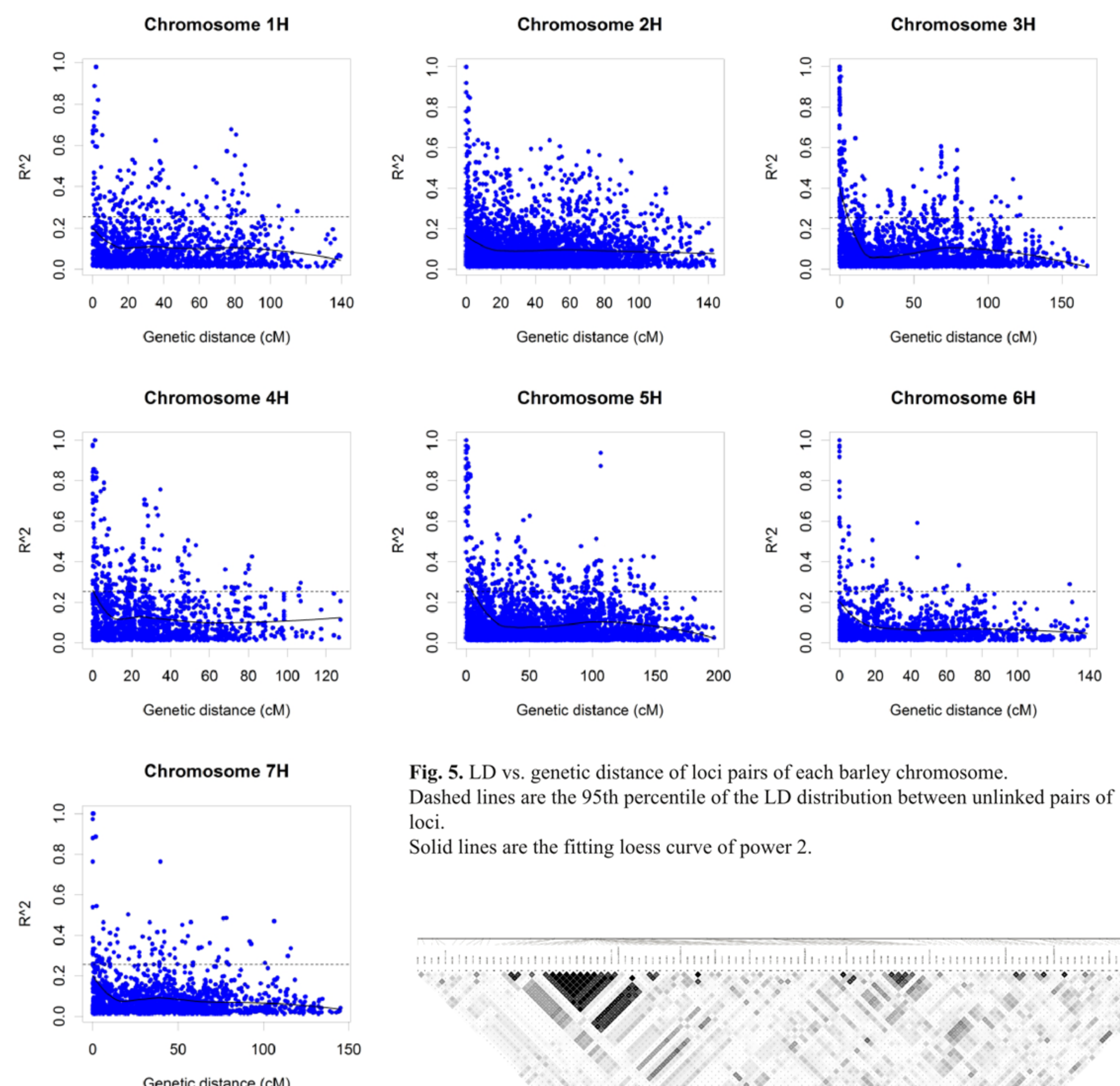
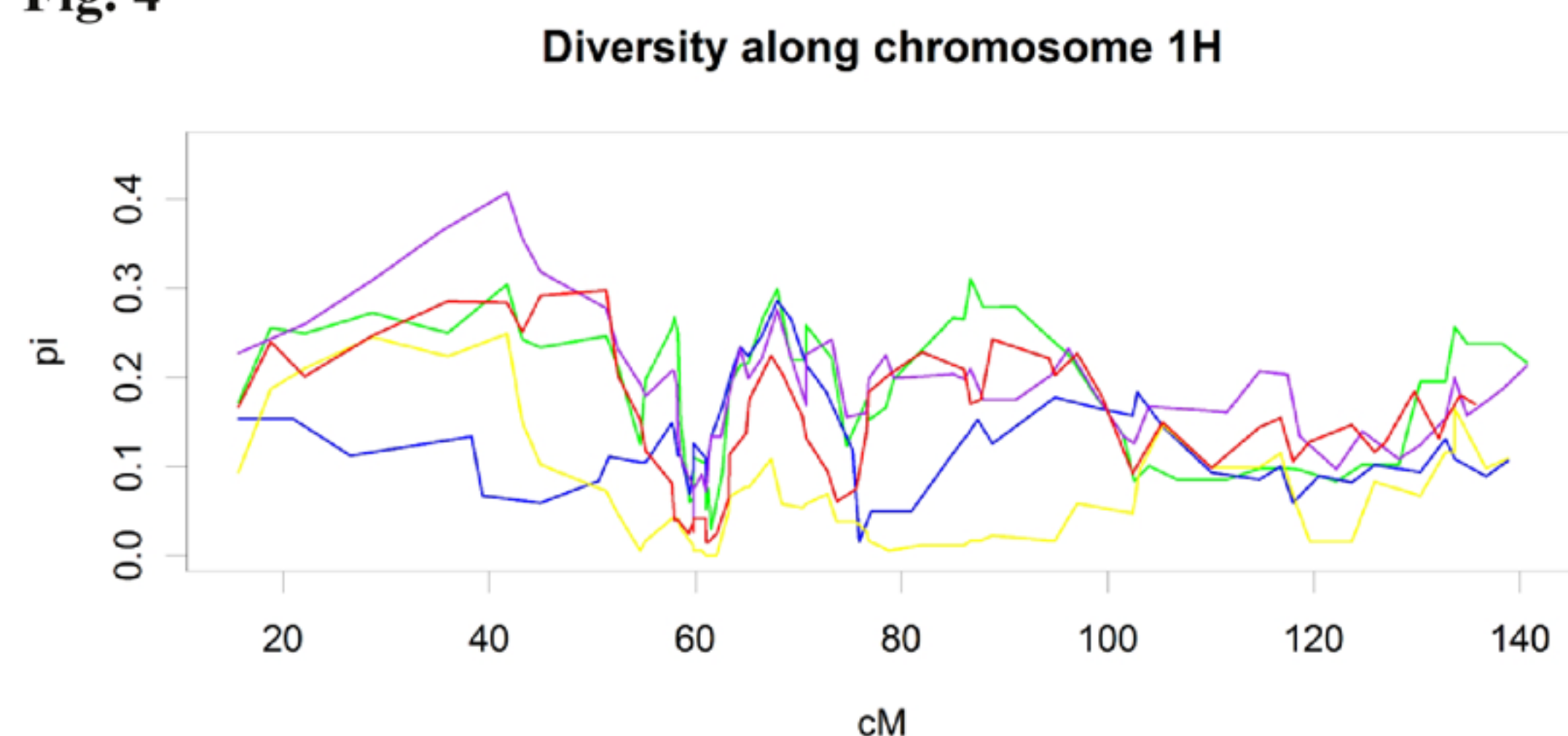


Fig. 5. LD vs. genetic distance of loci pairs of each barley chromosome. Dashed lines are the 95th percentile of the LD distribution between unlinked pairs of loci. Solid lines are the fitting loess curve of power 2.

CONCLUSIONS

This set of landraces shows a geographical population structure that follows a gradient in drought, with high LD in centromeric regions and a rapid LD decay along the chromosomes, which makes it a useful population to fine map QTLs.

REFERENCES

FALUSH et al. (2003). *Genetics*, 164, 1567-1587.
RUSSELL et al. (2003). *Theor Appl Genet*, 107:413-421.

ACKNOWLEDGEMENTS

We thank Robbie Waugh, David Marshall and Jordi Comadran for interesting discussion on the results. This work was supported by the ERA-PG grant 06.061A-EXBARDIV and the Generation Challenge Program. SCRI is funded by the Scottish Govern.

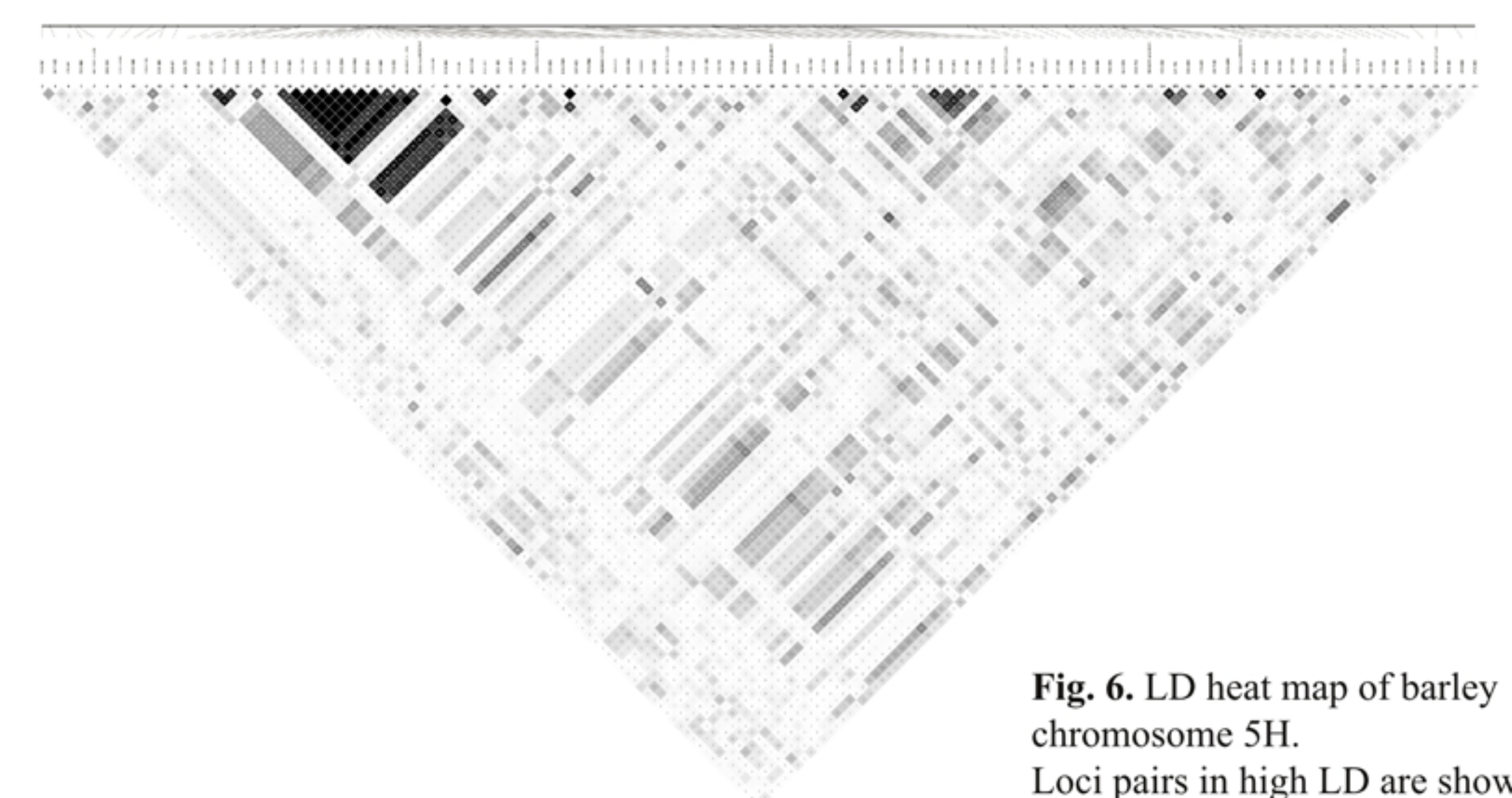


Fig. 6. LD heat map of barley chromosome 5H. Loci pairs in high LD are shown as dark areas.