

# Using Molecular Allelic Variation to Understand Domestication Processes and Conserve Diversity in *Brassica* Crops

G.J. King  
Comparative Genetics and Genomics Group,  
Horticulture Research International,  
Wellesbourne, Warwick CV35 9EF, UK

**Keywords:** *Brassica*, domestication, biodiversity, allelic variation, QTL, quantitative genetics

## Abstract

The harvested curd of cauliflower (*Brassica oleracea* var. *botrytis* L.) consists of proliferating, arrested inflorescence and floral meristems. The origins and events leading to the domestication of this important crop trait remain unclear. A similar phenotype observed in the *ap1-1/cal-1* mutant of *Arabidopsis thaliana* resulted in speculation that the orthologous genes from *B. oleracea* may be responsible for this characteristic trait. A recent molecular and genetic study has led to development of a genetic model based on segregation of recessive alleles at specific, mapped loci of the candidate genes *BoCAL* and *BoAPI*. This model, although incomplete, accounts for differences in stage of arrest between cauliflower and Calabrese broccoli (*B. oleracea* var. *italica* Plenck), and predicts the intermediate stages of arrest similar to those observed in Sicilian Purple types. There appears to be a reduction in allele diversity at two key loci associated with this crop type. Surveys of ex situ genetic resource collections have demonstrated the association of alleles of *BoCAL-a* with curding phenotypes of *B. oleracea*. Strong correlations exist between specific alleles of *BoCAL-a* and discrete inflorescence morphologies. These complementary lines of evidence suggest that the cauliflower curd arose in southern Italy from a heading Calabrese broccoli via an intermediate Sicilian crop type. The close association of *BoAPI-a* and *BoAPI-c* with the self-incompatibility locus *S* may have contributed to the observed reduced number of *S*-alleles within the modern cauliflower gene-pool. The results indicate that it is important to consider the effect of strong selection for crop type on genetic erosion of loci linked in coupling to genes underlying the crop phenotype.

## INTRODUCTION

The assessment and conservation of biodiversity in the context of crop species is currently loosely defined in terms of both genetic resources and the underlying genetic diversity. For conservation and utilisation of genetic diversity it is useful to be able to distinguish between a) allelic variation at one or more loci and b) unique combinations of alleles which result in unique or distinct phenotypes, or c) the ability of populations to undergo rapid adaptation. The clear definition of these different aspects of genetic diversity is particularly important when considering strategies and techniques for collection, evaluation, and conservation. This paper will outline an example of how variation at a small number of loci requires careful interpretation in the context of crop domestication.

The characteristic curd phenotype of cauliflower consists of proliferating, arrested, inflorescence meristems. This phenotype is seen in a wide range of forms across the crop types defined as *Brassica oleracea* L. vars. *botrytis* and *italica*, including the convars Romanesco, DiJesi, Macerata, and the classic white cauliflower. Each of these types is closely associated with eco-geographic distribution in Italy, which is the centre of diversity for the crop (Massie et al., 1996). Curd appearance differs significantly across the range of morphotypes (Watts, 1965), dependant to a large extent on differing iteration intervals during curd development (Kieffer et al., 1998). Most curd forms are heading, with the notable exception of white sprouting broccoli (Gray, 1982). Colour is the second most distinguishing factor between curds of different convars. Thus, in addition to the

'classic' white curds, whose persistence is attributed to lower expression of peroxidase genes (Dickson and Lee, 1980), accumulation of chlorophyll, anthocyanin, and beta-carotene result in, respectively, the green curds of macerata and romanesco types (Crisp and Angell, 1985), Sicilian purple types, and orange curds. However, the characteristic of all curds is the arrest prior to floral induction. This arrest stage is temperature sensitive (Wurr and Fellows, 1998), with eventual bolting and the production of wild type flowers.

## THE HISTORY OF CAULIFLOWER DOMESTICATION

The known history of the domestication of cauliflowers is incomplete, and has relied primarily on ancient text and drawings to reconstruct the sequence of events (Crisp, 1982; Crisp and Tapsell, 1993; Massie, 1998). As with all historical sources however, meaning can be lost during translation and interpretation. Even picture sources are scarce and little of what could be positively identified as a cauliflower exists before the sixteenth century (Zeven, 1996). There are conflicting theories of cauliflower phylogeny, although there is widespread support for the likelihood of the classic curd phenotype of cauliflower originating within the last two thousand years (Crisp, 1982; Helm, 1963; Thompson, 1976).

The most prevalent domestication theory appears to be of an Eastern European origin where both *B. cretica* (Gates, 1953) and *B. cypria* (Giles, 1941) have been proposed as the precursor of the cauliflowers. Following introduction into Italy, possibly around 1490 (Thompson, 1976), a secondary centre of diversity arose with development of distinct local forms isolated by geography and strong trade barriers between the many feudal states (Massie, 1998).

Classical breeding investigations have attempted to characterise the genetics underlying the curding phenotype of cauliflowers, with limited success (Crisp, 1982; Pease, 1926). Crosses between cauliflower and Calabrese broccoli often produce intermediate phenotypes, suggesting curding is a multifactorial trait. However, segregation studies suggest that, although polygenic, curding is more likely to be controlled by a few key loci rather than through an accumulation of many 'less significant' mutations (Crisp, 1982). Until recently a simple genetic model accounting for the inflorescence phenotype of the *Brassica* cauliflower had not been proposed.

## CANDIDATE GENES FROM ARABIDOPSIS HAVE LED TO A GENETIC MODEL FOR CURD FORMATION

An *Arabidopsis* mutant was identified whose inflorescence structure was enhanced from an *apetala1-1* phenotype (Irish and Sussex, 1990) to a form strongly resembling the curd phenotype of the *Brassica* cauliflower (Bowman et al., 1993). The mutant gene, which in conjunction with *APETALA1* (*API*), is responsible for this phenotype was named *CAULIFLOWER* (*CAL*). These two genes are closely related members of the MADS-box gene family of putative transcription factors, and are associated with both floral organ identity and the switch from inflorescence to floral meristem fate. A *B. oleracea* orthologue of the *Arabidopsis* *CAL* gene, *BoCAL* (syn. *BoCAL-a*), was cloned Kempin et al. (1995), and an allele of this gene from the *Brassica* cauliflower, termed *BobCAL* (syn. *Bocal-a*), was found to contain a premature stop codon.

The similarity between the *ap1-1/cal-1* double mutant *Arabidopsis* and the *Brassica* cauliflower curd, along with the discovery of the 'mutated' copy of *BoCAL* in *B. oleracea* var. *botrytis*, suggested that the curding phenotype of cauliflower may be at least in part due to a mutation at a locus containing *BoCAL*.

A simple genetic model for domestication of cauliflower curds has subsequently been proposed (Smith and King, 2000). The model is based on segregation of recessive alleles at specific, mapped loci of the candidate genes *BoCAL* and *BoAPI*. This accounts for differences in stage of arrest between cauliflower and Calabrese broccoli (*B. oleracea* var. *italica* Plenck), and predicts the intermediate stages of arrest similar to those observed in Sicilian Purple types. Although the model is incomplete in accounting for all stages of developmental arrest subject to environmental variation, it does provides a useful basis in

accounting for major loci at which key alleles have been selected and maintained during recent breeding history.

### **DISTRIBUTION OF *BOCAL-A* ALLELES IN THE *B. OLERACEA* GENE-POOL**

The geographical distribution of *Brassica* crop morphotypes has been examined on several occasions (Branca and Iapichino, 1997; Crisp and Gray, 1984). A recent taxonomic study of 40 curd and leaf traits in Italian cauliflowers and broccolis confirmed that historical eco-geographic groupings and relationships still exist (Massie et al., 1996). Thus, analysis of the genetics thought to underlie these characters may allow a reconstruction of the events leading to the domestication of the curd phenotype. Association of alleles of *BoCAL-a* with curding phenotypes of *B. oleracea* has also been demonstrated through a survey of crop accessions (Smith and King, 2000). Strong correlations exist between specific alleles of *BoCAL-a* and discrete inflorescence morphologies. These complementary lines of evidence suggest that the cauliflower curd arose in southern Italy from a heading Calabrese broccoli via an intermediate Sicilian crop type.

### **REDUCTION IN ALLELE DIVERSITY ARISING FROM SPECIFIC CROP SELECTION**

Diploid brassicas are facultative outbreeders. In *B. oleracea*, a strong sporophytic self incompatibility system operates whereby the rejection of self-pollen on the stigma surface is controlled by a single locus *S*, with up to 60 functional alleles (Ockendon, 1982). Incompatibility genes at this locus control the recognition and rejection of self-pollen prior to sexual fertilisation. This breeding system results in relatively heterogeneous wild and landrace populations, with hybridization limited only by eco-geographic isolation (Massie, 1998). Retention and development of distinct curd types may therefore be largely dependant upon selection by man, as curding plants have a lower 'fitness', based on seed set, than the related broccoli types (Crisp, 1982). Therefore the genetic model for domestication does not require distinct mutational events to have arisen recently in order for development of the curding habit. Rather it relies on bringing together, either by intentional or serendipitous combination, a group of alleles already present in the *B. oleracea* genepool (Smith and King, 2000). This is consistent with the suggestion by Purugganan et al. (2000) that the *Bocal-a* mutation detected by Kempin et al. (1995) may not have originated within the cauliflower genepool.

We have recently established that within a 14cM region on the smallest linkage group of *B. oleracea* (O6), lie two copies of *BoAPI* (*-a* and *-c*), which flank the *S*-locus (Ryder et al., 2001). There is thus the possibility that linkage drag arising from selection for *BoAPI* alleles associated with curd formation may lead to a reduction in the number of *S*-alleles within the cauliflower gene pool.

Information about the distribution of *S* alleles amongst the different crop types of *B. oleracea* has been collated over a number of years (Ockendon, 2000). The distribution amongst accessions representing the crop types cabbage, kales, Brussels sprouts and broccoli is illustrated in Figure 1a. The corresponding distribution amongst cauliflower accessions (Figure 1b) appears to be considerably reduced. Care has to be taken in the interpretation of such data, as the sampling strategy was not designed to provide equal weightings for this comparative purpose. However, the results do indicate that this is a worthwhile approach, and should be taken in to account when there is a requirement to introduce additional *S*-alleles into the cauliflower genepool for breeding purposes. Further work on the distribution of markers over the whole of linkage group O6 will provide greater insights into the constraints on allelic diversity imposed by selection for the cauliflower crop type.

### **CONCLUSIONS**

The selection of unique combinations of alleles at key regulatory loci affecting arrest of floral development appears to have contributed to domestication of the cauliflower phenotype. There is initial evidence for a reduction in allele diversity at two

key loci associated with this crop type. At *BoCAL-a* on linkage group O3 there is a strong selection for a very small number of alleles. The close association of *BoAPI-a* and *BoAPI-c* with the self-incompatibility locus *S* may have contributed to the observed reduced number of *S*-alleles within the modern cauliflower gene-pool. The results indicate that in order to maximise allelic variation in genetic conservation it is important to consider the effect of strong selection for loci which are linked in coupling to genes responsible for determining crop phenotype.

## ACKNOWLEDGEMENTS

Research on genetic and genomic analysis of *Brassica* crops is funded at HRI by the UK Biotechnology and Biological Sciences Research Council and the Department of Environment, Food and Rural Affairs.

## Literature Cited

- Bowman, J.L., Alvarez, J., Weigel, D., Meyerowitz, E.M. and Smyth, D. 1993. Control of flower development in *Arabidopsis thaliana* by *APETALA1* and interacting genes. *Development* 119: 721-743.
- Branca, F. and Iapichino, G. 1997. Some wild and cultivated Brassicaceae exploited in Sicily as vegetables. *Plant Gen. Res. News*. 110: 22-28.
- Crisp, P. 1982. The use of an evolutionary scheme for cauliflowers in the screening of genetic resources. *Euphytica* 31: 725-734.
- Crisp, P. and Angell, S.M. 1985. Genetic control of green curd colour in cauliflower. *Ann. Appl. Biol.* 107: 601-603.
- Crisp, P. and Gray, A.R. 1984. Breeding old and new forms of purple heading broccoli. *Cruciferae News*. 9: 17-18.
- Crisp, P. and Tapsell, C.R. 1993. Cauliflower, *Brassica oleracea* L. p.157-178. In: Kalloo G, Bergh B O (eds) Genetic improvement of vegetable crops. Pergamon Press, Oxford.
- Dickson, M.H. and Lee, C.Y. 1980. Persistent white curd and other curd characters of cauliflower. *J. Amer. Soc. Hort. Sci.* 105: 533.
- Gates, R.R. 1953. Wild cabbages and the effects of cultivation. *Journal of Genetics* 51: 365-72.
- Giles, W.F. 1941. Cauliflower and Broccoli. What they are and where they come from? *J. Royal Hort. Soc.* 66: 265-277.
- Gray, A.R. 1982. Taxonomy and evolution of broccoli (*Brassica oleracea* var. *italica*). *Econ. Bot.* 36: 397-410.
- Irish, V.F. and Sussex, I.M. 1990. Function of the *apetala-1* gene during *Arabidopsis* floral development. *Plant Cell* 2: 741-753.
- Helm, J. 1963. Morphologisch-taxonomische gliederung der kultursippen von *Brassica oleracea*. *Kulturpflanze* 11: 92-210.
- Kempin, S.A., Savidge, B. and Yanofsky, M.F. 1995. Molecular basis of the cauliflower phenotype in *Arabidopsis*. *Science* 267: 522-525.
- Kieffer, M., Fuller, M.P. and Jellings, A.J. 1998. Explaining curd and spear geometry in broccoli, cauliflower and 'romanesco': quantitative variation in activity of primary meristems. *Planta* 206: 34-43.
- Massie, I.H. 1998. Patterns of variation in the Italian landrace cauliflower and broccoli. p.223. PhD thesis, Univ. Lond., London
- Massie, I.H., Astley, D. and King, G.J. 1996. Patterns of genetic diversity and relationships between regional groups and populations of Italian landrace cauliflower and broccoli. (*Brassica oleracea* L. var. *botrytis* L. and var. *italica* Plenck) *Acta Hort.*, 407:45-53.
- Pease, M.S. 1926. Genetic Studies in *Brassica oleracea*. *J. Genet.* 16: 363-385.
- Purugganan, M.D., Lowman, A.C. and Suddith, J. 2000. Variation and selection at the *CAULIFLOWER* floral homeotic gene accompanying the evolution of domesticated *Brassica oleracea*. *Genetics* 155: 855-862.
- Ockendon, D.J. 1982. An S-allele survey of cabbage (*Brassica oleracea* var. *capitata*). *Euphytica* 31:325-331.

- Ockendon, D.J. 2000. The S-allele collection of *Brassica oleracea*. *Acta Hort.* 539:25-30
- Ryder, C.D., Smith, L.B., Teakle, G.R. and King, G.J. 2001. Contrasting genome organisation: Two regions of the *Brassica oleracea* genome compared with collinear regions of the *Arabidopsis thaliana* genome. *Genome* 44: 808-817.
- Smith, L.B. and King, G.J. 2000. The distribution of *BoCAL-a* alleles in *Brassica oleracea* is consistent with a genetic model for curd development and domestication of the Cauliflower. *Molecular Breeding* 6: 603-613.
- Thompson, K.F. 1976. Cabbages Kales etc. *Brassica oleracea* (Cruciferae). p.49-52. In: Simmonds NW (ed) *Evolution of crop plants*. Longman Group Ltd, London.
- Watts, L.E. 1965. Investigations on the inheritance and responses to selection of riceyness in early summer cauliflower. *Euphytica* 15: 111-115.
- Wurr, D.C.E. and Fellows, J.R. 1998. Leaf production and curd initiation of winter cauliflower in response to temperature. *J. Hortic. Sci. Biotech.* 73: 691-697.
- Zeven, A.C. 1996. Sixteenth to eighteenth century depictions of cole crops, (*Brassica oleracea* L.), Turnips (*B. rapa* L. cultivar group vegetable turnip) and Radish (*Raphanus sativus* L.) from Flanders and the present day Netherlands. *Acta Hort.* 407, 29-33.

**Figures**

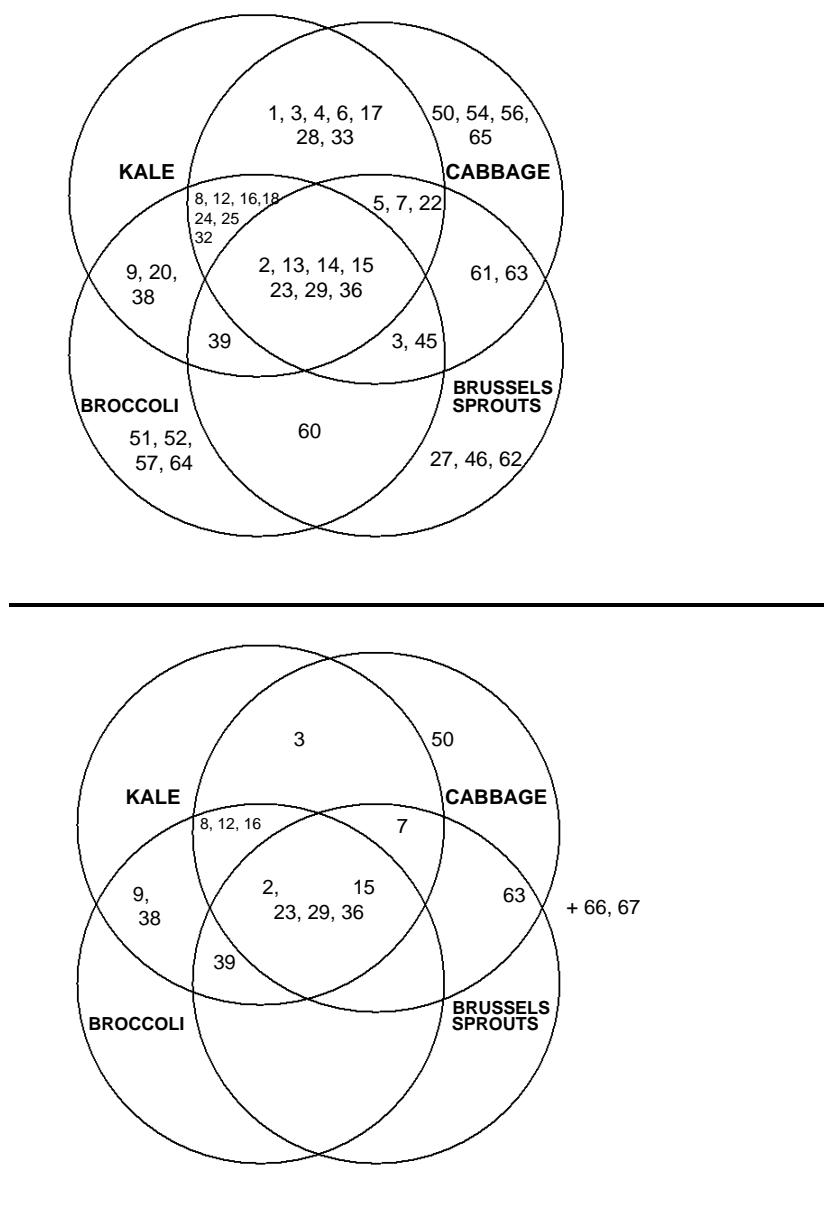


Fig. 1. Distribution of S-alleles within the *Brassica oleracea* genepool. a) distribution amongst accessions of kale, cabbage, broccoli and Brussels sprouts. b) distribution amongst accessions of cauliflower, superimposed on other crop types. Data are collated from Ockendon (2000).